Multichannel Pulse Oximetry: Effectiveness in Reducing HR and SpO$_2$ error due to Motion Artifacts
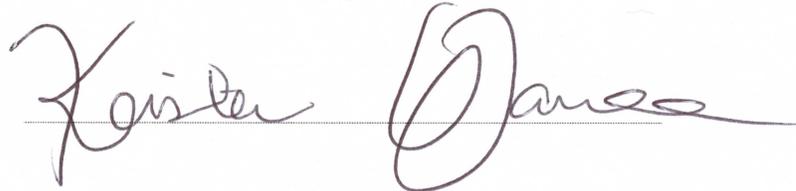
A Thesis

Submitted to the Faculty of

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Master of Science

By

Kristen M Warren

January 21st, 2016

Professor Yitzhak Mendelson, Major Advisor

WPI, Professor of Biomedical Engineering

Professor Ki Chon, Committee Member

University of Connecticut, Department Head of Biomedical Engineering

Professor Patrick Flaherty, Committee Member

UMass-Amherst, Mathematics and Statistics, Assistant Professor

ABSTRACT

Pulse oximetry is used to measure heart rate (HR) and arterial oxygen saturation (SpO_2) from photoplethysmographic (PPG) waveforms. PPG waveforms are highly sensitive to motion artifact (MA), limiting the implementation of pulse oximetry in mobile physiological monitoring using wearable devices. Previous studies have shown that multichannel pulse oximetry can successfully acquire diverse signal information during simple, repetitive motion, thus leading to differences in motion tolerance across channels. In this study, we introduce a multichannel forehead-mounted pulse oximeter and investigate the performance of this novel sensor under a variety of intense motion artifacts. We have developed a multichannel template-matching algorithm that chooses the channel with the least amount of motion artifact to calculate HR and SpO_2 every 2 seconds. We show that for a wide variety of random motion, channels respond differently to motion, and the multichannel estimate outperforms single channel estimates in terms of motion tolerance, signal quality, and HR and SpO_2 error. Based on 31 data sets of PPG waveforms corrupted by random motion, the mean relative HR error was decreased by an average of 5.6 bpm when the multichannel-switching algorithm was compared to the worst performing channel. The percentage of HR measurements with absolute errors ≤ 5 bpm during motion increased by an average of 27.8 % when the multichannel-switching algorithm was compared to the worst performing channel. Similarly, the mean relative SpO_2 error was decreased by an average of 4.3 % during motion when the multichannel-switching algorithm was compared to each individual channel. The percentage of SpO_2 measurements with absolute error ≤ 3 % during motion increased by an average of 40.7 % when the multichannel-switching algorithm was compared to the worst performing channel. Implementation of this multichannel algorithm in a wearable device will decrease dropouts in HR and SpO_2 measurements during motion. Additionally, the differences in motion frequency introduced across channels observed in this study shows precedence for future multichannel-based algorithms that make pulse oximetry measurements more robust during a greater variety of intense motion.

ACKNOWLEDGEMENTS

I express my gratitude to the following people for their contribution and support in my research:

My advisor, Prof. Yitzhak Mendelson, for his guidance and patience over the past two years, and for believing in me and pushing me to finish.

My thesis committee, Prof. Ki Chon and Prof. Patrick Flaherty, for their knowledge, perspective and suggestions.

Corinne Nawn for allowing me to be involved in her MQP project and introducing me to pulse oximetry.

Duy Dao for working closely with me during his time here at WPI and helping me figure out the devices and algorithms needed to complete this project.

Joshua Harvey for being an excellent lab mate over the last year, helping me recruit subjects, and providing valuable insight, discussions, support, and laughter during my days in lab.

My family and friends for believing in me and supporting me in this endeavor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF ABBREVIATIONS

BPF   Band-Pass Filter

DSP   Digital Signal Processing

ECG   Electrocardiogram

FFT   Fast Fourier Transform

Hb   Deoxyhemoglobin

HbO_2 Oxygenated Hemoglobin

HPF   High-Pass Filter

HR   Heart Rate

IR   Infrared

LED   Light Emitting Diode

LPF   Low-Pass Filter

MA   Motion Artifact

MCNL   Multichannel Noise Level

MCPO   Multichannel Pulse Oximeter

PD   Photodetector

PI   Performance Index

PO   Pulse Oximetry/Pulse Oximeter

PPG   Photoplethysmogram / Photoplethysmography

PSD   Power Spectral Density

R   Ratio-of-Ratios

RD   Red

SaO_2 Arterial Oxygen Saturation

SpO$_2$ Estimate of Arterial Oxygen Saturation

SNR   Signal-to-Noise Ratio

SQI   Signal Quality Index

1. INTRODUCTION

Pulse oximetry uses light absorption to measure arterial blood oxygen saturation (SpO_2) and heart rate (HR) from photoplethysmographic (PPG) waveforms. PPG waveforms are sensitive to motion artifacts (MA) that can corrupt the waveform and result in erroneous HR and SpO_2 calculations. The primary cause of MA in pulse oximetry is believed to be due to changes in the light path during sensor movements [1]. Pulse oximetry is widely used in hospitals where motion artifacts are generally less pronounced compared to mobile health applications. Motion artifacts are difficult to filter out since they do not have a predetermined frequency range and their spectral content often overlaps with the frequency band of the PPG waveform. If motion artifact persists long enough and has a frequency in the range of normal HR, the calculated HR and SpO_2 can be highly inaccurate. Clinicians have cited motion artifacts in pulse oximetry as the most common cause of false alarms, loss of signal, and inaccurate readings [2].

The primary approach to reduce motion artifact is the implementation of software-based algorithms that attempt to extract a clean PPG waveform from a motion-corrupted PPG signal. Studies showed that multichannel pulse oximetry is advantageous over single channel pulse oximetry in obtaining diverse signal information during low-motion artifact conditions [3-5]. Furthermore, studies have attempted to better characterize the effects of motion artifact in pulse oximetry, and have shown that intense, aperiodic, random movements generate the most errors [6]. In this study we use a custom made 6 photodetector (PD), forehead-mounted pulse oximeter and investigate the performance of this wearable sensor under a variety of intense motion. We introduce an advanced multichannel-switching algorithm that selects the channel with the least amount of motion artifact to calculate HR and SpO_2 every 2 seconds. We show that for a wide variety of random motion, channels respond differently to motion, and the multichannel estimate outperforms single channel estimates in terms of motion tolerance, signal quality, and HR and SpO_2 errors.

This thesis is divided into the following sections: Chapter 2 discusses the basic principles of pulse oximetry, how HR and SpO_2 values are calculated from raw PPG signals, the two operating modes of pulse oximetry (2.1), the effects of motion artifacts, and the methods currently used to reduce these effects (2.2), an overview of previously designed multichannel pulse oximeters (2.3), and a description our forehead –mounted multichannel pulse oximetry device (2.4).

Chapter 3 outlines the research objectives of this study. Chapter 4 describes the data collection protocol and the type of data collected. Chapter 5 outlines the methods of data analysis used in this thesis including: motion quantification methods (5.1), frequency-based SNR (5.2), HR error metrics (5.3) and SpO_2 error metrics (5.4). The multichannel template-matching algorithm we developed is explained in detail in section 5.1.1, and the analysis of the accompanying accelerometer data is explained in 5.1.2. Chapter 6 presents the results of this including the time-domain response to motion across channels (6.1), the measured accelerometer motion (6.2), the frequency-based SNR values (6.3), the multichannel noise level (MCNL) values during motion and a comparison with the noise detection algorithm previously used in our lab (6.4), the HR errors during motion (6.5), and the SpO_2 errors during motion (6.6). Chapter 7 includes a discussion of the results and implications of our findings, Chapter 8 summarizes the conclusions of this study, and Chapter 9 provides recommendations for future multichannel-based algorithms.

2. BACKGROUND

2.1   Pulse Oximetry and Photoplethysmography

Pulse oximetry is a non-invasive method used to measure arterial blood oxygen saturation levels based on the difference in optical absorption properties between reduced hemoglobin (Hb) and oxyhemoglobin (HbO_2). Figure 1. Optical absorbance spectra of HbO_2 and Hb; HbO_2 absorbs less red (RD) light than Hb, and absorbs more infrared (IR) light than Hb. Two LEDs, typically around 660 nm and 940 nm, are used to shine light into the blood vessels and surrounding tissue.
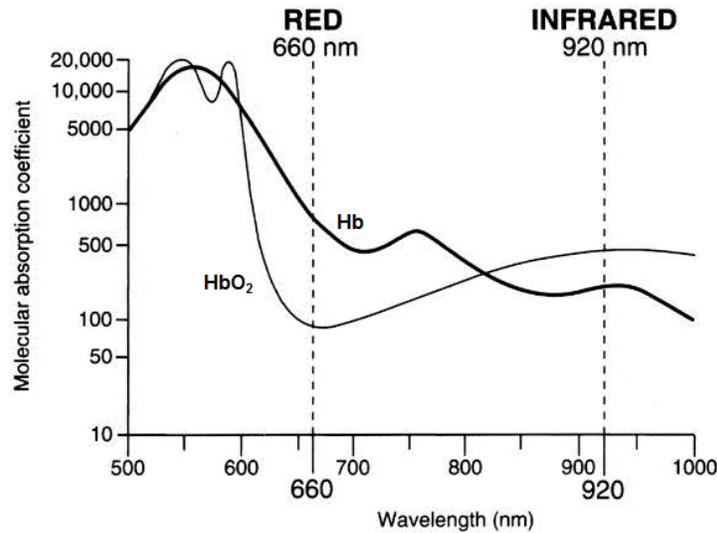
Figure 1. Optical absorbance spectra of HbO_2 and Hb, taken from [7].

The relative amounts of RD and IR light absorbed by the blood correspond to the blood oxygen level. Tissue, bones, and non-pulsatile blood absorb some of the light, generating a DC offset in the PPG signal as shown in Figure 2. Arterial blood pulses with each heartbeat, creating an AC signal superimposed on top of a large DC offset. The signal produced by this absorption process is called a photoplethysmogram (PPG) waveform. Figure 2. Variations in light absorption due to blood and surrounding tissue, illustrating the PPG signal generation.depicts the composition of the PPG signal.
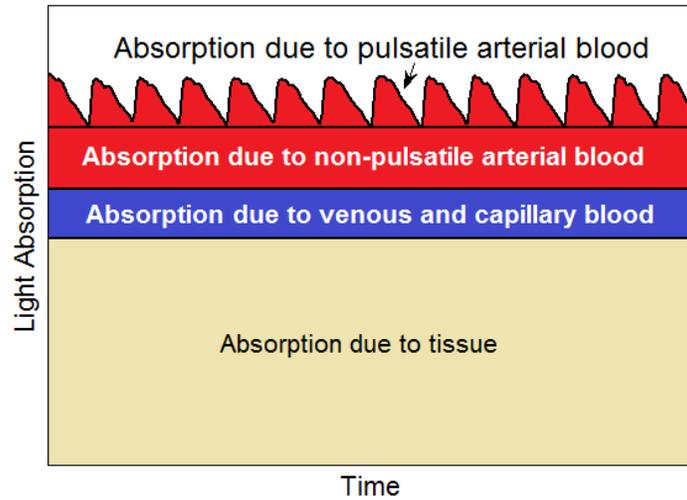
Figure 2. Variations in light absorption due to blood and surrounding tissue, illustrating the PPG signal generation.

The AC signal is a measure of the pulsatile blood, with period corresponding to the cardiac cycle of the heart. A typical AC IR PPG waveform taken from a forehead sensor during rest is shown below in Figure 3.
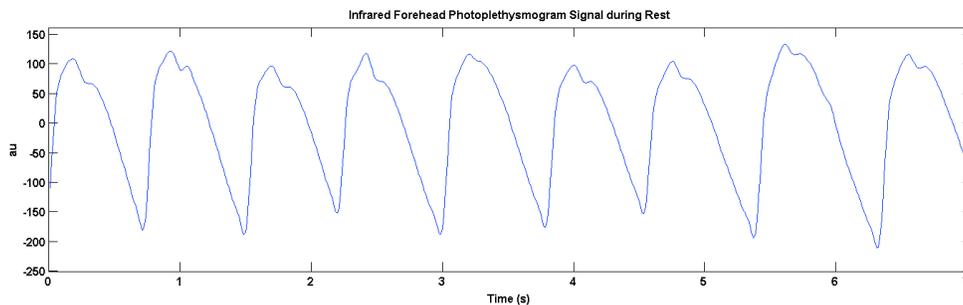
Figure 3. Typical AC PPG signal acquired during rest.

2.1.1 SpO$_2$ MEASUREMENTS

Blood oxygen levels estimated by a pulse oximeter are a function of the ratio of RD to IR light absorption. In order to minimize the effects of physiological and physical differences between patients, including different LED intensities, thickness or density of a tissue region, and color of the skin, the AC amplitude of each signal is normalized by the DC amplitude, creating a "ratio-of-ratios" (**R**) as follows.

4

$$R = \frac{\left(\dfrac{AC_{rd}}{DC_{rd}}\right)}{\left(\dfrac{AC_{ir}}{DC_{ir}}\right)} \tag{1}$$

SpO$_2$ is then clinically calibrated empirically as a linear function of the ratio-of-ratios, and a calibration curve specific to a single device is generated according to Equation 2, where A and B are the coefficients of the best fit linear regression line.

$$SpO_2 = A - B * R \tag{2}$$

Once A and B coefficients are determined for a particular device, a calibration curve is created, allowing all SpO$_2$ measurements to be calculated based on the linear equation or a look-up table. A typical regression line of real SpO$_2$ and **R** values is plotted in Figure 4.
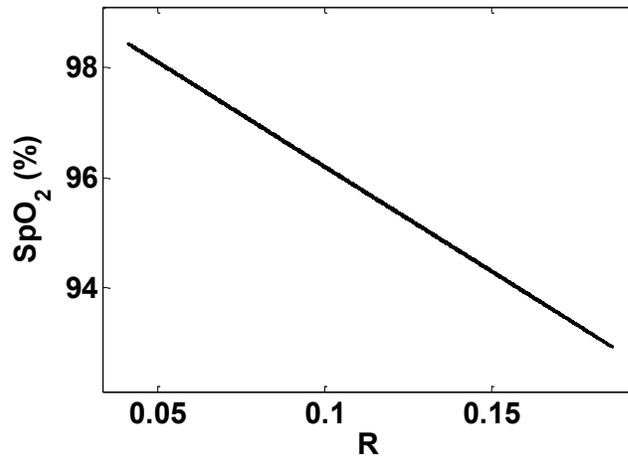
Figure 4. Typical calibration curve of SpO$_2$ based on R values.

To separate the PPG waveform into AC and DC components, a low-pass filter (LPF) and a band-pass filter (BPF) are implemented. A 0.5 Hz LPF is used to extract the DC component, and a 0.5-12 Hz BPF is used to extract the AC component. To smooth out the AC portion of the signal, the RMS value of the AC signal is taken to calculate SpO$_2$. Using the peak-to-peak amplitude of the AC PPG waveform resulted in SpO$_2$ measurements with high variance during rest and sporadic SpO$_2$ measurements with high error. Using the RMS of the AC PPG signal was found to result in more stable SpO$_2$ measurements and overall less measurement error.

2.1.2 HEART RATE MEASUREMENTS

The pulsatile peaks of the PPG signal correspond to cardiac beats, allowing peak detection based algorithms to estimate HR from the PPG waveform. Figure 5 shows a typical ECG signal and the corresponding RR intervals and a typical PPG waveform with the peak-to-peak intervals (PP) lined up with the R waves in the ECG signal.

Figure 5. ECG waveform lined up with PPG waveforms, showing that the RR interval of ECGs correspond to the P-P intervals of PPG.

A peak detection algorithm is used to find the peaks and troughs of the PPG waveform. The peak-to-peak period, T, is then used to measure HR in bpm using Equation 3.
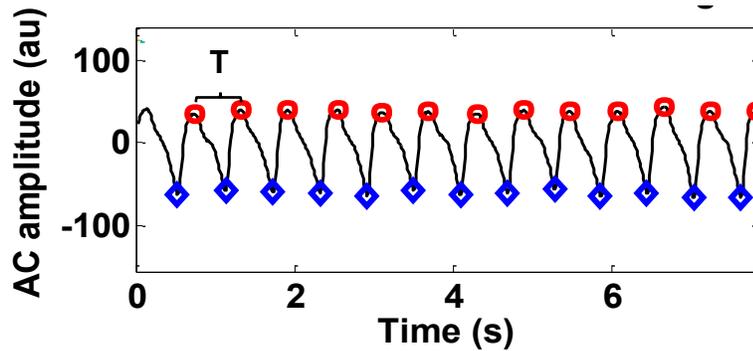
Figure 6. Peak and trough detection of an infrared PPG waveform.

$$HR = \frac{60}{T} \tag{3}$$

6

2.1.3   OPERATING MODES AND MEASUREMENT SITES

Pulse oximeters can operate in two different modes: transmittance and reflectance. Most pulse oximeters operate in transmittance mode; light is transmitted through a part of the body, such as an ear lobe or finger, and a PD on the opposite side from the LEDs detects the transmitted light. In reflectance mode, the PDs are placed adjacent to the LEDs, and the PD subsequently detects the light reflected by the bone. Figure 7 depicts the configuration of a typical transmittance-type pulse oximeter worn on the finger (left), and a reflectance-type pulse oximeter resting on top of the skin (right).

Figure 7. Transmittance mode and reflectance mode configuration for pulse oximetry.

Reflectance mode pulse oximetry allows for more sensor placement sites, including the forehead, which is less susceptible to motion artifacts and easier to wear during physical activity than a finger sensor.

2.1.4   LIMITATIONS OF PULSE OXIMETRY

Pulse oximetry is highly accepted and accurate in clinical settings when a patient is at rest. During motion, low perfusion, and in the presence of electromagnetic interference, the PPG waveform is compromised, resulting in lower accuracy and more frequent error. Means have been taken to limit the effects of low perfusion and electromagnetic interference, but for active patients, motion artifacts are the primary limiting factor to using pulse oximetry in mobile health applications.

2.2   MOTION ARTIFACTS

The primary causes of MA include changes in the light path due to sensor movement in relation to the skin and introduction of venous blood pulsation. As shown in Figure 8, when the sensor is displaced in the plane that the sensor touches the skin, the light path changes and the light then travels differently to the photodiode. This motion of the sensor relative to the skin causes a

change in the optical coupling between the tissue and the sensor, which can cause a change in the PPG signal [9].
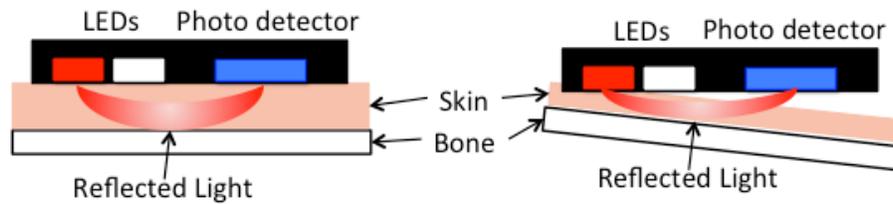
Figure 8. Light path changes during sensor-skin displacement due to motion (adapted from [8]).

Additionally, pulse oximetry assumes that the only pulsatile component is due to arterial blood pulsations. Arteriovenous (AV) shunting causes venous blood to change with arterial pulsing [3]. The presence of venous pulsation becomes more prevalent during motion, yielding false low SpO_2 values due to the lower oxygen saturation of venous blood.

2.2.1 EFFECTS OF MOTION ARTIFACT

Motion artifacts can affect the amplitude and frequency of the PPG waveform, making the peaks corresponding to the cardiac cycle unidentifiable. Particularly, if the motion artifact signal lasts long enough and has a frequency in the range of normal heart rate, the peak detection algorithm will measure motion frequency as HR and use incorrect AC amplitudes to calculate SpO_2. Figure 9 below shows an example AC PPG waveform that starts out clean and is then corrupted by motion artifact. Once motion artifact is present in the waveform, it becomes difficult to separate the physiological signal from the effects of motion.

Figure 9. Clean PPG waveform corrupted by motion artifacts

It is clear that it would be difficult to extract the specified amplitude or frequency that corresponds to the true physiological signal from the motion-corrupted portion of the PPG waveform shown in Figure 9. The HR calculation algorithm would rely on the detection of false peaks, and the SpO_2 algorithm would incorrectly use the amplitude of the motion noise, resulting in erroneous HR and SpO_2 measurements.

2.2.2 CURRENT APPROACHES TO LIMIT MOTION ARTIFACT

There are a number of software-based methods currently being used to limit the effects of motion artifacts on the PPG signal. The primary approach to reduce the effects of motion artifact is to try and extract a clean PPG waveform from the motion-corrupted PPG signal [10-25]. Fixed frequency filtering is helpful, but not very effective when MA has no predetermined frequency band. Baseline subtraction, use of frequency banks, moving average filtering, and removal of corrupted signal segments have shown improvement against MA in some cases, but are not robust against motion which may have varied dynamics [26-29].

Numerous studies have investigated the use of adaptive noise cancellation (ANC) to selectively filter out MA based on a specified reference signal. Reference signals used include: on-board accelerometers [10-13], a reference signal synthesized from the motion corrupted PPG signal [14-17], and a reference signal measured by an adjacent photoelectric device [18,19]. Correlations have been found between the distorted PPG data and the accelerometer data after a time delay [10], and the adaptive filters remove some of the noise that was correlated to the corrupted PPG, but the reconstruction is not perfect. Similarly, sensor motion measured using self-mixing interferometry more directly measures the light changes with respect to sensor movement and does reduce some motion artifacts, but does not fully suppress them [19].

Separate reference signals are not always an accurate representation of the signal corruption, which can lead to unintentional filtering of a portion of the relevant PPG waveform.

Alternate algorithms have been developed to extract the clean PPG waveform from the motion-corrupted signal based on fundamental components of the PPG signal. These methods include: principle component analysis (PCA) [20], independent component analysis (ICA) [21-23], and singular spectral analysis (SSA) [24]. Adaptive filtering with accelerometer data has been tested against the use of Independent Component Analysis (ICA). ICA outperforms adaptive filtering in terms of artifact removal [30]. ICA, however, has permutation and scale ambiguities, not always preserving the correct PPG. The use of a temporally constrained ICA in combination with an adaptive filter works better than either method individually, and can successfully output clean PPG signals. The ICA is performed and then provided as a reference for the adaptive filter, thus extracting the clean PPG signal, which has attributes of the ICA result [23]. Work in this lab has been done on the use of Iterative Motion Artifact Removal (IMAR) using Singular Spectral Analysis (SSA) and compared with time-domain ICA. The IMAR method outperforms ICA in terms of HR and SpO_2 errors [24]. Most recently, algorithms based on filtering out the motion frequency as calculated from the accelerometer spectra have been useful in separating motion signal from PPG signal [31-33]. These algorithms have been proven effective during motion, but are currently designed to use one signal to extract information from, and when the single channel is too corrupted to reconstruct, or when the motion frequency overlaps with the HR frequency preventing motion from being filtered out, HR and SpO_2 information may be lost, leading to dropouts during monitoring.

2.3   MULTICHANNEL PULSE OXIMETERS

Multichannel devices have been used to improve HR and SpO_2 measurements by capturing multiple PPG waveforms simultaneously. An PPG sensor placed on the sternum was developed that uses PCA in the frequency domain to find the most likely SpO_2 estimation [4]. This PPG sensor has a total of 16 RD LEDs, 16 IR LEDs, and 8 PDs on one board, as shown in Figure 10. In the study, the investigators only used the 8 pairs of LEDs and 4 PDs circled in black in Figure 10.
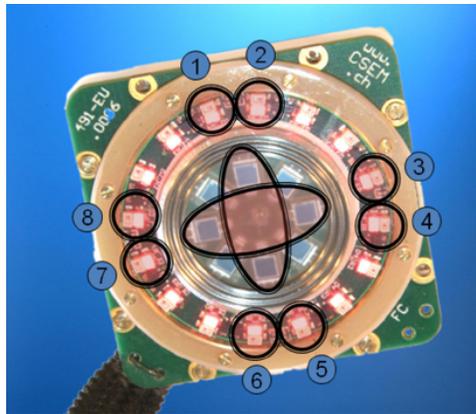
Figure 10. The 16-channel sternum sensor used by [4]. Only the 8 channels circled in black were used in their analysis.

These investigators found that "the spatial diversity of the channels considerably increases the available photoplethysmographic information…however this requires a more sophisticated control and signal processing" [4]. Frequency information from an external BIOPAC ECG sensor and an on-board accelerometer was used to design a specified filtering window for each measurement. Spatio-temporal PCA was used to determine the most likely **R** value between channels, which was then used to calculate SpO_2. Both a BIOPAC finger clip sensor and a NELLCOR finger clip sensor were used for reference SpO_2 measurements. Ten subjects were subjected to hypoxia using a re-breathing system to slowly lower SpO_2 values from 100 % down to 70 % and back up to 100 %. They found that multichannel SpO_2 estimates were more robust than single channel SpO_2 estimates, especially when considering the poor perfusion seen in PPG waveforms taken from the sternum.

A 3-channel reflectance earpiece PPG sensor was developed and tested during standing, sitting, and walking [5]. This sensor featured 2 IR LEDs and 3 PDs in a device that hooks around the ear, shown in Figure 11. This study only analyzed multichannel HR measurements, so only IR LEDs were used.
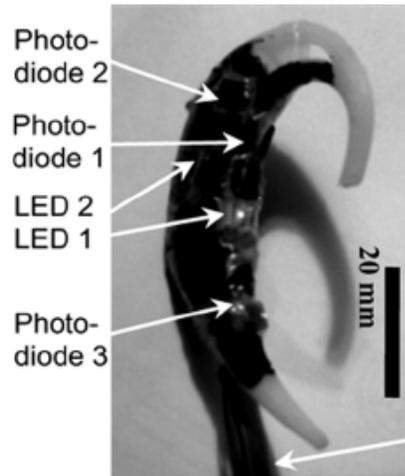
Figure 11. The 3PD ear sensor used in [5].

Ten subjects were tested wearing the ear sensor and a Nellcor PPG sensor for reference HR measurements. Subjects were instructed to perform normal daily activities consisting of standing, sitting, and walking, and were then instructed to walk on a treadmill at a speed starting at 4km/hr, increasing to 7 km/hr, and then decreasing back down to 4 km/hr. To measure the motion frequency, the motion-induced current was measured by reading the output from the PDs when the LEDs were turned off. An adaptive notch filter was implemented at the measured motion frequency with a bandwidth of 0.4 Hz to reduce noise contribution. These investigators found that motion-induced current was channel-specific, and that the channel with the highest power around the HR frequency varied between experimental runs.

These studies showed that multichannel pulse oximetry is advantageous over single channel pulse oximetry in obtaining diverse signal information during low-motion and periodic motion artifact conditions. However, studies have attempted to better characterize the effects of motion artifact in pulse oximetry, and have shown that intense, aperiodic, random movements generate the most errors [6]. Previously, we have shown that in a 6-channel prototype reflectance-based forehead pulse oximeter, during short up-down, left-right, and circular head motion, channels responded differently to motion [3]. These types of movements are periodic and are not very intense in amplitude. Users looking to wear a portable pulse oximetry sensor would not be experiencing only simple, low-amplitude movements, but a variety of different aperiodic movements. Examining how multichannel pulse oximetry responds to a wider variety of motion

12

would help to further assess the benefit of this unique design, and assist in developing future signal extraction algorithms that make use of multichannel waveforms to improve motion tolerance in pulse oximetry.

2.4   FOREHEAD MULTICHANNEL PULSE OXIMETER

We have developed a custom reflectance-type, forehead-mounted, multichannel pulse oximeter (MCPO) shown in Figure 10. Six PDs are positioned concentrically around two pairs of Red (660 nm) and IR (940 nm) light emitting diodes (LEDs) at an equidistant separation distance of 10 mm [34]. An opaque ring was incorporated to minimize direct light shunting between the LEDs and PDs. The RD LEDs are shown in red in Figure 10 below, the IR LEDs are shown adjacent to the RD LEDs, and channels 1 through 6 are labeled in white, starting at the top of the sensor and incrementing clockwise.

Figure 12. 6-PD forehead pulse oximeter.

The sensor and battery are enclosed in a plastic casing. The sensor housing is attached to an elastic band worn as a headband, allowing the sensor to rest comfortably on the forehead. The sensor is also equipped with an on-board tri-axial accelerometer (Acc). When the sensor is placed on the forehead, the x-direction of the Acc corresponds to motion perpendicular to the transverse plane, the y-direction perpendicular to sagittal plane, and the z-direction perpendicular to the coronal plane. Previous data have shown that variations in sensor position and vasculature heterogeneity of the underlying tissue can cause measurement errors, as well as light diffusion by the subcutaneous tissues predominantly in the direction perpendicular to the emitting surface of the LEDs [3]. Different motions change the optical coupling between the sensor and the skin, yielding 6 independent channels with slightly different motion-corrupted PPG waveform characteristics.

3. RESEARCH OBJECTIVES

The goal of this research was to investigate the feasibility of using multichannel pulse oximetry to estimate HR and SpO_2 during random motion of varying amplitudes and frequencies.

Objective 1: To collect random motion data from the prototype pulse oximetry sensor with a variety of frequencies and amplitudes.

Objective 2: Validate differences in channel response to motion.

Objective 3: Implement a channel-multiplexing algorithm to choose the cleanest channel from which to calculate HR and SpO_2 estimates.

Objective 4: Determine if multichannel pulse is significantly better than using single channel pulse oximetry during motion in terms of HR and SpO_2 errors.

4. DATA COLLECTION

Data were collected from 15 healthy volunteers between the ages of 22 and 32. Worcester Polytechnic Institute IRB approved the study protocol and informed consent was required by all subjects prior to data recording.
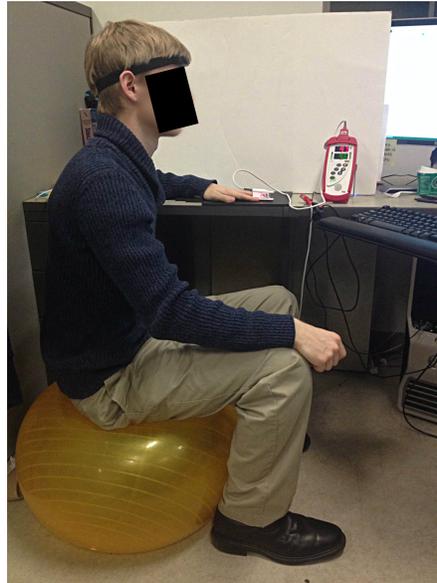
Figure 13. Experimental setup for generating random motion.

Subjects were instructed to bounce on an exercise ball with varying amounts of movement while wearing the 6PD MCPO forehead sensor and a reference Masimo-57 Radical (Masimo SET[®], Masimo Corporation, CA, USA) finger type transmittance pulse oximeter that was kept motionless by resting the left hand on a table, as shown in Figure 2. The type of movement each subject performed was free to interpretation by each person, resulting in different types of motion for each data set. Furthermore, by asking each subject to bounce on an exercise ball, each data set had increasing and decreasing HR, allowing us to track changes in HR. And, since no specified motion, walking, or running speed was given to each subject, the frequencies of motion will vary from subject to subject, yielding a variety of motion corruption data. Each subject was asked to alternate between 3 minutes of rest and 5 minutes of bouncing on the exercise ball for a total of 19 minutes. Six pairs of PPG waveforms corrupted by random motion artifacts were obtained from the forehead-mounted MCPO sensor. PPG waveforms from the MCPO were sampled at 80 Hz. Reference HR and SpO_2 measurements were obtained from the Masimo pulse

oximeter every 2 seconds. All data were transferred to a PC and processed offline with MATLAB. The data were aligned by removing the first few measurements from either the Masimo estimates or the MCPO measurements until the MCPO HR data followed the Masimo HR measurements during rest. It was assumed that the SpO_2 measurements were aligned when the HR measurements were aligned. The six raw IR PPG waveforms, the six raw RD PPG waveforms, the on-board tri-axial accelerometer signals, and the Masimo reference data were saved to MATLAB for further processing.

5. DATA ANALYSIS

The particular motion protocol we followed was meant to test the response of multichannel signals to a variety of non-uniform motion. It was important to quantify the level of motion present in the PPG waveforms, both across different data sets and during data collection across channels to determine which channel was the least affected by motion artifacts. To represent overall motion, RMS values from the on-board tri-axial accelerometer data were calculated for each data set. Furthermore, because HR calculations are highly dependent on the frequency content of the infrared PPG waveform, and the arterial oxygen measurements are highly dependent on the amplitudes of the PPG waveforms, to determine how motion affects the multichannel PPG waveforms, frequency-domain signal-to-noise-ratios (SNR) were calculated for each data set and channel.

A measure of signal quality had to be implemented as a metric to choose between channels. Previous noise detection algorithms have implemented template matching to measure signal quality in terms of PPG morphology. To quantify instantaneous noise level for our multichannel pulse oximeter, a multichannel template-matching algorithm was developed that matches beats in a specified window to an average template representative of a clean PPG morphology. The output from this algorithm is a number between 0 and 1, representing the amount of noise in the signal at a specified time point. A MCNL of 0 means the signal has no noise and an MCNL of 1 means that the PPG waveform is completely corrupted by motion noise. This resulting multichannel noise level (MCNL) will be measured during rest and during motion across all data sets and channels.

Acceptable HR and SpO_2 measurements during motion were taken from Masimo's specifications during motion given that they considered the leading "motion-tolerant" pulse oximeter on the market and, therefore, were used as reference measurements in this study. Masimo specifications claim a $\pm$ 5 bpm HR error during motion and a $\pm$ 3 % SpO_2 error during motion. HR and SpO_2 error measurements are quantified according to three different parameters: performance index, accuracy and accuracy, defined as the percent of measurements within the above specified tolerances during motion, the mean of the absolute relative error, and the standard deviation of the absolute relative error, respectively. The performance index, accuracy, and precision were

compared between each individual channel and the channel-multiplexing estimates. These parameters are explained in greater detail in the following sections.

5.1   MOTION QUANTIFICATION

5.1.1 MULTICHANNEL TEMPLATE MATCHING

To quantify the noise level in each channel at a particular time point, we developed a multichannel template-matching algorithm. An algorithm has been developed based on creating a template from single channel IR PPG waveforms using an average of beats over a specified window [35]. In our algorithm, six IR PPG waveforms are used to create a template over a single time period, allowing more robust template formation. First, the raw PPG waveforms were digitally filtered with a 12th order, 0.5 Hz cut-off, Butterworth low-pass filter and a 12th order, 0.5 to 12 Hz Butterworth band-pass filter to separate the raw PPG signals into their DC and AC components, respectively. The six IR AC components served as inputs to the multichannel template-matching algorithm.

a) Template Formation

Peak-trough detection was performed on each of the six IR AC components over a 12 s window. A 12 s window was chosen to ensure that an accurate HR could be calculated from a single channel, and that an accurate template could be formed from the data in the window. From the beats in the window, an average peak-peak period was calculated across all channels and was defined by a variable L. Peak-peak period values that are outside of $L \pm 0.2$ s were removed to include only clean beats in the beat period calculation. The average peak-peak period is updated by taking the average of the remaining peak-peak period values. Individual beats were segmented at the indices of each trough. Beats across all channels for the entire 12 s of data were re-sampled to be the same length (L) and normalized to have maximum amplitude of one and zero minimum amplitude. A temporary template is created from the average of all individual beats in the 12 s window. The correlation between this template and all beats in the current window was computed, and if the correlation of a beat in the current window with this template was less than 0.95, the beat was removed from further calculations. Analysis of the correlation between beats in a window during rest showed that the lowest correlation during rest was 0.9572, so a 0.95 correlation was assumed to be the cutoff correlation for clean data verses noise-

corrupted data. If the number of beats removed were greater than 1/3 of the total number of beats in the current window, the template from the previous window was used. If less than 1/3 of the beats are removed, taking the average of the remaining "good" beats in the window forms a new template. The average correlation of all individual beats in the window is again calculated with the latest template. If the average correlation across all beats in the window with this template is less than 0.98, the template from the previous window is used. If the average correlation across all beats in the window is greater than 0.98, the template from the previous step is maintained.

b) Multichannel Noise Calculation

With the final formation of the template for a given window, the template matching algorithm systematically overlays beats from each individual channel, cuts off each beat at L/2 in order to capture only the systolic morphology of the PPG beats, which are more indicative of clean signals, and normalizes each beat such that the minimum amplitude is 0 and the maximum amplitude is 1. If the noise is such that there are no beats detected in a given channel, the multichannel noise level (MCNL) is set to 1. Otherwise, the correlation between the template and each beat in a given channel is calculated, and the overall MCNL is calculated as (1-C), where C is the average correlation between beats in a single channel and the template. The specifics of the overall process are described in Figure 14.

TEMPLATE FORMATION

MULTICHANNEL NOISE DETECTION
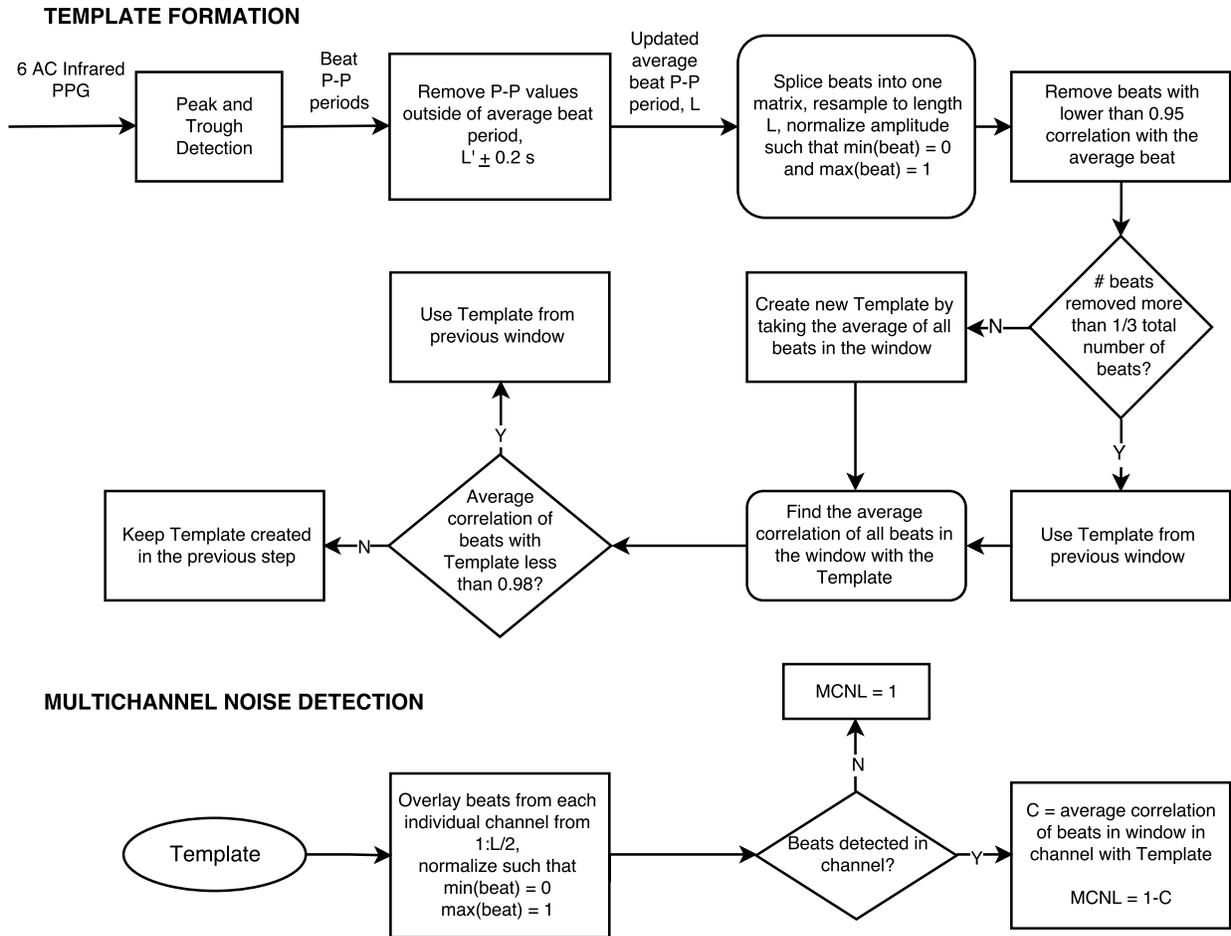
Figure 14. Processing of the data with the multichannel template-matching algorithm to obtain noise level (MCNL) for each channel.

Figure 15 shows sample beats overlaid from all 6 channels during rest and motion. Beats taken from clean data have high correlation (C) with the average beat template, and are similar across all channels. On the contrary, beats corrupted with motion artifact have low correlation with the average "beat". In this case, the template from the previous window is used.
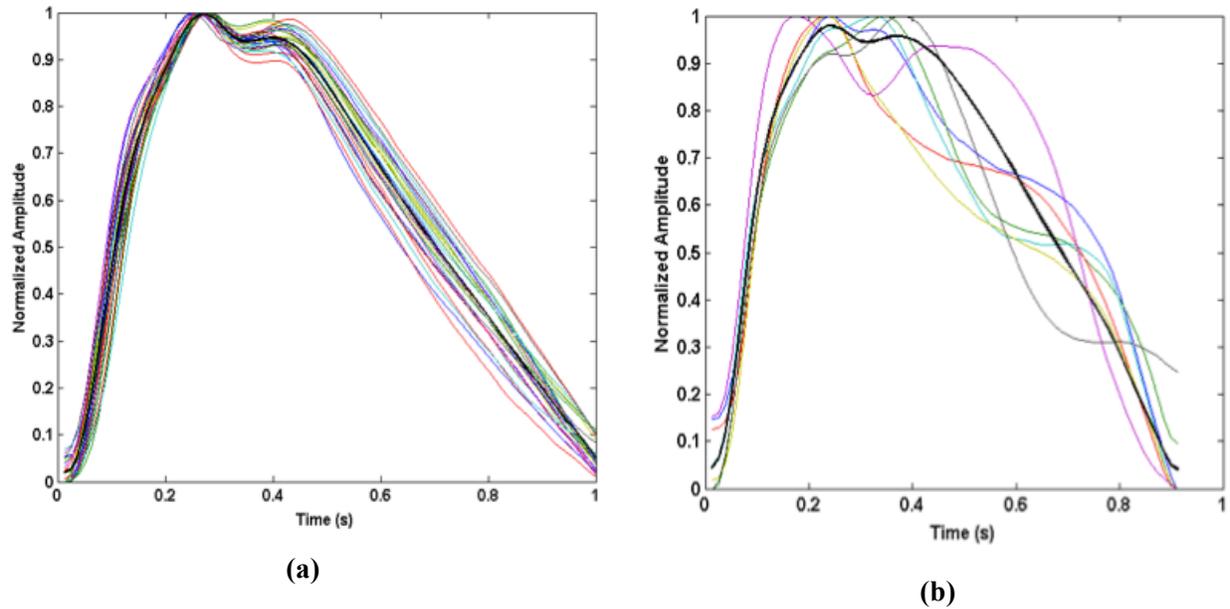
Figure 15. Beat detection and overlay for a multichannel device during rest (a) and motion (b).

Figure 15 shows beats present during rest on the right and the average beat used as the template plotted in black. During motion, only the beats with a high correlation with the template are shown on the right; the template from the previous window is shown in black. Also during motion, some channels have higher correlation with the template than other channels, yielding different MCNL levels for each channel. Figure 16 shows beats overlaid in the window separated by channel during rest and motion. During rest, beats across all channels are have very similar morphology, are highly correlated with one another, and are highly correlated with the template shown in Figure 16. During motion, some channels have beats that remain highly correlated with the template, while some channels are highly corrupted with motion artifact. Therefore, the average correlation coefficient, and MCNL, will differ between channels, allowing the "best" channel to be chosen. The multichannel HR estimate was calculated by dividing 60 by the average peak-to-peak period of the individual waveforms in the 12 s window only from the channel with the lowest MCNL, in accordance with Equation 3.
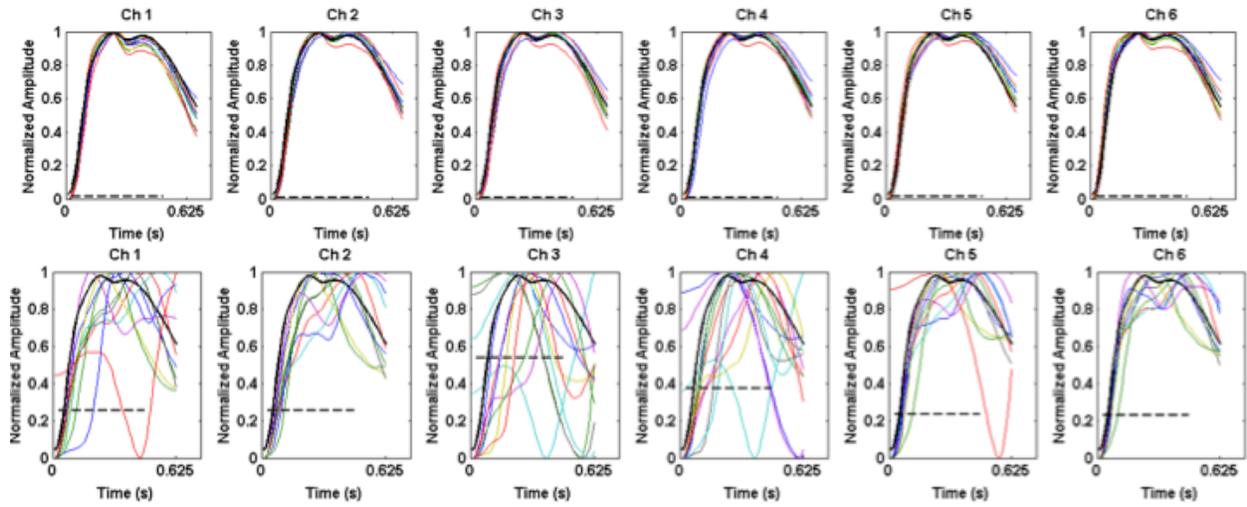
Figure 16. Beats in window overlaid from each individual channel during rest (top) and motion (bottom).

Because HR measurements are more dependent upon the frequency content of the PPG waveform and the SpO_2 measurements are more dependent on the amplitude of the PPG waveforms, the channel chosen to calculate HR was not necessarily the same channel that was chosen to calculate SpO_2. To choose the best channel for calculating SpO_2, a weighted "noise level" was added based on the scaling factor used when the individual PPG waveforms were normalized to maximum amplitude of 1. The correlation of the beat with the template (C) was multiplied by the percent amplitude difference from the mean amplitude in the window. First, the average amplitude of each PPG waveform was calculated as A. This average amplitude was then divided by the maximum value of each individual waveform in the window. This relative (scaled) amplitude, a, was calculated for each waveform i in the 12-second window. Then, the relative amplitude of each waveform was either multiplied or divided by the correlation number to obtain an updated, amplitude-weighted correlation number, C'. The calculations for this amplitude-weighted correlation are denoted in Equations 4-6.

$$A = \text{mean(maximum value of each waveform in 12 s window for a single channel)} \qquad \textbf{(4)}$$

$$a(i) = \frac{A}{\max(\text{waveform}(i))} \qquad \textbf{(5)}$$

$$C'(i) = \begin{cases} \text{if}(a(i)<1) & \dfrac{C(i)}{a(i)} \\ \text{else if } (a(i)\geq1) & C(i) * a(i) \end{cases} \tag{6}$$

The channel chosen to calculate SpO_2 is still the minimum of (1-C), but with the updated amplitude-weighted correlation number C'. The SpO_2 measurements from the channel with the lowest amplitude-weighted MCNL value were calculated by using the RMS of the AC/DC ratio for IR and the RMS of the AC/DC ratio for RD. This RMS-based ratio was used in the formula, $125 - 38 * \mathbf{R}$. These A and B coefficient numbers were based on preliminary calibration data, but are not the final calibration numbers for our sensor.

5.1.2 ACCELEROMETER AMPLITUDE

The accelerometer signals measured during motion provide a way to quantify the amount of motion that was introduced into any particular data set. The RMS value of all 3 axes of the on-board accelerometer was used to quantify motion, according to Equation 7.

$$\mathbf{Accel_{RMS}} = \sqrt{\frac{1}{3}((\mathbf{Accel_X})^2 + (\mathbf{Accel_Y})^2 + (\mathbf{Accel_Z})^2)} \tag{7}$$

Although these calculations do not translate directly into the level of noise introduced into the corrupted PPG waveforms during motion, they help to distinguish between high motion and low motion data sets based on the corresponding RMS values.

5.2   FREQUENCY DOMAIN SNR

Frequency domain SNR represents the power of the PPG signal inside of the HR frequency band relative to the power outside the HR frequency band, which corresponds to motion induced frequencies. First, we compute the power spectral density (PSD) of the entire waveform during motion, and then calculate the mean heart rate (MHR) during motion according to the reference readings obtained from the Masimo finger sensor. The first harmonic of the MHR is subtracted before integrating. We integrate the PSD inside the ± 0.2 Hz frequency band around the MHR

during motion and divide it by the integration of the PSD outside of the ± 0.2 Hz frequency band of the MHR, from 0.67 Hz to 3 Hz. The lower limit of 0.67 was chosen to eliminate the respiratory rate frequency in the calculation. An upper limit of 3 Hz was chosen because 3 Hz corresponds to a 180 bpm heart rate, which was deemed the upper limit for HR calculations. Furthermore, the frequencies introduced by the experimental protocol did not exceed 3 Hz.

$$SNR_{frequency} = 10\log_{10}\left(\frac{\int_{MHR*0.8Hz}^{MHR*1.2Hz} PSD_{PPG}}{\int_{0.67Hz}^{MHR*0.8Hz} PSD_{PPG} + \int_{MHR*1.2Hz}^{3Hz} PSD_{PPG}}\right)[dB] \qquad (8)$$

5.3   HEART RATE ERROR METRICS

HR error was defined as the absolute error between the HR calculated by the multichannel device and the HR given by the Masimo-57 reference sensor, as shown in Equation 9. Absolute error was chosen over a relative, percent error so that a 3 bpm error on a 90 bpm heart rate would be the same as a 3 bpm error for a 140 bpm heart rate.

$$Err_{HR} = \left|HR_{MCPO}\text{-}HR_{Masimo}\right| [bpm] \qquad (9)$$

This HR error was analyzed during motion using three parameters: accuracy, precision, and performance index. Performance Index is an indication of the number of measurements taken that are within a specified tolerance. The Masimo-57 sensor claims HR error tolerance of ± 5 bpm during motion; the upper limit used for performance index calculations matches the Masimo specifications during motion. Performance Index was chosen as an error metric because a higher PI from a channel or sensor indicates that that sensor or channel will experience less dropouts in readings and less overall false HR measurements. Accuracy was chosen as an error metric because accuracy is the degree to which HR measurements match the true, or accepted, values. High degrees of accuracy are important for medical device specifications, as clinical accuracy is of significant importance in medical device instrumentation. Precision refers to the exactness of measurements, or repeatability. Precision was chosen as the third error metric as it is a measure of the ability of a medical device to make the same measurements over time of the same true HR.

24

5.3.1 PERFORMANCE INDEX

Performance index (PI) was defined as the percentage of the measurements that have absolute relative HR errors lower than 5 bpm during motion. The higher in performance index a channel is, the more "correct" measurements were taken during motion. PI is defined in Equation 10.

$$PI = \frac{\text{\# measurements } (\text{Err}_{HR} \leq 5 \text{ bpm})}{\text{total \# of measurements}} \, [\%] \tag{10}$$

5.3.2 ACCURACY

Accuracy was defined as the offset that a HR measurement has in relation to the reference device. In this thesis, we assume that the accuracy is the mean absolute relative HR error during motion in relation to the measurements read by the Masimo-57 finger sensor. Accordingly, accuracy is defined in Equation 11.

$$\text{Accuracy} = \text{AVERAGE}(\text{Err}_{HR}) \, [\text{bpm}] \tag{11}$$

5.3.3 PRECISION

Precision was defined as the ability of a device to make consistent measurements, the spread of measurements taken around one true value, or the repeatability to measure a single true value multiple times. Precision is measured as the standard deviation of the absolute relative error taken in relation to the reference Masimo-57 oximeter. Accordingly, precision was defined in Equation 12.

$$\text{Precision} = \text{STD}(\text{Err}_{HR}) \, [\text{bpm}] \tag{12}$$

5.4 SpO$_2$ ERROR METRICS

Given that SpO$_2$ measurements depend on the clinical calibration curves for each individual device, and that subjects were not subjected to hypoxia in this study, SpO$_2$ measurements from the multichannel device were compared against a constant blood oxygen level of 98 %. Similar to HR error, SpO$_2$ error was defined as the absolute error against a constant oxygen saturation of 98 %, as shown in Equation 13; absolute error was chosen over relative error so that the error calculation would not be dependent on the oxygen saturation level.

$$\text{Err}_{SpO2} = \left| \text{SpO}_{2\,MCPO} - 98 \right| \, [\%] \tag{13}$$

Similar to HR error analysis, Performance Index was chosen as an error metric because a higher PI from a channel or sensor indicates that that sensor or channel will experience less dropouts in readings and less overall false SpO_2 measurements. Accuracy was chosen as an error metric because accuracy is the degree to which SpO_2 measurements match the true, or accepted, values. High degrees of accuracy are important for medical device specifications, as clinical accuracy is of significant importance in medical device instrumentation. Precision refers to the exactness of measurements, or repeatability. Precision was chosen as the third error metric as it is a measure of the ability of a medical device to make the same measurements over time of the same true SpO_2 level.

5.4.1 PERFORMANCE INDEX

Performance index (PI) was defined as the percentage of the measurements that have SpO_2 errors lower than 3 % during motion. The higher percentage a channel or device has in PI, the more "correct" measurements are taken during motion. Hence, PI was defined according to Equation 14.

$$PI = \frac{\# \text{ measurements } (Err_{SpO2} \leq 3\ \%)}{\text{total } \# \text{ of measurements}} [\%] \tag{14}$$

5.4.2 ACCURACY

Accuracy was defined as the offset that a SpO_2 measurement has in relation to the reference device. In this thesis, we defined accuracy as the mean absolute relative SpO_2 error during motion in relation to a SpO_2 of 98%. Accuracy was defined according to Equation 15.

$$\text{Accuracy} = \text{AVERAGE}(Err_{SpO2}) [\%] \tag{15}$$

5.4.3 PRECISION

Precision was defined as the ability for a device to make consistent measurements, the spread of measurements taken around one true value, or the repeatability to measure a true value multiple times. In this thesis, precision was determined based on the standard deviation of the absolute relative error taken in relation a SpO_2 of 98 %. Precision is defined in Equation 16.

$$\text{Precision} = \text{STD}(\text{Err}_{\text{SpO2}}) \ [\%] \tag{16}$$

6. RESULTS

6.1 TIME-DOMAIN PPG WAVEFORM DIFFERENCES DURING MOTION

Though the range of differences in signal corruption varies from data set to data set, for all 31 data sets we collected, the IR PPG waveforms were similar during rest across all channels, and differed in the levels of signal corruption between channels during motion. Figure 17 shows typical IR PPG waveforms recorded during rest and motion from data set 10.

(a) (b)
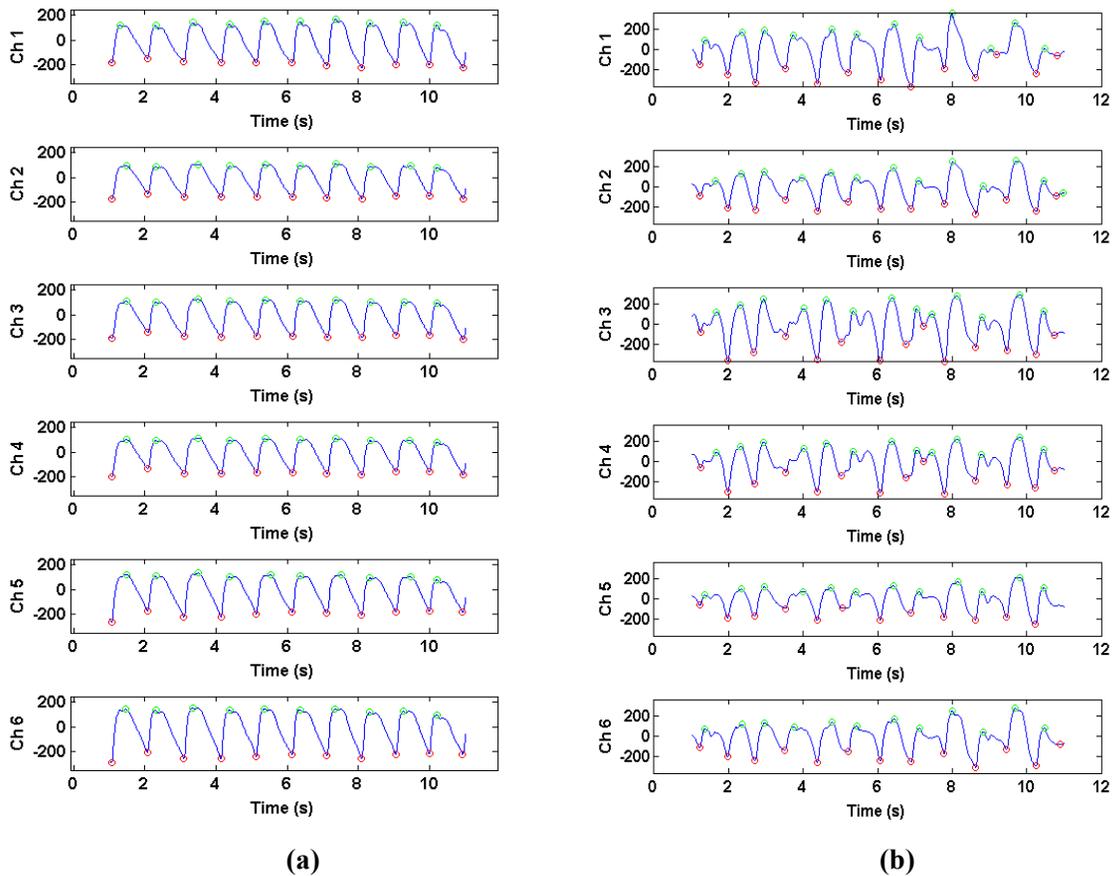
Figure 17. Typical differences in IR PPG waveforms from all six channels recorded during rest (a) and motion (b) in a 12-second window for data set 10.

It is clear from Figure 17b that a channel with greater motion corruption, such as channels 3 and 4, will have worse peak-detections than channels with less motion corruption, as seen in channels 1 and 6. Therefore, the HR estimations from these channels will be better during motion.

6.2 ACCELEROMETER AMPLITUDES DURING MOTION

The RMS amplitudes of the on-board tri-axial accelerometers provide an indirect measure of how much noise was introduced into the corrupted PPG waveforms in each data set, as summarized in Figure 18.
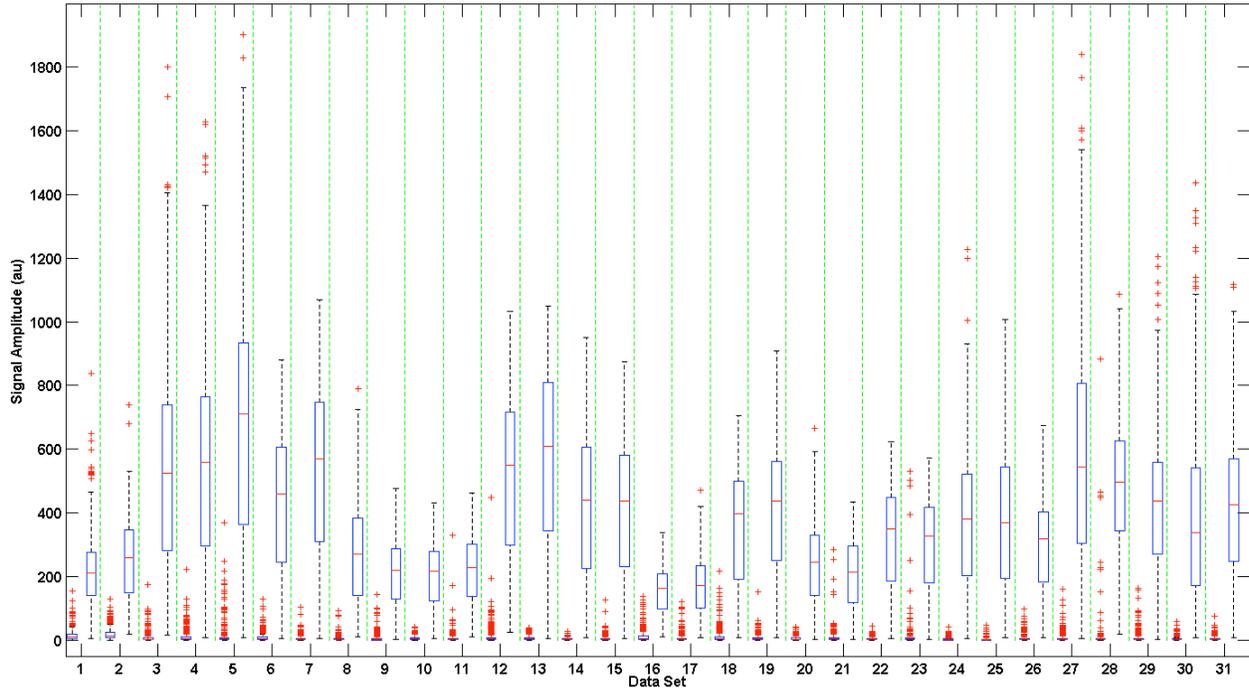
Figure 18. Box and Whisker plot of RMS accelerometer amplitudes across all data sets during rest (left side of column) and motion (right side of column). RMS values were calculated using Equation 8.

Generally, the larger the accelerometer RMS amplitudes are, the more intense is the motion performed by the subject during testing. The accelerometer amplitude is a good indication of the level of motion artifacts introduced, but is indirectly related to the motion artifact introduced into the corrupted PPG waveform. As seen in Figure 16, about half of the data sets have a relatively small range of RMS accelerometer amplitudes during motion, and median RMS accelerometer amplitudes below 500 au. Generally, these data sets showed greater differences in motion between channels. However, the pressure exerted by the headband to secure the sensor and the signal amplitude of the PPG, amongst other variables, can affect the quality of the recorded PPG waveforms.

29

6.3   FREQUENCY SNR DURING MOTION

The PSD of each AC IR waveform was calculated using Welch's method. The frequency-based SNR was calculated by integrating the area under the PSD of the IR PPG waveform during motion within a frequency band of ± 0.2 Hz around the MHR divided by the area outside of this range between 0.67 Hz and 3 Hz, as shown in Equation 8. The second harmonic of the HR (2*MHR) was subtracted out before calculations so as not to be added to the noise calculation. Table 1 shows the SNR values during motion for all 31 data sets and for each of the six channels.

Table 1. Frequency-based SNR of the IR AC PPG waveforms for all 6 channels during motion.

Data Set	Ch 1	Ch 2	Ch 3	Ch 4	Ch 5	Ch 6
Frequency Band Signal-to-Noise Ratio (SNR) [dB] during motion						
1	7.44	6.93	8.85	11.55	11.69	8.13
2	3.24	3.50	2.70	2.65	2.28	2.62
3	11.10	10.30	9.63	9.79	9.74	10.03
4	11.36	10.13	9.31	10.02	10.43	9.94
5	9.99	6.98	5.34	6.53	6.71	6.13
6	16.24	9.21	12.87	14.10	12.13	14.15
7	15.02	9.68	15.42	18.19	13.85	11.74
8	10.28	6.24	6.58	7.55	7.22	10.27
9	7.51	7.60	7.97	7.77	8.36	7.58
10	5.47	5.45	2.71	2.83	4.68	5.27
11	5.79	6.08	3.65	3.56	5.43	5.76
12	10.96	11.55	16.08	18.04	17.23	13.24
13	9.20	11.88	15.06	15.59	14.01	11.33
14	14.38	15.93	17.45	16.35	13.44	16.08
15	15.26	17.21	20.62	18.07	12.49	17.33
16	5.61	4.00	3.49	3.87	3.88	4.48
17	4.18	2.07	1.49	2.06	2.14	2.97
18	8.88	8.51	11.15	12.08	11.49	8.95
19	8.89	8.83	12.08	12.92	12.28	9.56
20	5.31	1.85	-2.79	-1.97	1.08	5.53
21	-1.30	-0.78	-3.81	-4.64	-6.02	-6.94
22	11.88	14.94	10.54	9.38	8.84	9.46
23	9.84	13.17	10.95	9.81	7.54	7.62
24	-5.07	-4.37	-1.88	-4.54	-6.52	-7.96
25	-6.08	-3.22	-3.32	-6.72	-7.02	-4.44
26	-2.68	0.88	2.21	3.56	2.40	-1.22
27	-7.71	-8.24	-6.84	-6.37	-7.81	-8.42
28	-3.63	-3.19	-0.91	0.11	-0.64	-3.10

29	-0.39	0.08	-1.76	-2.44	-2.21	-0.80
30	0.97	-0.41	-2.48	-2.06	-0.80	1.39
31	-11.96	-8.01	-6.04	-7.96	-9.50	-10.46

The motion frequency differences across channels can be visualized in comparison to the power spectral density (PSD) of the accelerometer signal during motion. Figures 19 and 20 show the PSD of the PPG waveform across all six channels, and the PSD of the tri-axial accelerometer waveform from two typical data sets where channels have different amplitudes of motion frequency. When the forehead sensor is placed on the forehead, the x-direction corresponds to motion perpendicular to the transverse plane, the y-direction perpendicular to sagittal plane, and the z-direction perpendicular to the coronal plane. The majority of the motion present in our data sets is in the x-direction, as shown in Figures 19 and 20.
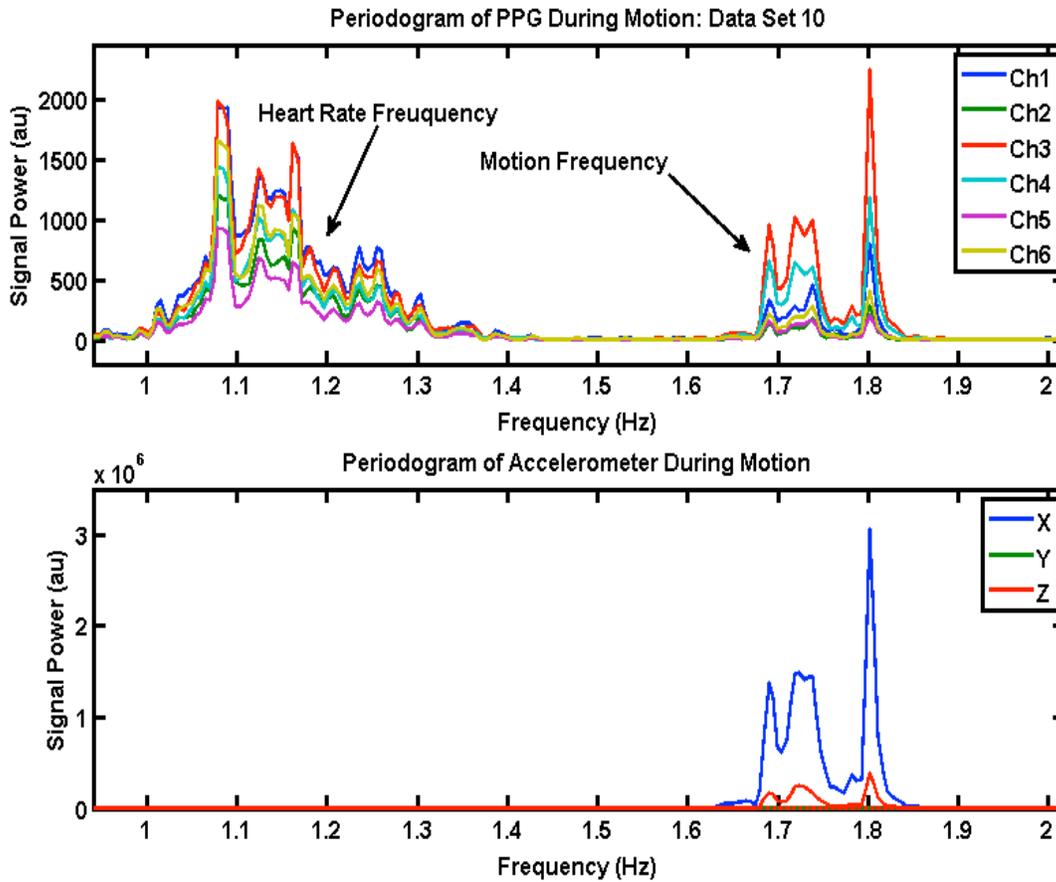
Figure 19. Power Spectral Density (PSD) of all six channels during motion for data set 10. The PSD of the accelerometer is plotted below to show the dominant motion frequencies.

Figure 19 shows where the MHR frequency is present in the PPG waveform and the motion frequency present in the PPG waveforms across all channels, corresponding to the spectra of the accelerometer during motion. In the spectra on the right, channels 2 and 5 show less amplitude in the PSD at the motion frequency than the remaining four channels.
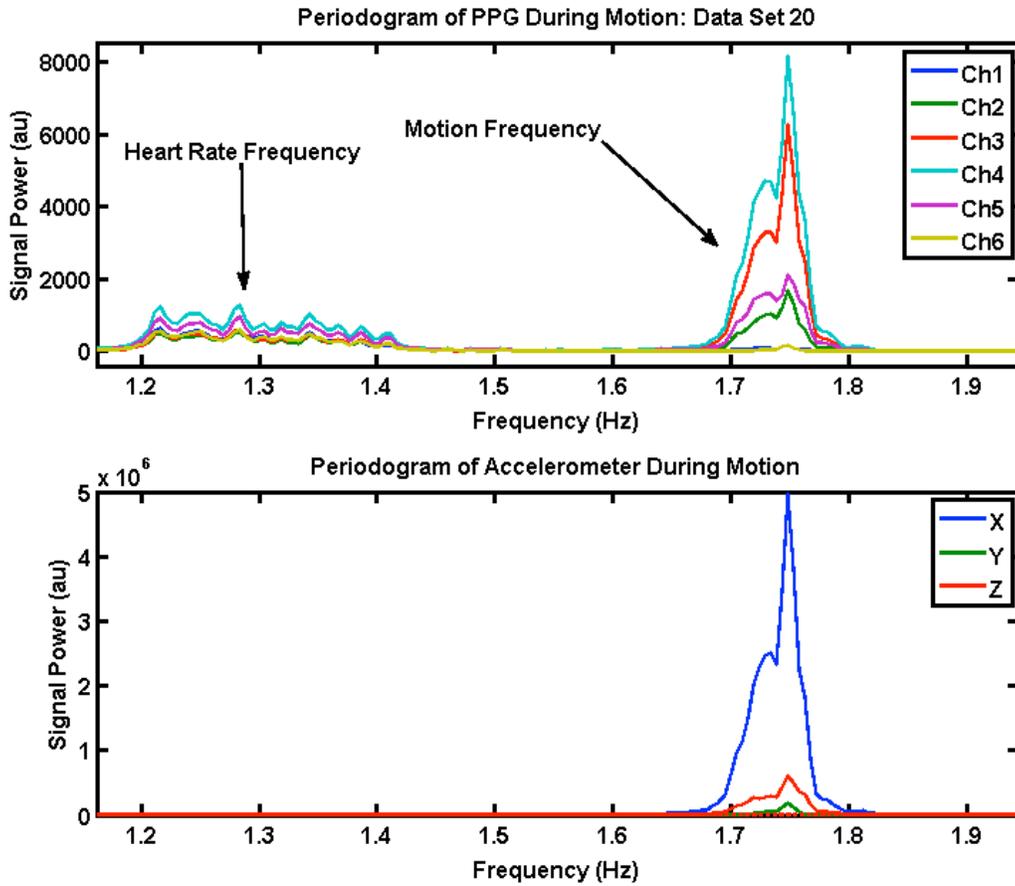
Figure 20. Power Spectral Density (PSD) of multiple channels during motion for data set 20. The PSD of the accelerometer is plotted below to show the dominant motion frequencies.

Figure 20 shows the PSD of the six AC PPG waveforms during motion and the PSD of the accelerometer data for data set 20. In this data set, channels 1 and 6 hardly show any amplitude in the motion frequency relative to the other 4 channels. Figure 21 shows the entire time-varying spectrogram from data set 14, channel 5. The prominent motion frequency, as calculated from the accelerometer PSD, is 1.72 Hz.
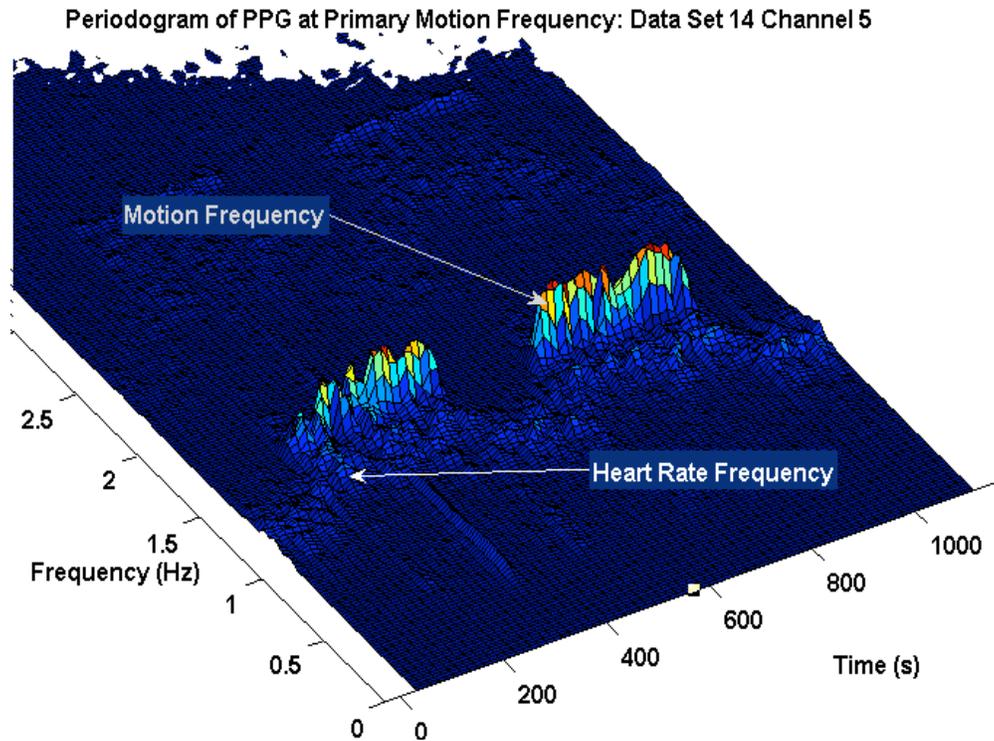
Figure 21. Time-varying spectrogram of the AC IR PPG waveform from channel 5. The heart rate frequency during motion and the dominant motion frequency, around 1.72 Hz, are labeled in white.

Figure 22 shows the PSD centered on the dominant motion frequency of 1.72 Hz for each channel. From these plots, it is evident that the motion frequency is greater in channels 3 and 4 and smaller in channels 1, 2, and 6, and that the amount of motion in each channel changes over time.
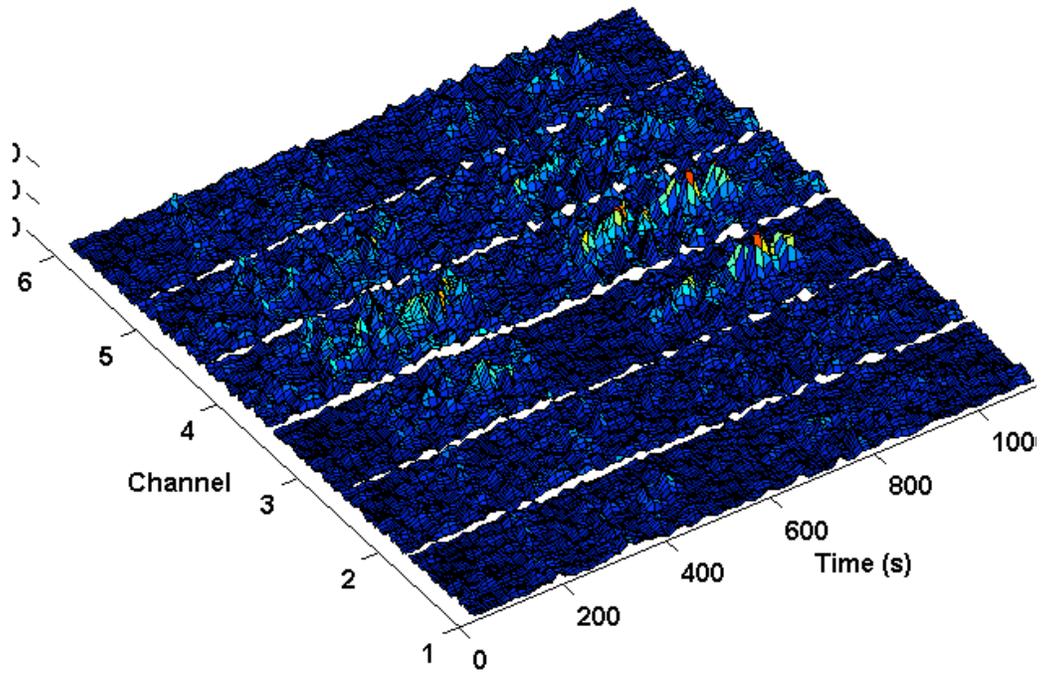
Figure 22. Time-varying spectrogram of AC IR PPG waveforms from all channels only centered on the dominant motion frequency around 1.72 Hz.

6.4 MULTICHANNEL TEMPLATE MATCHING NOISE LEVEL (MCNL)

6.4.1 MCNL VS VFCDM

Previous to development of the MCNL algorithm, noise detection performed in this lab was implemented using the Variable Frequency Complex Demodulation Method (VFCDM). To compare noise detection performance of the old method and our new method, VFCDM noise levels and MCNL values were compared for our motion data sets. Figure 23 shows the AC portion of the IR PPG waveform plotted against both the MCNL and VFCDM noise level. The accelerometer is plotted below both noise levels for motion reference. Although the trend of the noise level for each algorithm differs, both algorithms respond to and detect noise in the PPG waveform and both algorithms are sensitive to changes in the PPG waveform.
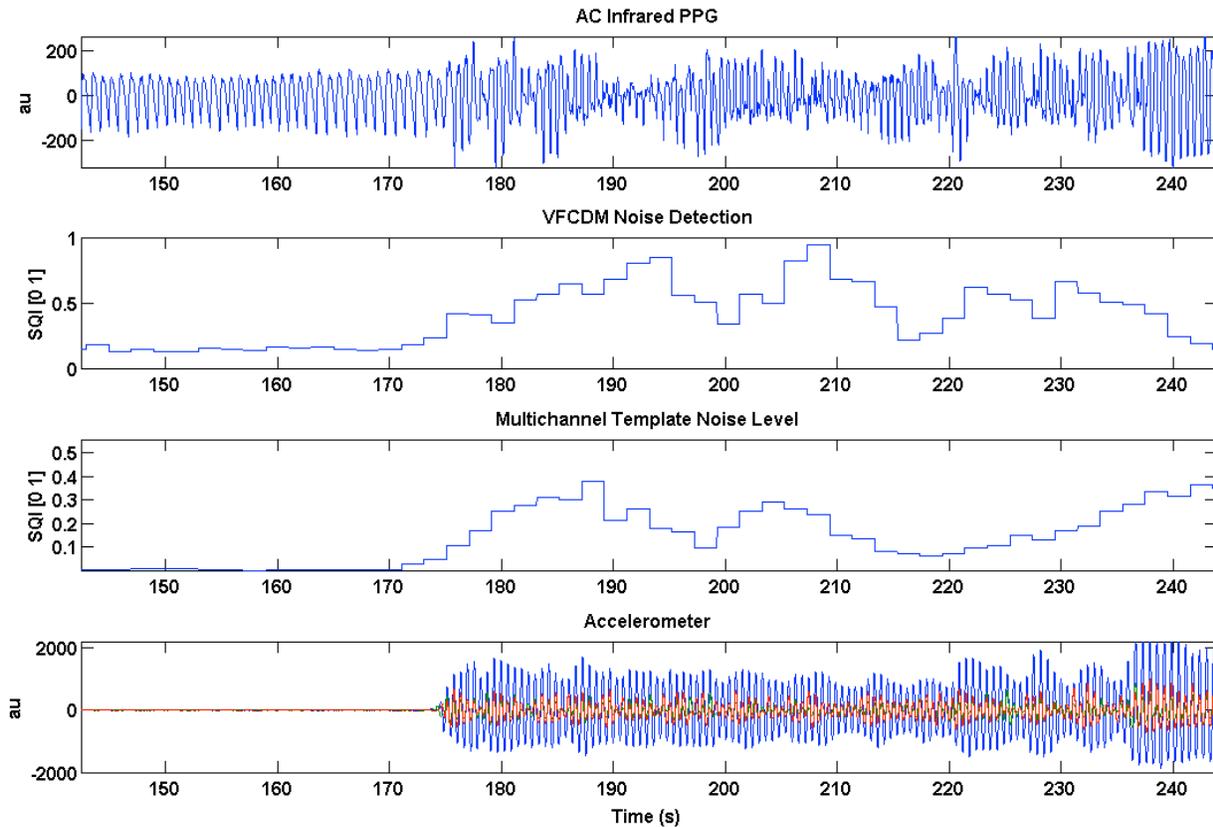
Figure 23. Time-series of the IR AC PPG, the VFCDM approach, the noise level output by the multichannel template matching algorithm (MCNL), and the accelerometer data for motion reference.

Generally, the MCNL was similar to the output from the VFCDM algorithm: low during rest, and higher during motion with varying levels of noise across channels. Figure 24 shows the same

four plots shown in Figure 23, but over the duration of an entire data set. The trends of both noise detection algorithms followed the same pattern, but spread across a different range of noise levels—the VFCDM noise level ranged from 0.07 to 1 during motion whereas the MCNL values tended to range between 0.002 and 0.4 during motion.
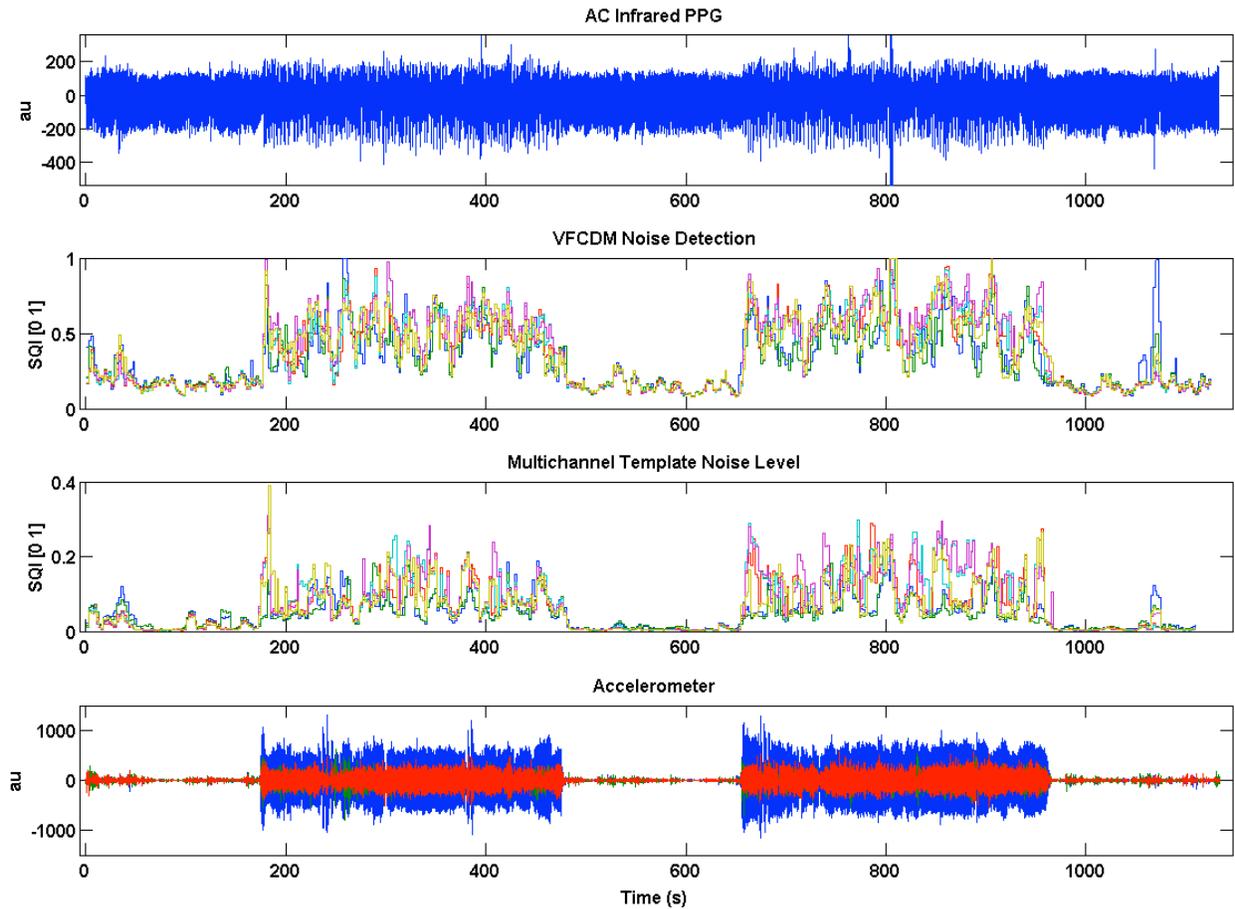
Figure 24. Time-series of the IR AC PPG, the VFCDM approach, the noise level output by the multichannel template matching algorithm (MCNL), and the accelerometer data for motion reference for the duration of an entire data set.

For the purpose of our algorithm, the noise level of each channel in relation to the other five channels is more important than the overall range of noise level output by our multichannel noise detection algorithm. Furthermore, the VFCDM was higher in computational time, making it less efficient than the MCNL computation. Because of this, it was determined that the MCNL was sufficient for detecting motion noise and as the basis for choosing channels.

6.4.2   MCNL VALUES DURING MOTION

The overall MCNL across all six channels was averaged and the corresponding Box-and-Whisker plots of the noise level during rest (left) and motion (right) are plotted in Figure 25 for all 31 data sets.
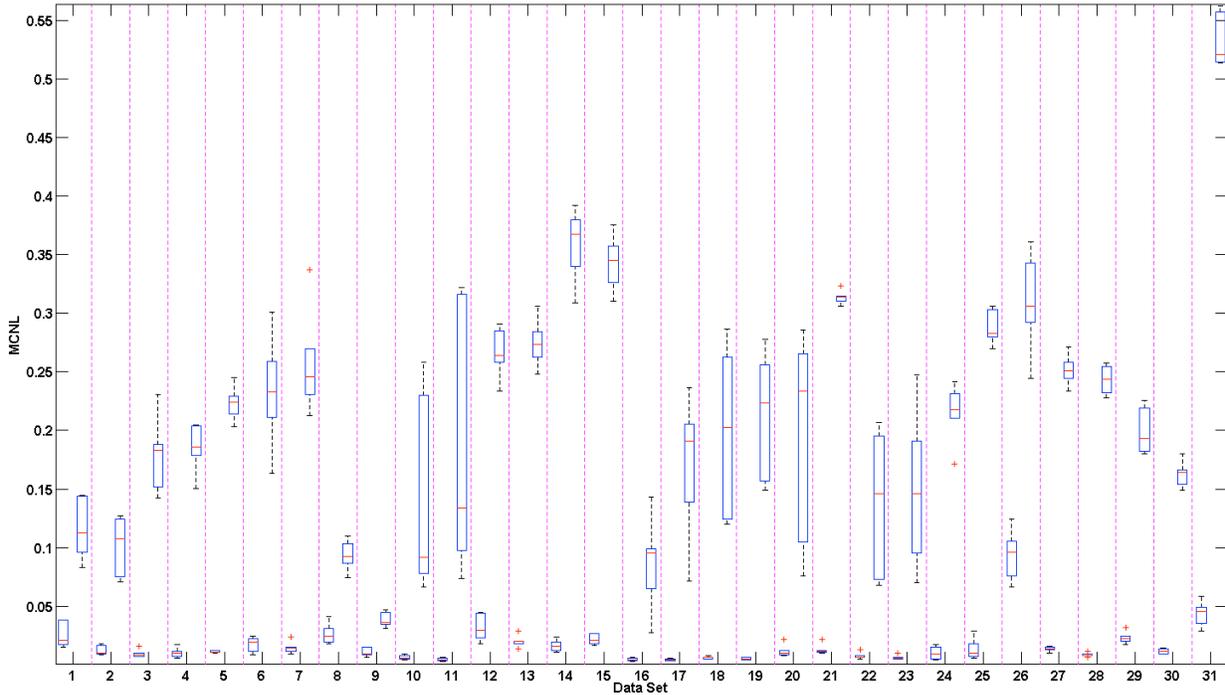
Figure 25. Box-and-Whisker plots of the average Multichannel Noise Level (MCNL) across all data sets. The average MCNL across all channels during rest and motion are shown on the left and right hand side of each column, respectively.

Multichannel noise level (MCNL) plots were used to illustrate how clean or corrupted different channels were during motion based on their respective signal morphology. During clean segments of the PPG waveforms that were recorded during rest, beats across all channels showed a relatively high degree of correlation (C) and low MCNL values. The differences in the medians and ranges of the MCNL, calculated by the multichannel template-matching algorithm during motion, showed high variation across data sets, hence displaying the variety of motions manifested by our experimental protocol. The data sets with low MCNL values during motion tended to have low accelerometer RMS amplitudes.

Figure 26 shows the time-series MCNL of data set 20 across all channels, where each color indicates a different channel. The MCNL outputs low values during rest, high values during

motion, and varying values across channels during motion. In data set 20, we noticed that the channels were more consistently separated in terms of noise level, and it is clear that the channel shown in blue has the lowest noise level of all six channels during the majority of the motion.
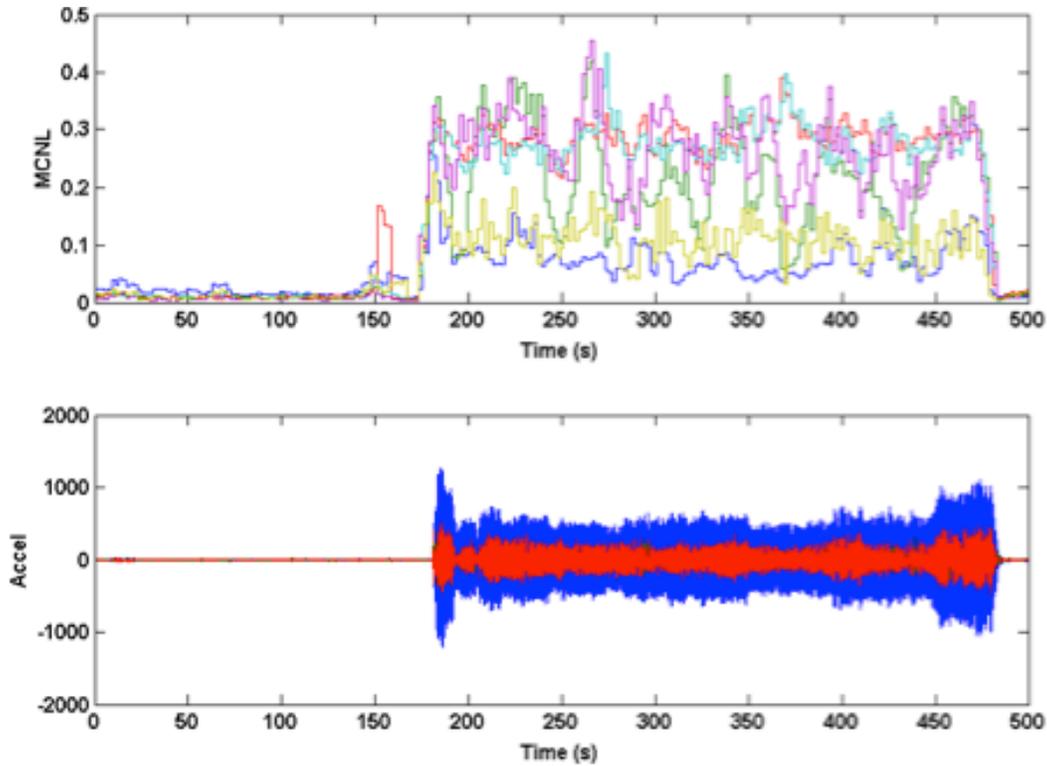
Figure 26. Multichannel template-matching noise level (MCNL) plotted for the duration of the first motion segment of data set 20. Accelerometer data is plotted below the MCNL to indicate where motion occurs.

Figure 27 shows the MCNL plotted for a portion of rest and the entire duration of the first motion segment across all six channels. The filtered accelerometer signal is shown beneath the MCNL plot for motion reference. In data set 24, the noise levels across channels were closer together during motion. The channel shown in red was the lowest during a lot of the motion, but not at every time point.
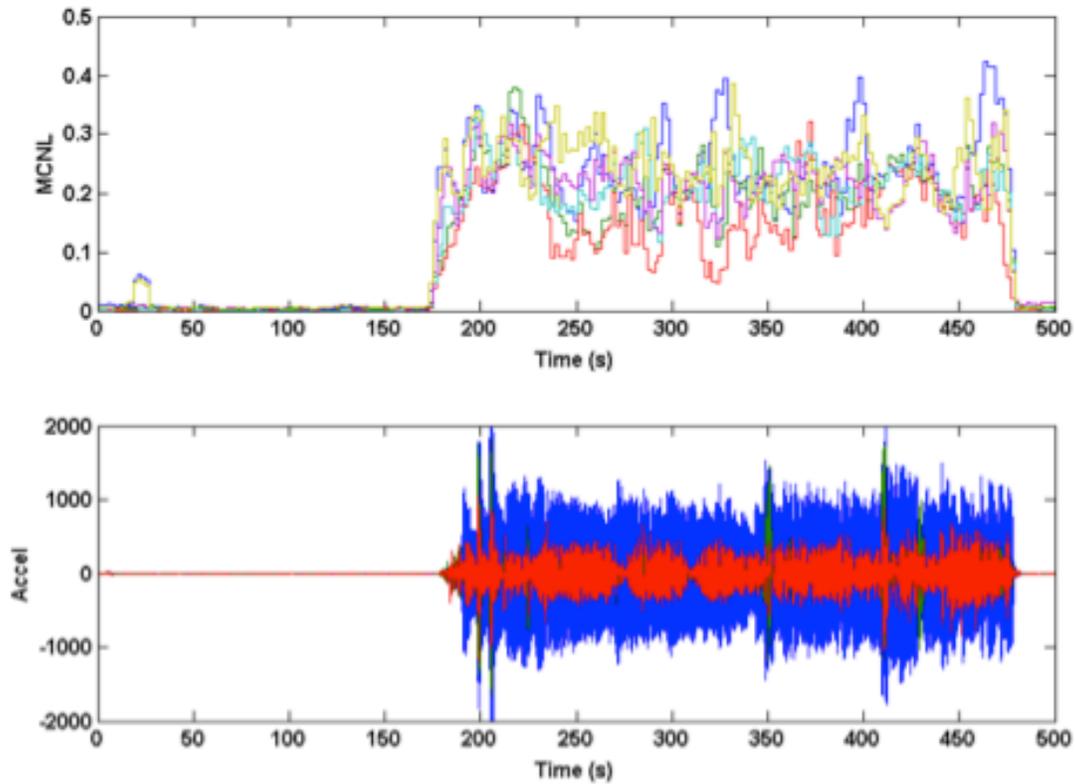
Figure 27. Multichannel template-matching noise level (MCNL) plotted for the duration of the first motion segment from data set 24. Accelerometer data is plotted below the MCNL to indicate where motion occurs.

6.4.3 SPREAD OF CHANNELS CHOSEN DURING MOTION

To see how the channel-switching algorithm affected measurements taken during motion, the number of times that any individual channel had the lowest MCNL value during motion was calculated for each data set. Figure 28 shows the histograms of the spread of channels chosen during motion for fifteen different data sets.
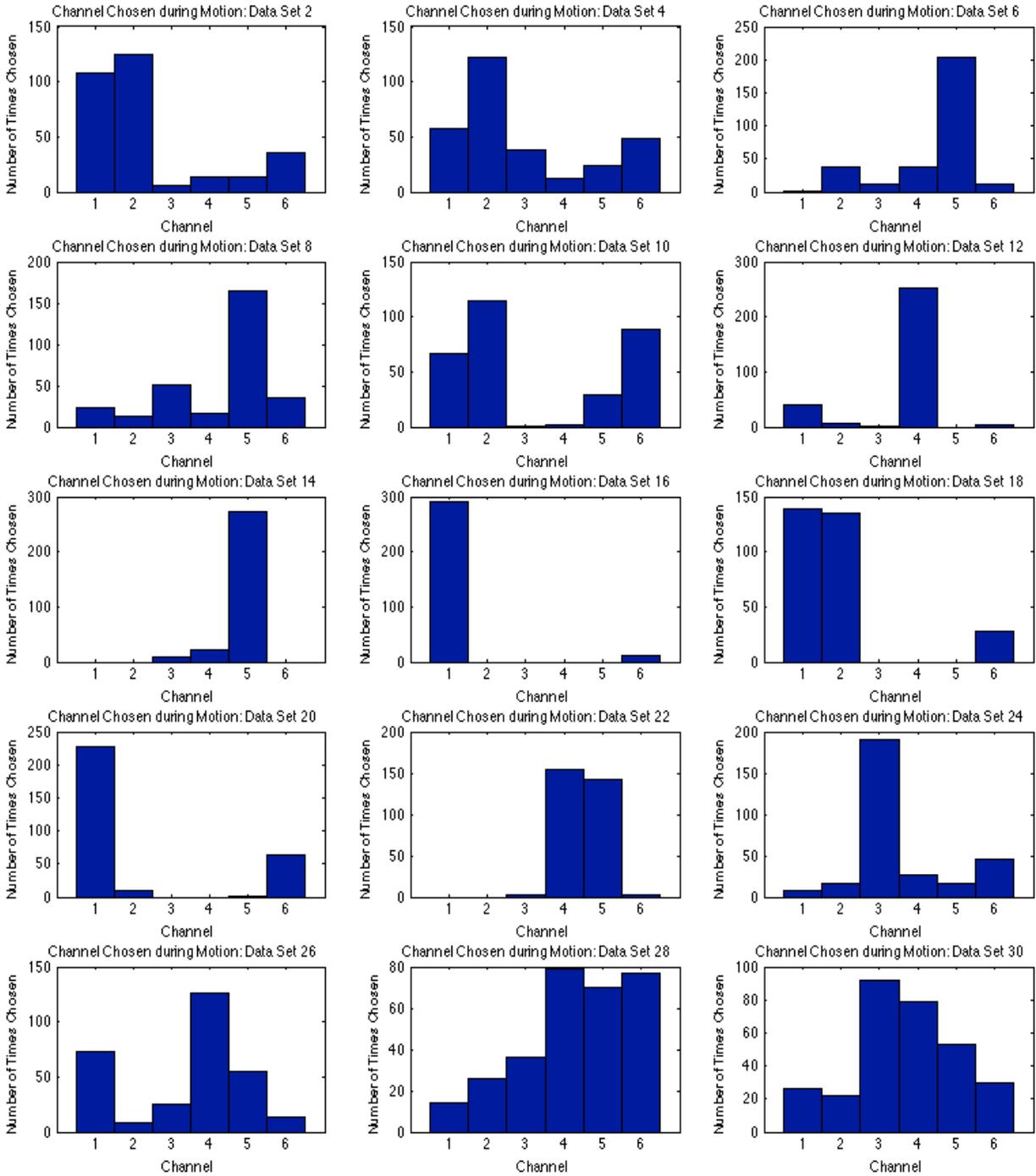
Figure 28. Histograms of the number of times each channel was chosen during motion for 15 of the 31 data sets.

Figure 28 shows the spread of measurements taken from each channel during motion for fifteen of the thirty-one data sets. For a majority of the data sets, most measurements were taken from one or two channels during motion. The channel chosen during motion did differ between data

sets, as shown in Figure 28. For example, data sets 6 and 14 used channel 5 for most of the measurements, while data set 20 used channel 1 for most of the measurements and data set 24 used channel 3 for most of the measurements. And, some of the data sets used all six channels more evenly, as seen in the histograms for data sets 28 and 30.

6.5   HR ERRORS DURING MOTION

HR errors during motion were calculated in comparison to the Masimo reference sensor according to 3 separate statistical parameters: performance index (PI), accuracy, and precision. The PI, accuracy, and precision of each channel was calculated and put in order from worst performance to best performance for each data set. To compare the multichannel estimate against the single channel estimates for each parameter, six one-sided t-tests were performed; estimates from each individual channel were compared against the multichannel estimate for each parameter. The multichannel estimate (MC) corresponds to the HR calculated by switching between channels every 2 seconds using the multichannel template-matching algorithm. A confidence value of 95 % ($\alpha = 0.05$) was used to find t-critical values for each t-test. Data sets 21 and 31 showed extremely low SNR, high accelerometer amplitudes during motion, high MCNL during motion and high HR errors across all channels. Furthermore, the PPG waveforms across all channels were completely corrupted by motion artifacts; therefore, these data sets were eliminated from the statistical calculations.

6.5.1   PERFORMANCE INDEX

Table 2 summarizes the performance index calculated for each data set across all 6 channels, and the corresponding multichannel HR estimates. For each data set, the performance index of each channel was put in order of worst PI to best PI during motion. The mean performance index per each channel, the corresponding multichannel estimate, and the mean and standard deviation of the difference between the multichannel and each individual channel are given below.

Table 2. Performance Indices of each individual channel and the multichannel estimate for HR measurements during motion.

Performance Index: Percentage of HR error less than or equal to 5 bpm [%]

Data Set	Worst		← Channel →			Best	MC
1	95.7%	95.7%	96.0%	96.0%	97.4%	97.7%	98.0%
2	96.7%	97.4%	98.0%	98.0%	99.3%	99.3%	98.7%

3	83.4%	87.4%	90.7%	91.4%	93.0%	95.7%	94.0%
4	83.4%	88.7%	89.1%	89.7%	90.1%	91.1%	91.7%
5	33.4%	48.0%	48.7%	69.2%	69.9%	74.8%	64.6%
6	85.1%	88.1%	89.4%	90.1%	93.0%	94.4%	94.7%
7	59.6%	65.2%	69.5%	77.5%	81.1%	83.8%	77.5%
8	64.6%	65.6%	65.9%	66.9%	68.2%	69.2%	65.6%
9	87.1%	87.4%	87.7%	88.1%	88.7%	88.7%	88.4%
10	63.6%	69.2%	88.1%	89.1%	90.1%	91.7%	90.1%
11	62.9%	63.2%	95.4%	98.3%	99.3%	99.7%	99.7%
12	5.6%	6.0%	7.6%	7.9%	11.9%	22.5%	9.3%
13	61.3%	63.9%	65.6%	67.2%	70.9%	78.8%	68.9%
14	33.1%	34.4%	34.4%	34.8%	36.1%	36.8%	36.4%
15	26.5%	26.5%	28.1%	28.8%	28.8%	29.5%	29.1%
16	85.8%	91.7%	92.1%	93.7%	96.7%	98.0%	98.0%
17	45.7%	58.6%	60.9%	62.9%	73.8%	85.8%	85.4%
18	8.3%	11.3%	14.6%	79.8%	81.5%	90.1%	89.1%
19	0.7%	1.3%	2.6%	61.6%	70.5%	76.2%	76.5%
20	1.7%	2.0%	46.7%	76.8%	95.7%	96.0%	95.7%
21*	0.3%	0.3%	0.3%	0.3%	1.0%	1.0%	0.3%
22	25.5%	71.5%	72.2%	85.4%	85.8%	86.1%	85.4%
23	32.5%	60.6%	85.4%	92.7%	94.7%	94.7%	94.7%
24	8.9%	19.2%	20.5%	30.8%	37.7%	58.3%	50.7%
25	5.3%	5.6%	5.6%	9.9%	11.6%	15.9%	11.9%
26	1.3%	4.6%	15.2%	44.7%	54.0%	80.5%	61.3%
27	23.8%	27.2%	31.1%	40.7%	43.7%	43.7%	38.7%
28	15.6%	23.8%	35.8%	59.3%	62.3%	69.2%	61.9%
29	28.8%	34.8%	51.7%	59.9%	59.9%	63.9%	60.6%
30	23.5%	24.8%	28.5%	32.1%	40.1%	50.3%	39.4%
31*	3.6%	4.0%	4.0%	4.0%	4.6%	5.6%	3.6%
Mean (all)	40.4%	46.1%	52.3%	62.2%	65.5%	70.0%	66.4%
Mean (excl. 21 & 31)	**43.1%**	**49.1%**	**55.8%**	**66.3%**	**69.9%**	**74.6%**	**70.9%**
Mean diff.	27.81%	21.80%	14.61%	4.57%	1.04%	-3.67%	
Std Err diff.	5.04%	4.69%	3.86%	1.31%	0.79%	0.93%	
T-statistic	5.521	4.648	3.789	3.483	1.310	-3.942	
T- critical	1.701						
α	0.05						

The performance index for the multichannel switching estimate was better than five of the six channels across all six data sets. The PI was worse than the best performing channel by 3.7 %, but better than the worst performing channel by 27.8 %. In order of performance index, we found that the multichannel estimate was significantly higher in PI than four of the six channels.

6.5.2 ACCURACY

The accuracy, or accuracy, was defined as the mean of the absolute relative error. The accuracy was calculated for all 6 channels and for the multichannel switching estimate. Six one-sided t-tests were performed to compare each individual channel with the multichannel estimate in order to test the multichannel approach to the single channel based method for calculating HR in a conventional pulse oximeter. Table 3 summarizes the accuracy of each channel and of the multichannel estimate (MC) during motion.

Table 3. Mean absolute HR errors for each individual channel and the multichannel estimate during motion.

Data Set	Worst		← Channel →			Best	MC
1	1.8	1.8	1.7	1.7	1.7	1.6	1.5
2	1.7	1.6	1.5	1.4	1.4	1.3	1.2
3	3.6	2.8	2.7	2.5	2.5	2.2	2.3
4	3.1	2.8	2.7	2.7	2.6	2.5	2.4
5	9.5	7.5	7.1	4.5	4.3	4.1	5.0
6	2.9	2.7	2.6	2.3	2.1	2.0	2.0
7	4.7	4.4	4.0	3.5	3.2	3.1	3.6
8	5.0	4.9	4.8	4.6	4.5	4.5	4.7
9	2.5	2.5	2.5	2.4	2.4	2.4	2.5
10	9.5	7.8	3.1	3.1	2.9	2.8	2.9
11	13.0	12.2	2.3	1.6	1.3	1.2	1.2
12	10.7	10.7	10.7	10.5	10.0	9.2	10.5
13	4.7	4.4	4.3	4.1	4.0	3.5	4.2
14	9.5	9.5	9.5	9.4	9.4	9.2	9.5
15	11.6	11.6	11.4	11.4	11.2	11.2	11.3
16	4.6	3.1	2.9	2.3	1.8	1.6	1.5
17	16.4	12.2	11.2	10.4	6.5	2.7	2.8
18	13.9	13.9	12.7	3.4	3.3	2.6	2.6
19	13.7	13.6	13.6	5.5	4.7	3.9	3.9
20	24.4	24.1	11.4	4.3	1.8	1.7	1.8
21*	40.4	39.4	39.3	38.8	38.8	38.8	39.3
22	8.2	4.1	4.0	2.8	2.7	2.6	2.8
23	7.7	5.8	2.9	2.1	2.0	2.0	2.0
24	20.4	15.1	14.2	11.1	10.1	7.1	10.0
25	22.9	22.8	22.2	20.5	20.2	17.5	20.0
26	44.5	42.6	34.0	19.2	15.8	6.1	17.7
27	24.5	23.6	22.4	19.4	19.4	19.0	22.6
28	22.1	20.0	13.7	9.9	9.2	6.7	8.7

29	10.4	9.5	7.4	6.8	6.7	6.1	6.9
30	12.8	11.6	11.3	10.5	9.1	7.8	9.3
31*	22.9	22.8	22.7	22.7	22.3	22.0	22.6
Mean (all)	13.0	12.0	10.2	8.2	7.7	6.7	7.7
Mean (excl. 21 & 31)	**11.7**	**10.6**	**8.8**	**6.7**	**6.1**	**5.1**	**6.1**
Mean diff.	5.63	4.53	2.68	0.58	-0.02	-1.00	
Std Err diff.	1.27	1.19	0.76	0.30	0.19	0.41	
T-Statistic	4.434	3.809	3.520	1.919	-0.092	-2.428	
T-Critical	1.701						
α	0.05						

The multichannel estimate was as low or lower than five of the six channels in accuracy of HR measurements over the remaining 29 data sets. A one-sided t-test was performed to compare the single channel estimates with the multichannel estimates. The multichannel estimate was statistically better in accuracy than four of the six channels. Data analysis showed that the average accuracy of HR error was 5.6 bpm lower than the worst performing channel and at worst, 1 bpm worse than the best performing channel.

6.5.3 PRECISION

The precision was defined as the standard deviation of the absolute relative error. Table 4 shows the precision calculated for all 6 channels and for the multichannel switching estimate (MC). A one-sided t-test was performed to compare the multichannel estimate against each individual channel to compare against a single-channel pulse oximeter.

Table 4. Precision of absolute HR errors for each channel and the multichannel estimate during motion.

	Standard Deviation of HR error during motion [bpm]						
Data Set	**Worst**	**←**	**Channel →**			**Best**	**MC**
1	2.5	2.0	1.9	1.8	1.6	1.6	1.6
2	1.5	1.4	1.4	1.3	1.3	1.2	1.2
3	3.5	2.9	2.6	2.1	1.9	1.8	2.2
4	3.2	3.2	2.8	2.7	2.6	2.6	2.6
5	6.4	5.9	5.7	3.7	3.4	3.4	4.3
6	2.6	2.4	2.2	2.2	2.0	2.0	1.9
7	3.7	3.5	3.2	2.8	2.8	2.6	3.2
8	5.1	4.9	4.8	4.4	4.3	4.3	4.6
9	3.9	3.9	3.9	3.9	3.9	3.8	4.0
10	13.2	11.1	6.5	6.2	6.2	6.2	6.2

11	17.8	16.9	5.6	4.1	1.2	1.0	1.1
12	4.5	4.0	3.8	3.4	3.3	3.3	3.6
13	3.8	3.5	3.3	3.3	3.1	2.8	3.4
14	6.5	6.5	6.4	6.4	6.4	6.4	6.5
15	8.0	7.9	7.9	7.9	7.8	7.8	7.9
16	9.5	6.7	6.3	3.9	1.9	1.6	1.6
17	16.3	14.7	13.9	13.4	10.1	3.0	3.1
18	6.0	5.9	5.1	3.4	2.7	2.4	2.4
19	4.7	4.3	4.2	4.2	4.1	3.9	3.8
20	10.8	6.0	5.9	5.2	2.0	1.9	1.9
21*	7.1	6.4	4.0	4.0	3.9	3.7	6.0
22	4.0	3.3	2.8	2.7	2.7	2.5	2.8
23	5.0	4.6	2.4	2.0	1.9	1.9	1.9
24	9.3	9.1	9.0	8.2	8.1	7.4	10.2
25	10.0	9.9	9.9	9.2	9.0	8.8	10.0
26	18.4	17.4	17.2	13.8	10.7	10.2	21.5
27	23.6	23.2	22.8	21.9	21.7	21.5	23.9
28	14.2	13.7	12.6	12.6	11.6	9.5	11.1
29	8.7	8.0	7.4	7.0	6.6	6.5	7.6
30	8.3	7.9	7.4	7.4	6.9	6.8	7.2
31*	9.6	9.1	8.8	8.6	8.6	8.4	9.1
Mean (all)	8.1	7.4	6.5	5.9	5.3	4.9	5.7
Mean (excl. 21 & 31)	**8.1**	**7.4**	**6.5**	**5.9**	**5.2**	**4.8**	**5.6**
Mean diff.	**2.5**	**1.8**	**0.9**	**0.3**	**-0.4**	**-0.8**	
Std Err diff.	0.8	0.7	0.5	0.5	0.5	0.4	
T-Statistic	3.038	2.465	1.834	0.543	-0.835	-2.097	
T-Critical	1.701						
α	0.05						

Data analysis showed that the precision of the multichannel estimate was better than four of the six channels, and statistically lower than three of the six channels. The multichannel-switching estimate increased precision by up to 2.5 bpm when compared to the worst performing channel, and was at worst 0.8 bpm worse than the best performing channel.

Generally, the multichannel data can be divided into one of three scenarios for HR measurements: (1) during motion, the estimated HR values produce low-errors for all six channels, making the multichannel HR estimation approximately as good as any individual channel used in a conventional pulse oximeter, (2) during motion, HR estimations produce high-errors in some of the channels, and low-errors in other channels, making the multichannel

derived HR estimate better than any individual channel, and (3) during motion, all six channels are corrupted by motion artifacts such that all six channels have large HR errors, and the multichannel estimate is not better than any individual channel. The first scenario is depicted in Figures 29 and 30.
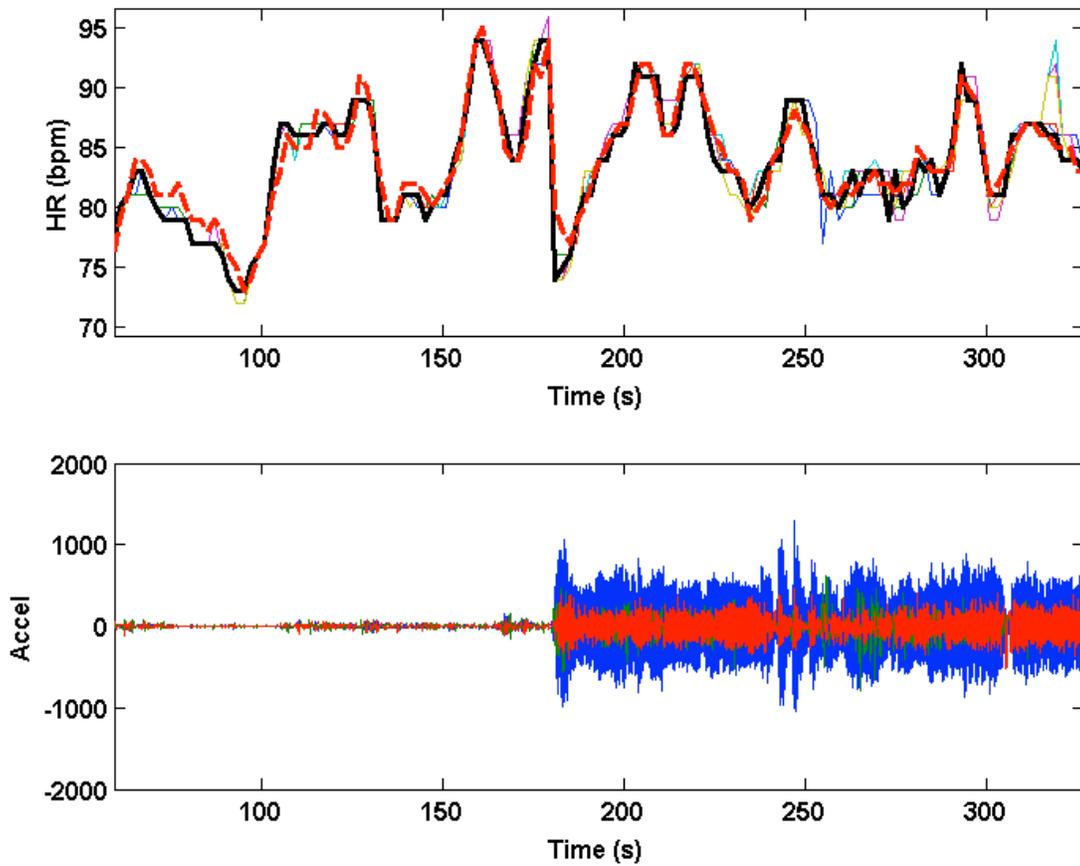
Figure 29. Multichannel HR Estimates from data set 2.
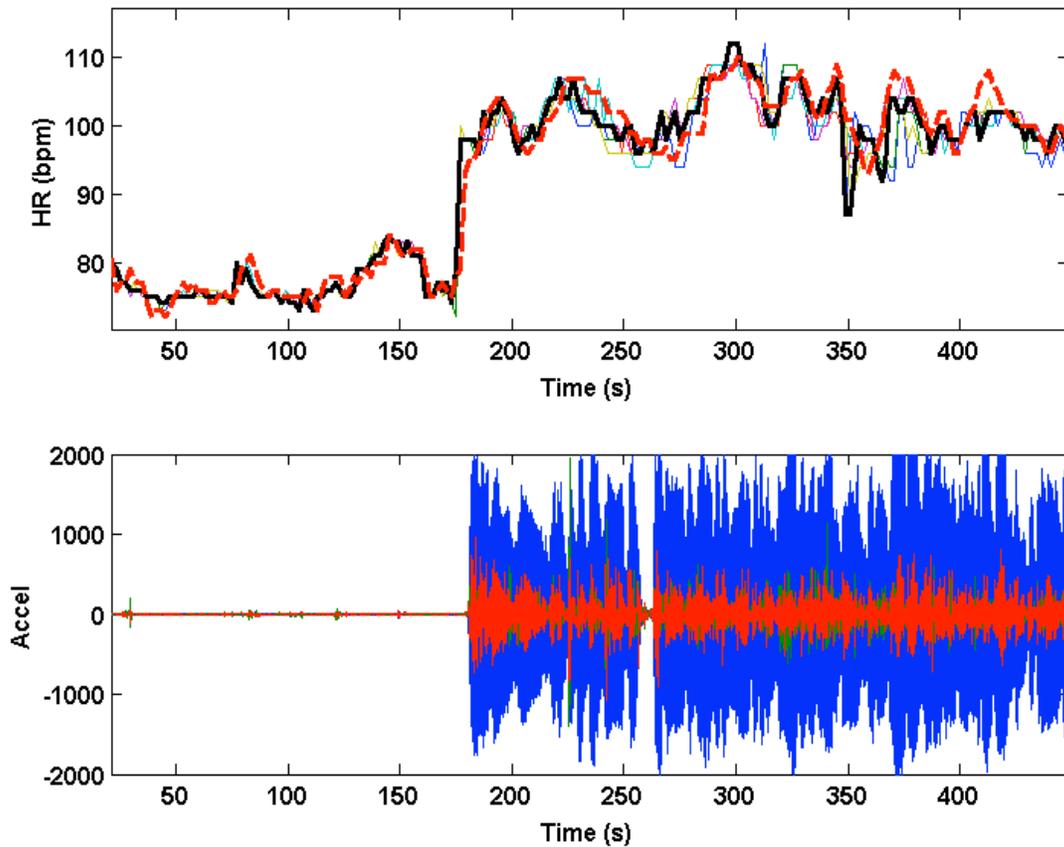
Figure 30. Multichannel HR Estimates from data set 3.

As stated earlier, we found that the signals recorded from the accelerometer do not necessarily indicate whether estimated HR values from all six channels will have a low-error or not. For example, for data sets 2 and 3 in Figures 29 and 30, where both data sets had low-error HR measurements across all channels, the overall accelerometer amplitudes during motion was vastly different. However, data sets 11 and 17, shown in Figures 31 and 32, have even lower accelerometer amplitudes than data sets 2 and 3 during motion, but have different levels of HR error across channels.
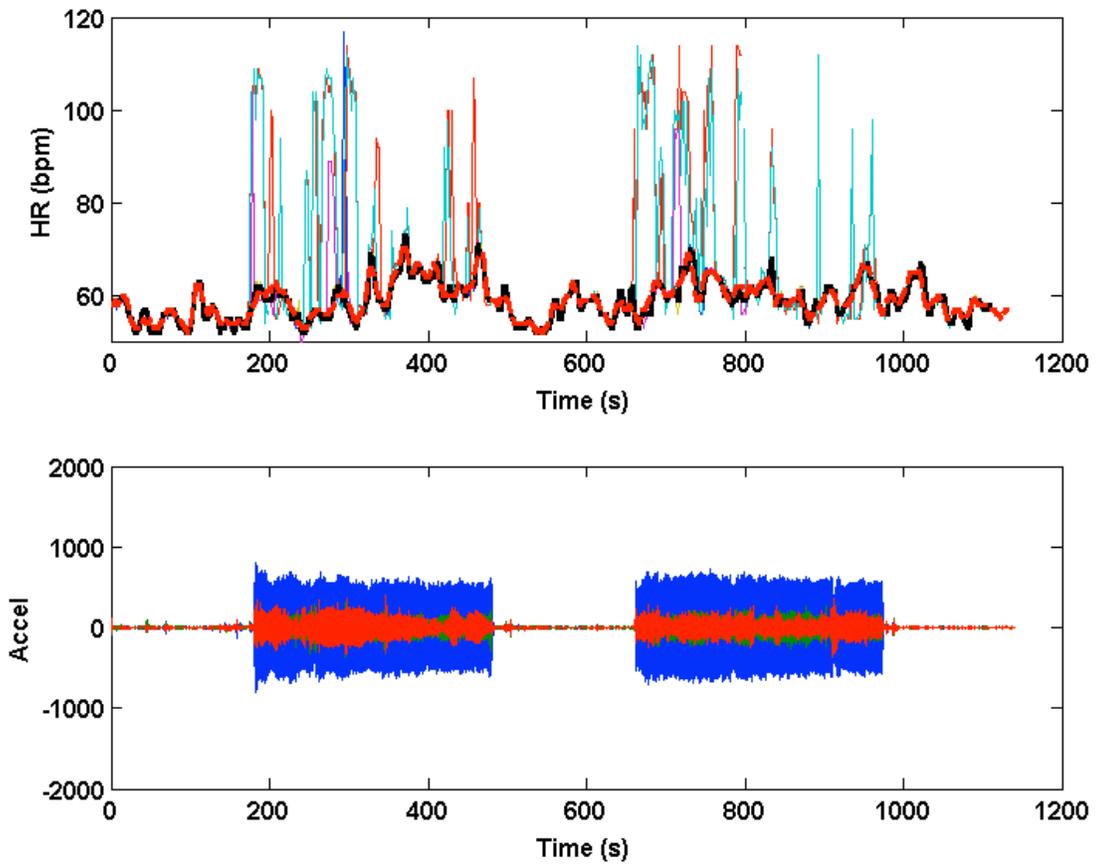
Figure 31. Multichannel HR estimates from data set 11.
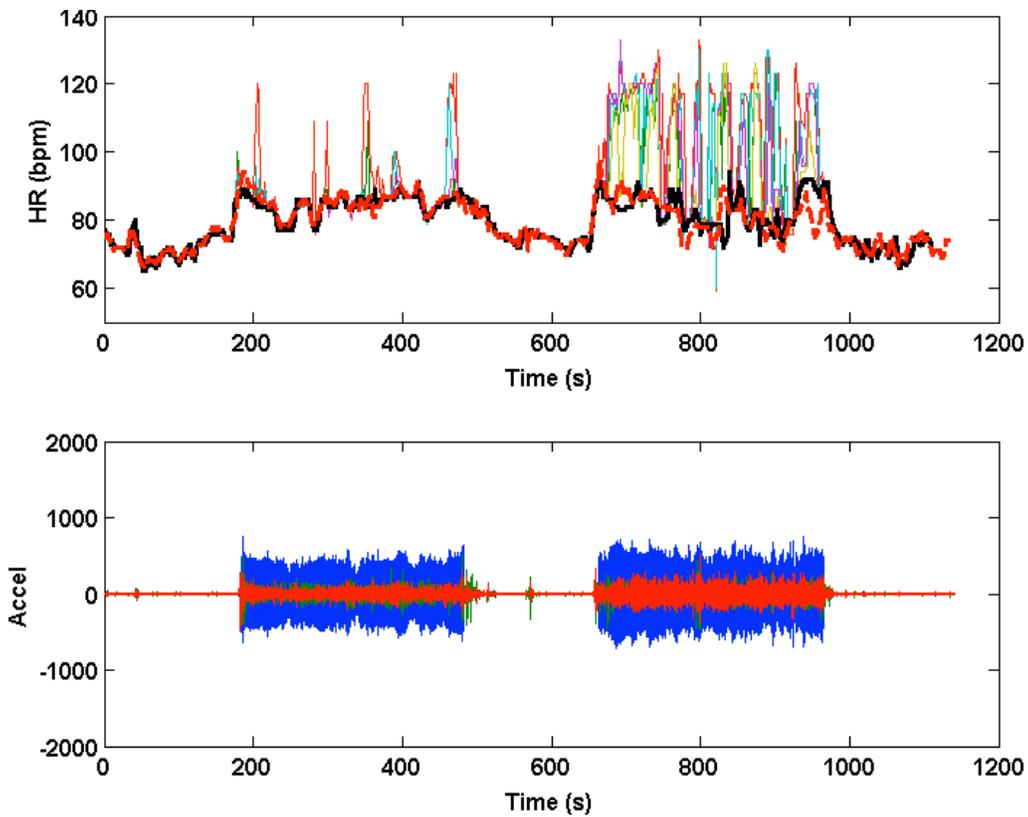
Figure 32. Multichannel HR estimates from data set 17.

This particular situation, where the multichannel was seen to be the most relevant, occurred in 9 to 10 data sets out of 31 sets. The final situation is depicted below in Figures 33 and 34: where all six channels had produced high HR errors during motion, and therefore the multichannel estimate was not successful in lowering the overall HR error lower during motion.
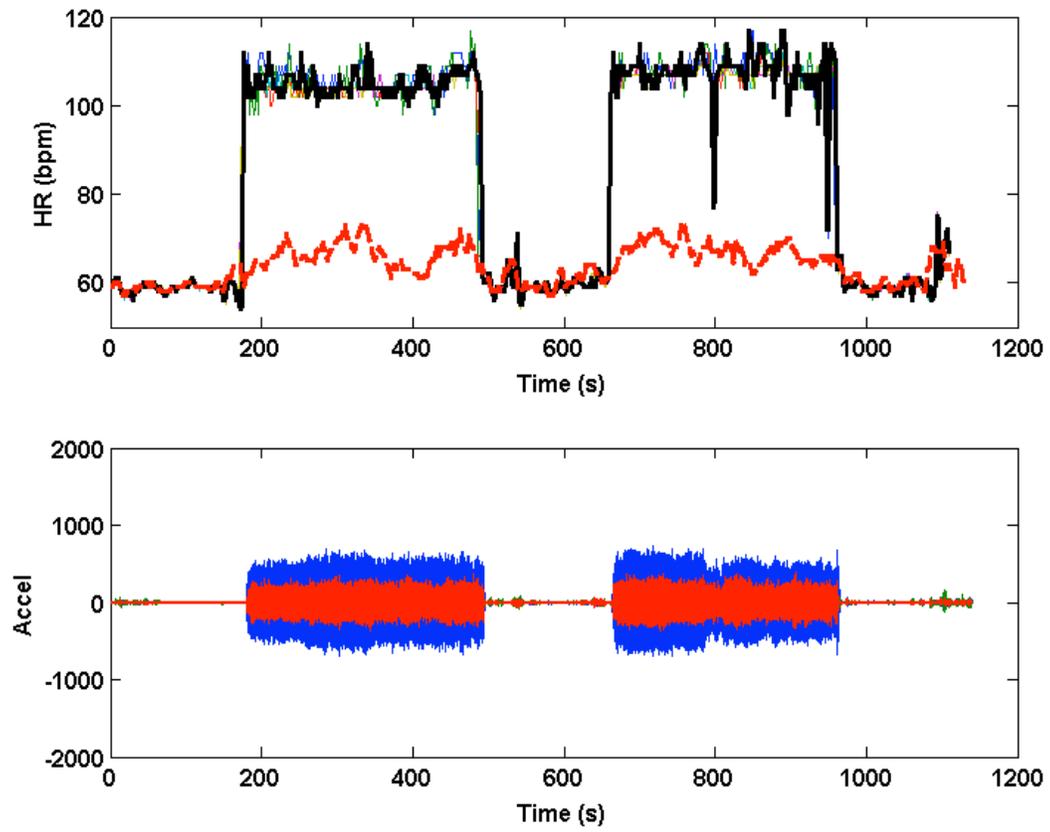
Figure 33. Multichannel HR estimates from data set 21.
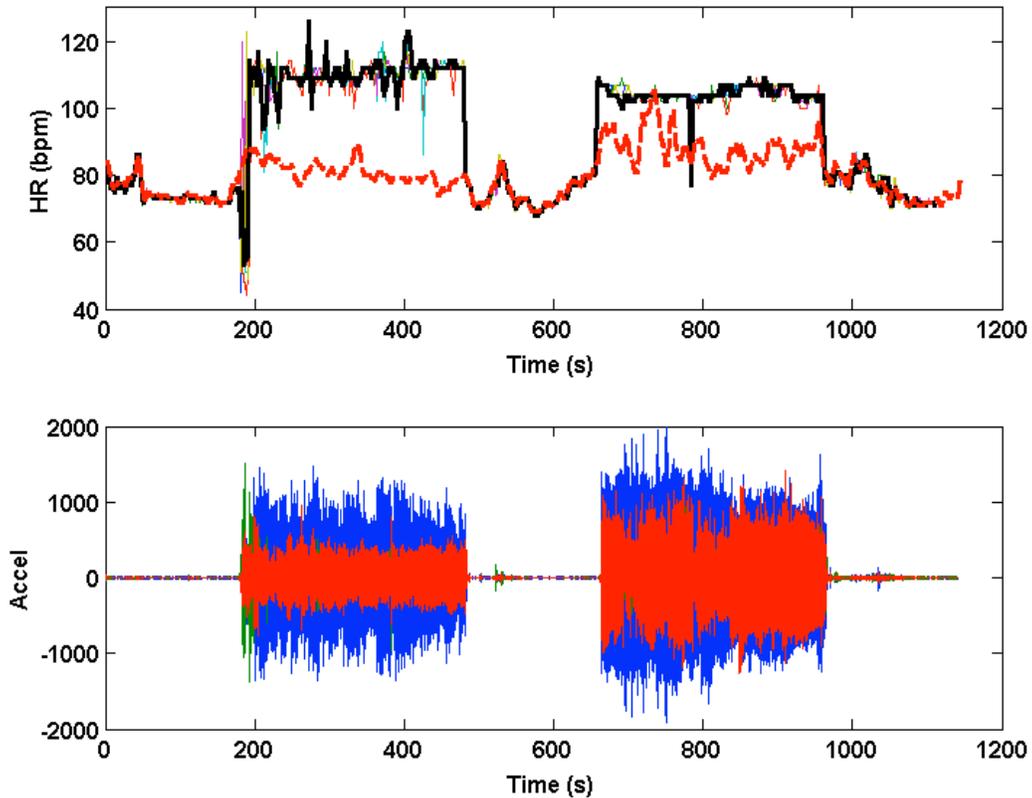
Figure 34. Multichannel HR estimates from data set 31.

Note that the accelerometer amplitudes in these data sets are lower than the accelerometer amplitudes seen in Figure 31, further suggesting that the overall amplitude level of the accelerometer is not a direct measurement of how much motion frequency will be introduced into the PPG waveforms, or how severely HR measurements during motion will be affected.

To further depict the accuracy (bias), precision, and differences in HR measurements, time-series of the HR measurements for the first 24 data sets during motion were plotted, as well as the correlation and the Bland Altman plots of the HR measurements taken with our multichannel device compared to the Masimo reference HR measurements. These plots for the first 24 data sets are shown in Appendix H.

6.6   SpO₂ ERRORS DURING MOTION

Since subjects in this study did not undergo hypoxia, the SpO_2 errors were calculated and compared to a constant SpO_2 value of 98 %. SpO_2 errors were separated into 3 distinct parameters: performance index (PI), accuracy, and precision. To compare the multichannel

estimates against the single channel estimates for each parameter, six one-sided t-tests were performed. Estimates from each individual channel were compared against the multichannel estimate for each parameter. The multichannel estimate (MC) corresponds to SpO_2 measurements taken by switching between channels every 2 seconds using the multichannel template-matching algorithm with amplitude-weighted correlation values. A confidence value of 95 % ($\alpha = 0.05$) was used to find t-critical values for each t-test. Data sets 6 showed relatively high-error SpO_2 measurements during rest, so this data set was eliminated in the statistical analysis. Since data sets 21 and 31 showed extremely low SNR, high accelerometer amplitudes during motion, high MCNL during motion, and high HR errors across all channels, these data sets were also excluded from the statistical analysis in calculating estimated SpO_2 errors. Data sets 27 and 30 showed less than 1% performance index for SpO_2 measurements during motion, so these data were also eliminated from the statistics calculations.

6.6.1 PERFORMANCE INDEX

Table 5 summarizes the performance index calculated for each data set across all 6 channels, and the multichannel SpO_2 estimates. The mean performance index per each channel, the corresponding multichannel estimate, and the mean and standard deviation of the difference between the multichannel and each individual channel are given below.

Table 5. Performance Index for SpO_2 during motion from all six channels and the multichannel estimate.

Performance Index: % of SpO_2 measurements with less than 3 % error during motion [%]

Data Set	Worst	←	Channel	→	Best		MC
1	64.57%	87.09%	95.03%	95.03%	96.69%	98.68%	95.03%
2	93.71%	97.02%	98.34%	98.34%	100.00%	100.00%	100.00%
3	2.98%	78.48%	85.43%	94.04%	97.68%	98.34%	83.44%
4	4.97%	42.05%	64.24%	75.17%	85.10%	98.01%	69.21%
5	2.65%	51.99%	64.57%	68.87%	74.50%	92.38%	78.81%
6*	24.17%	44.70%	47.02%	62.91%	90.07%	97.68%	43.05%
7	0.33%	32.12%	48.01%	73.18%	77.15%	98.68%	58.28%
8	14.57%	18.21%	51.66%	67.55%	97.02%	98.34%	80.13%
9	94.37%	99.01%	100.00%	100.00%	100.00%	100.00%	100.00%
10	90.07%	91.39%	94.37%	95.36%	96.03%	96.36%	95.36%
11	98.68%	99.01%	99.67%	100.00%	100.00%	100.00%	100.00%
12	0.00%	0.99%	15.23%	17.55%	43.71%	100.00%	79.80%
13	13.25%	22.19%	26.16%	34.11%	38.08%	53.64%	42.38%
14	0.00%	0.66%	98.34%	98.34%	98.34%	100.00%	91.06%

15	0.00%	0.33%	92.72%	100.00%	100.00%	100.00%	80.13%
16	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
17	99.01%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
18	98.68%	99.67%	100.00%	100.00%	100.00%	100.00%	100.00%
19	97.35%	100.00%	100.00%	100.00%	100.00%	100.00%	99.34%
20	1.99%	59.93%	91.72%	95.70%	100.00%	100.00%	100.00%
21*	0.00%	0.66%	6.29%	97.35%	100.00%	100.00%	19.21%
22	0.00%	16.56%	17.55%	79.14%	95.70%	97.68%	88.08%
23	0.99%	3.64%	39.40%	92.72%	94.70%	99.34%	95.03%
24	0.66%	0.99%	1.99%	5.63%	14.57%	19.54%	8.61%
25	1.99%	2.65%	2.65%	4.30%	4.97%	97.35%	30.46%
26	27.48%	30.13%	53.31%	58.61%	65.23%	84.11%	50.33%
27*	0.00%	0.00%	0.33%	0.33%	0.66%	0.66%	0.33%
28	0.00%	0.00%	0.66%	39.40%	42.38%	52.98%	30.79%
29	3.31%	6.29%	6.95%	7.62%	11.26%	13.58%	12.58%
30*	0.00%	0.00%	0.00%	0.00%	0.00%	0.66%	0.33%
31*	2.32%	6.95%	6.95%	68.87%	91.72%	93.05%	50.66%
Mean (all)	30.26%	41.70%	55.12%	68.71%	74.70%	83.58%	67.18%
Mean (excl. *)	**35.06%**	**47.71%**	**63.39%**	**73.10%**	**78.20%**	**88.42%**	**75.73%**
Mean diff.	40.7%	28.0%	12.3%	2.6%	-2.5%	-12.7%	
Std Err diff.	7.1%	6.1%	4.2%	3.0%	2.4%	3.1%	
T-statistic	5.74	4.60	2.90	0.88	-1.02	-4.06	
T-Critical 1.706							
α 0.05				*** Excluding data sets 6, 21, 27, and 31**			

For performance index in SpO$_2$ measurements, the multichannel estimate (MC) was higher than four of the six channels, and statistically higher than three of the six channels. When compared to the worst performing channel, the multichannel estimate increased the PI by an average of 40.7 % during motion.

6.6.2 ACCURACY

The accuracy, or accuracy, is defined as the mean of the absolute relative error. Table 6 shows the accuracy calculated for all 6 channels and for the multichannel switching estimate. Six one-sided t-tests were performed to compare each individual channel to the multichannel estimate to test the multichannel estimate against the single channel pulse oximeter scenario.

Table 6. Accuracy of absolute SpO$_2$ error during motion from all six channels and the multichannel estimate.

	Mean SpO$_2$ error during motion [%]				
Data Set	**Worst**	**←**	**Channel →**	**Best**	**MC**

55

1	2.73	1.53	1.47	1.31	1.09	0.92	1.26
2	1.38	1.28	0.68	0.64	0.43	0.30	0.57
3	8.02	2.40	1.85	1.49	1.02	0.78	1.99
4	6.31	3.14	2.68	2.31	2.21	0.86	2.53
5	10.43	2.82	2.74	2.71	2.49	1.61	2.33
6*	3.90	3.56	2.63	2.47	1.29	1.10	2.96
7	7.44	5.74	2.83	2.54	2.43	1.56	3.03
8	7.70	5.07	2.51	2.48	2.17	1.29	2.49
9	0.94	0.89	0.72	0.63	0.48	0.34	0.52
10	1.65	1.40	1.14	1.12	0.84	0.66	1.22
11	0.90	0.85	0.82	0.60	0.60	0.54	0.80
12	6.41	5.61	5.27	4.68	3.40	0.94	1.81
13	8.30	8.16	5.56	4.22	3.75	3.01	4.76
14	15.88	12.66	2.45	2.28	1.82	1.17	2.40
15	12.67	10.52	2.55	2.24	1.26	0.77	3.57
16	1.05	0.83	0.51	0.45	0.42	0.34	0.36
17	1.77	1.00	0.91	0.52	0.50	0.46	0.46
18	1.69	1.23	1.19	1.07	0.64	0.54	1.02
19	1.78	1.30	1.25	1.14	0.61	0.53	0.95
20	4.26	2.82	2.73	1.19	0.71	0.63	0.64
21*	9.96	7.16	3.87	2.68	2.21	1.43	4.28
22	12.94	6.39	5.12	2.25	1.45	1.24	1.83
23	14.04	7.30	4.35	1.28	1.25	1.12	1.25
24	13.67	13.60	7.43	7.23	5.66	3.47	8.24
25	8.58	8.34	6.41	6.19	4.85	1.15	4.65
26	5.09	4.56	3.03	2.95	2.37	1.29	3.35
27*	15.90	14.18	11.50	10.15	10.12	8.23	10.76
28	9.68	7.24	6.18	4.16	3.68	3.14	4.79
29	8.49	7.01	7.00	6.83	5.36	4.84	6.15
30*	13.58	13.22	11.78	11.37	10.66	8.43	10.25
31*	8.30	5.84	5.35	2.37	1.28	1.21	3.30
Mean (all)	7.27	5.41	3.69	3.02	2.49	1.74	3.05
Mean (excl.*)	**6.68**	**4.76**	**3.05**	**2.48**	**1.98**	**1.29**	**2.42**
Mean diff.	4.26	2.33	0.63	0.06	-0.44	-1.13	
Std Err diff.	0.76	0.51	0.23	0.16	0.16	0.23	
T-statistic	5.63	4.61	2.71	0.38	-2.83	-5.02	

T-Critical   1.706

α   0.05   *** Excluding data sets 6, 21, 27, and 31**

For accuracy of SpO_2 measurements, the multichannel estimate was better than four of the six channels, and statistically lower than three of the six channels. When compared to the worst

performing channel, the multichannel estimate increased the accuracy of SpO_2 measurements during motion by 4.3 %, but was worse than the best performing channel by 1.13 %.

6.6.3 PRECISION

The precision is defined as the standard deviation of the absolute error. Table 7 shows the precision calculated for all 6 channels, in order from highest to lowest, and for the multichannel switching estimate (MC). A one-sided t-test was performed to compare the multichannel estimate against each individual channel.

Table 7. Precision of the SpO_2 measurements during motion from all six channels and the multichannel estimate.

Precision: Standard Deviation of SpO_2 error during motion [%]

Data Set	Worst	←	Channel	→	Best		MC
1	1.75	1.19	0.96	0.84	0.83	0.76	0.95
2	1.41	1.05	0.97	0.55	0.25	0.19	0.37
3	2.35	1.54	1.04	1.02	0.97	0.84	1.44
4	2.04	2.04	1.25	1.09	1.00	0.83	2.01
5	3.32	2.54	2.48	2.02	1.67	1.04	2.77
6*	1.98	1.77	1.37	1.03	0.98	0.79	1.91
7	4.34	1.73	1.01	1.00	0.77	0.70	2.55
8	3.43	1.99	1.30	1.11	0.94	0.65	2.16
9	0.91	0.55	0.51	0.38	0.33	0.25	0.39
10	1.91	1.18	1.16	1.01	0.94	0.81	1.13
11	0.64	0.64	0.62	0.56	0.30	0.24	0.60
12	2.49	1.60	1.11	0.91	0.56	0.50	1.79
13	4.24	4.20	2.93	2.16	1.67	0.95	3.75
14	2.33	1.26	0.73	0.65	0.56	0.51	3.51
15	1.87	0.89	0.50	0.42	0.29	0.00	3.37
16	0.53	0.40	0.39	0.28	0.20	0.13	0.27
17	0.70	0.50	0.50	0.06	0.04	0.01	0.05
18	0.77	0.53	0.53	0.46	0.46	0.36	0.56
19	0.62	0.52	0.51	0.50	0.49	0.43	0.61
20	1.27	0.85	0.71	0.60	0.37	0.36	0.37
21*	1.36	1.36	0.56	0.51	0.50	0.41	2.02
22	3.04	2.70	1.45	1.42	0.95	0.88	2.05
23	2.91	2.76	2.19	2.07	1.84	0.61	2.01
24	3.63	2.41	2.36	2.20	2.12	0.58	5.05
25	2.76	1.53	1.37	1.27	1.25	0.74	2.63
26	2.17	2.08	2.08	1.94	1.73	1.18	2.56
27*	4.02	2.90	2.41	2.26	2.25	2.20	3.12

28	2.67	2.39	2.16	1.52	1.25	1.12	3.02
29	2.85	2.79	2.22	2.16	1.27	0.93	2.48
30*	5.21	4.29	3.52	3.30	2.83	1.83	3.67
31*	2.04	1.93	1.89	1.86	1.73	1.49	2.78
Mean (all)	2.31	1.75	1.38	1.20	1.01	0.72	2.00
Mean (excl.*)	**2.19**	**1.59**	**1.24**	**1.05**	**0.85**	**0.58**	**1.84**
Mean diff.	0.36	-0.25	-0.59	-0.78	-0.98	-1.26	
Std Err diff.	0.28	0.18	0.19	0.19	0.19	0.23	
T-statistic	1.27	-1.40	-3.17	-4.17	-5.14	-5.53	

T-Critical   1.706

α   0.05   *** Excluding data sets 6, 21, 27, and 31**

The precision of the SpO_2 errors was worse than five of the six channels, and was not statistically lower than any individual channel. The multichannel estimate was at worse less than 1.3 % lower than any individual channel, making the precision approximately on par with single channel measurements.

For all 31 data sets, the SpO_2 measurements differed between channels during motion. Similar to the multichannel HR estimates, multichannel SpO_2 measurements can be divided into three categories: (1) the multichannel SpO_2 estimates were better than the worst channel during motion, but the channel-switching algorithm had difficulty choosing the channel with the lowest-SpO_2 error (2) the multichannel SpO_2 estimate was equal to the channel with the lowest SpO_2 error during motion, and the channel-switching algorithm successfully chose the lowest-error channel, and (3) all six channels produced high-error SpO_2 measurements during motion, and the multichannel estimate was not able to obtain low-error SpO_2 measurements during motion.

Figures 35 and 36 show examples of two data sets where the multichannel SpO_2 estimate was better than the worst channel, but the algorithm was not always able to choose the channel with the lowest SpO_2 error. The accelerometer amplitude was high in both cases, and is plotted below the SpO_2 measurements as a motion reference.
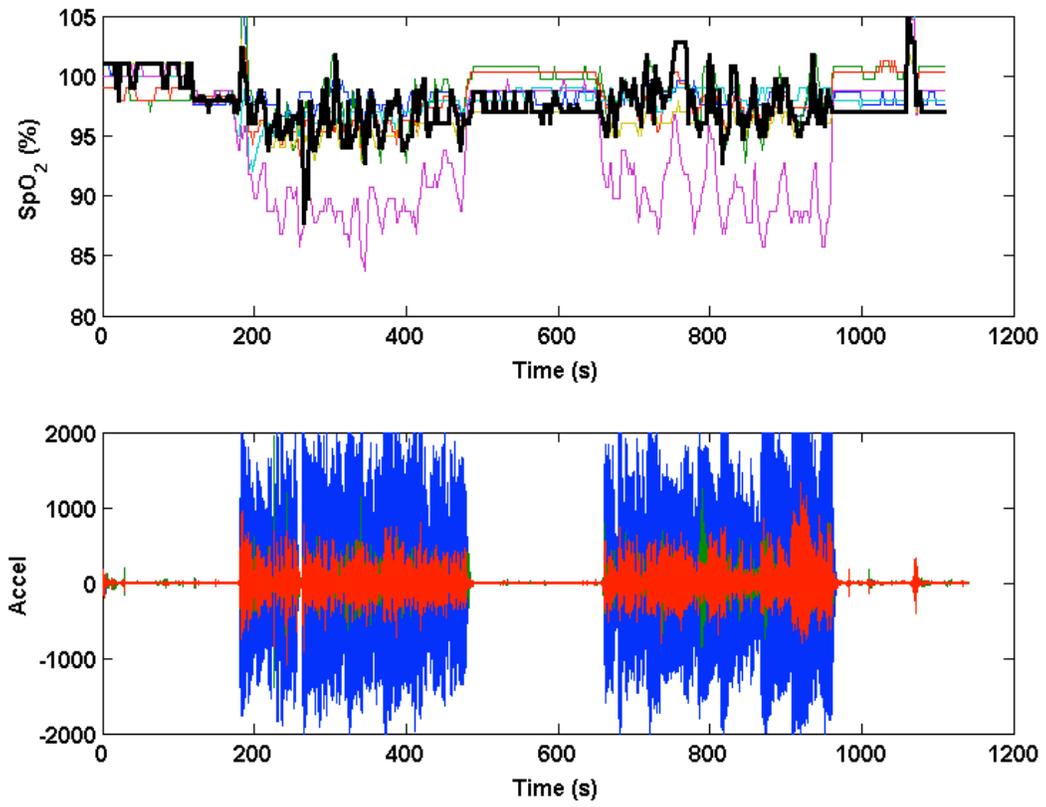
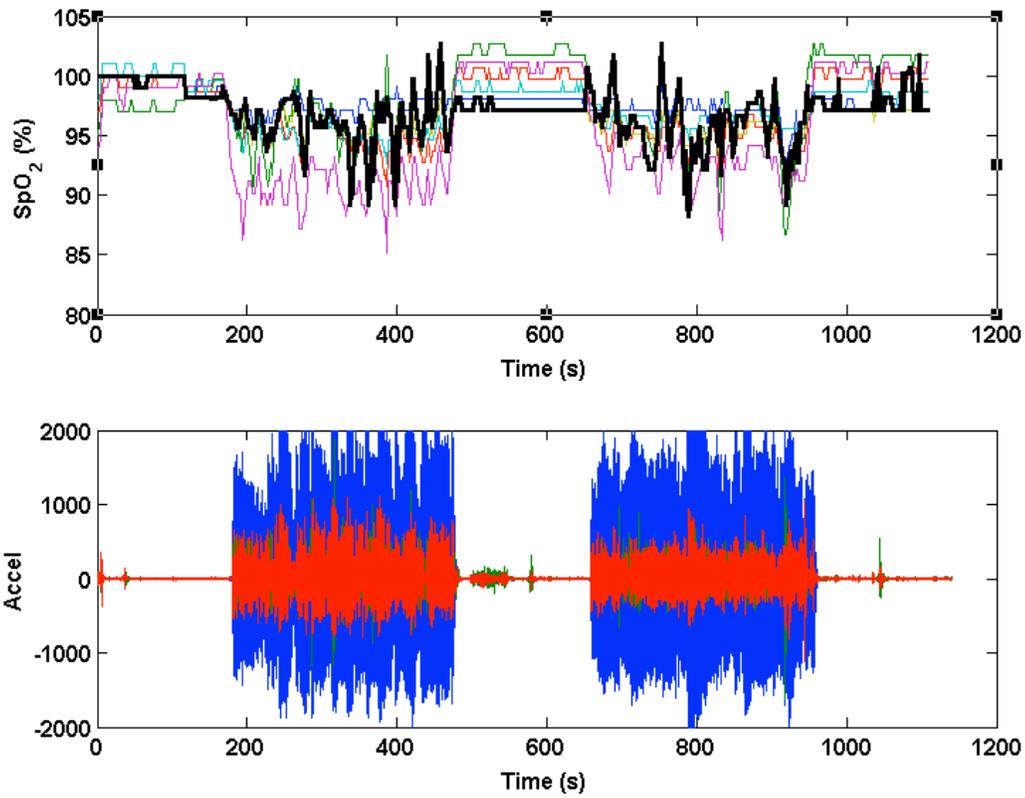**Figure 35. SpO₂ measurements taken during random motion from data set 3.**

Figure 36. SpO₂ measurements taken during random motion from data set 4.

Figures 37 and 38 show examples of two data sets where the multichannel SpO₂ estimate produced the lowest SpO₂ error during motion. The accelerometer amplitude was high in both cases, and is plotted below the SpO₂ measurements as a motion reference.
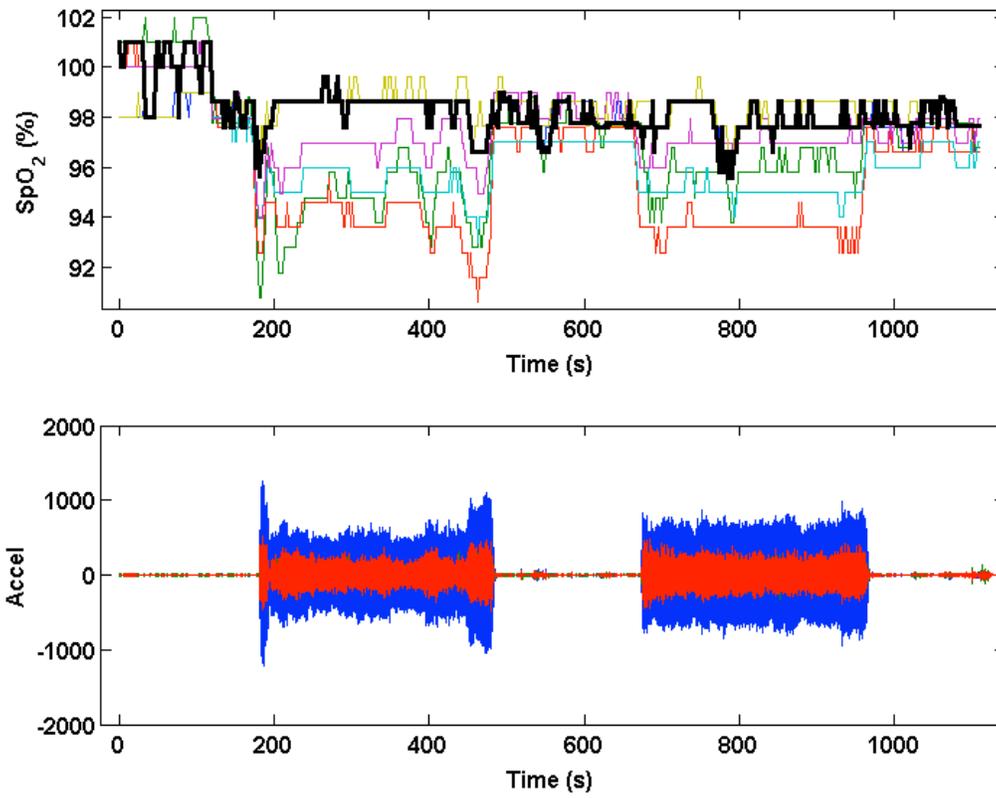
Figure 37. SpO$_2$ measurements taken during random motion from data set 20.
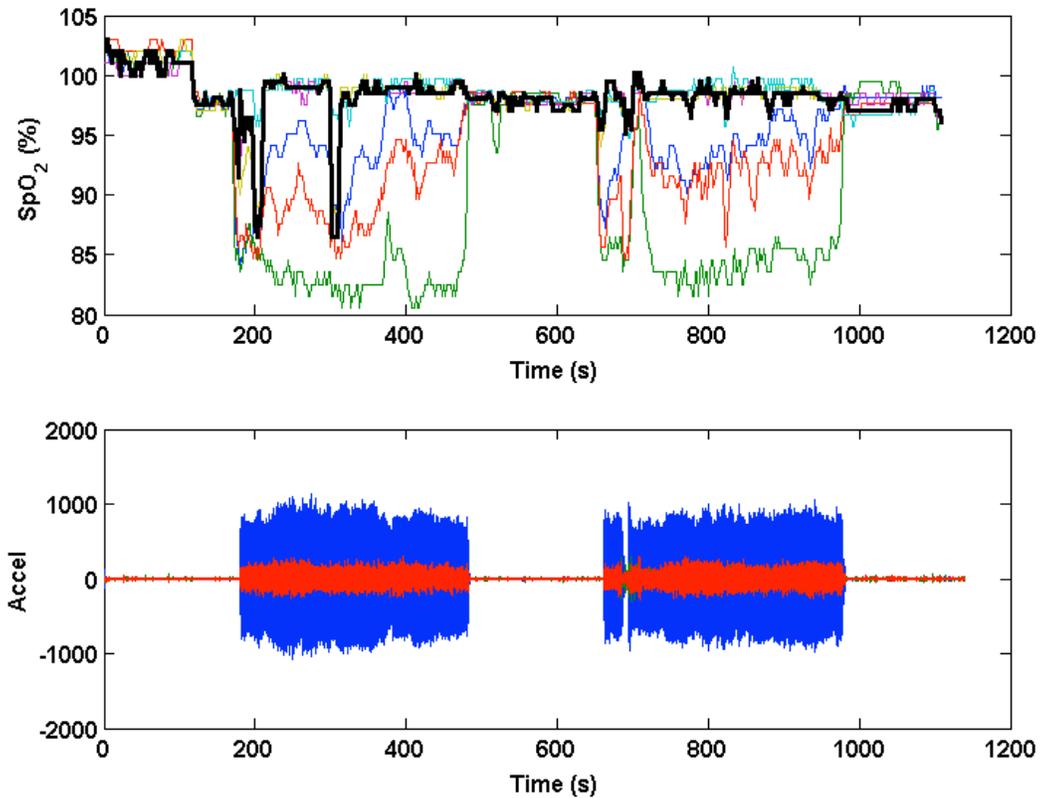
Figure 38. SpO$_2$ measurements taken during random motion from data set 23.

Figures 39 and 40 show examples of two data sets where the multichannel SpO$_2$ estimate was poor because all six channels produced high-error SpO$_2$ measurements during motion. The accelerometer amplitude is plotted below the SpO$_2$ measurements as a motion reference. The Masimo SpO$_2$ reference readings are plotted in red to show that the actual oxygen measurements did not drop during motion and to show where the baseline SpO$_2$ is in relation to the motion-corrupted estimates.
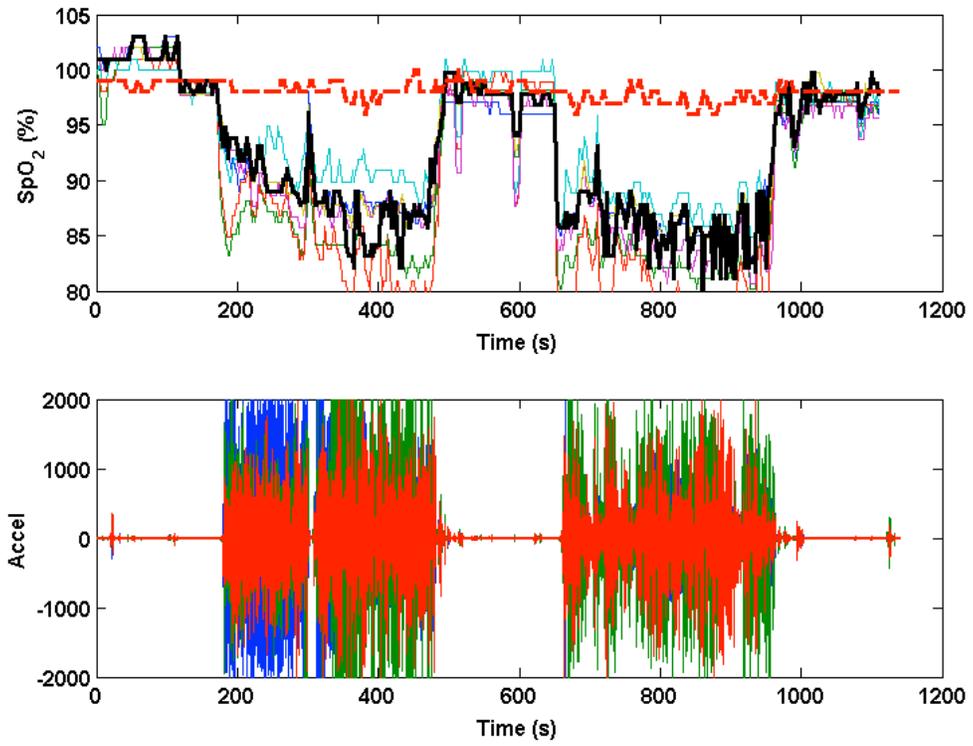
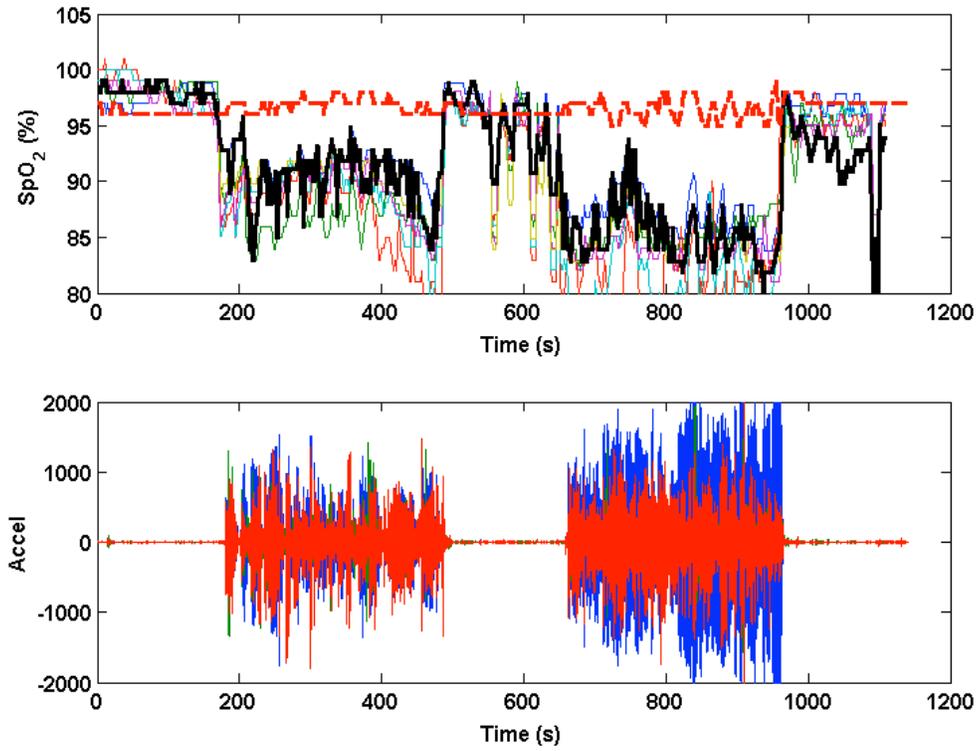**Figure 39. SpO₂ measurements taken during random motion from data set 27.**

Figure 40. SpO$_2$ measurements taken during random motion from data set 30.

7. DISCUSSION

Motion artifacts are the primary limiting factor in the utilization of pulse oximetry for mobile health applications. Motion artifacts are hard to quantify and filter out given the unpredictable nature of motion induced PPG signal corruption. In this thesis we performed experiments to generate random, aperiodic, motion corrupted data from the forehead using a reflectance-type, multichannel pulse oximeter sensor, and introduced a multichannel-switching algorithm based on previously developed template-matching algorithms. We hypothesized that motion artifacts would affect PPG waveforms in each channel during random motion differently, and that the multichannel-switching algorithm would outperform single channel estimates during motion in terms of HR error, SpO_2 error, and motion tolerance.

We collected 31 data sets of motion-corrupted PPG waveforms with a wide variety of amplitudes and frequencies. The range of motion introduced into our data is seen in both the RMS accelerometer data and the range of MCNL values output by our multichannel template-matching algorithm. Data analysis showed that channels responded differently to motion based on the frequency SNR differences between channels during motion. Depending on the severity of motion, we found that PPG waveforms in the time domain were visually different during motion across all data sets, as seen in Figure 15. Differences in motion frequency and amplitude across channels corroborate the benefits of multichannel pulse oximetry. In the case when the motion amplitude overwhelmed a single channel, but did not affect all six channels as severely, HR and SpO_2 measurements can still be obtained from the cleanest channel during motion. In the case where the motion frequency overlaps with the HR frequency and would be difficult to filter out if a single IR channel pulse oximeter was used, we found that HR and SpO_2 can still be extracted from the cleanest channel during motion with sufficient clinical accuracy when motion artifact did not affect all six channels to the same extent.

The Box and Whisker plots of the MCNL showed a wide range of MCNL values for data sets 10, 11, 17 - 20, 22, and 23. These data sets showed significant improvement when the multichannel switching estimate was implemented. We found that the multichannel approach shows the most improvement when channels differ significantly in signal quality and morphology, resulting in a high variance of MCNL values during motion. Particularly, shown in Fig. 10, when the MCNL is low for some channels during motion and high for other channels, the multichannel switching

algorithm can choose automatically the channel with the least amount of motion corruption from which to calculate the most accurate HR and SpO$_2$ values.

For HR measurements during motion, data sets 10, 11, 16, 17, 18, 19, 20, 22, and 23 showed significant improvement from individual channels compared to the multichannel estimates in absolute relative HR error. We found that one or more individual PPG channels were above the accepted tolerance in absolute relative HR error, while the multichannel estimate was at or below tolerance in absolute relative HR error. The accelerometer data showed that a wide variety of motion amplitudes were introduced across all 31 data sets. Of the 9 data sets where multichannel switching decreased the HR error significantly, only data sets 10, 11, 16, 17 and 20 had low median accelerometer amplitude. Hence, accelerometer amplitude is not an accurate measure of the level of motion artifact corruption present in the PPG waveform, or an indication of how the multichannel approach will affect estimated HR errors during motion. The frequency of the SNR values are also not a direct measure of how much motion is introduced into a signal, especially if the motion frequency overlaps the HR frequency, resulting in an incorrectly high SNR value. Nonetheless, the negative SNR values are indicators of poor signal quality, as seen in data sets 21, and 24 through 31. When the motion frequency power is greater than the HR frequency power—for instance when then SNR is negative—the multichannel approach failed to significantly improve HR measurement errors during motion.

The Box-and-Whisker plots of the MCNL showed a wide range of values for data sets 10, 11, 17 - 20, 22, and 23. These data sets showed significant improvement when the multichannel switching estimate was implemented. We found that the multichannel approach shows the most improvement when channels differ significantly in signal quality and morphology, resulting in a high variance of MCNL values during motion. Particularly, as shown in Fig. 10, when the MCNL is low for some channels during motion and high for other channels, the multichannel switching algorithm can choose automatically the channel with the least amount of motion corruption from which to calculate the most accurate HR values.

Although the benefit in performance index during motion in terms of HR error differed between data sets, when all data sets are considered, the multichannel switching estimate performed significantly better in performance index over the four worst performing channels. The PI was increased by 27.8 % when compared to the worst performing channel during motion across all 29

data sets. The benefit of multichannel-switching in accuracy of HR measurements during motion also varies across data sets, but was an improvement when all data sets were considered over the four worst performing channels during motion. The multichannel estimate increased accuracy of HR measurements by up to 5.6 bpm when compared to the worst performing channel over 29 data sets. For precision, the multichannel-switching algorithm was better than the worst channel by 2.5 bpm over 29 data sets, and was statistically better than three of the six channels. One benefit of this method and algorithm is that a user is initially unsure of what channel may be the worst corrupted or which channel will generate HR measurements with the lowest error. Generally, data sets with less severe motion and more significant differences in signal quality between channels will benefit more from the application of the proposed multichannel-switching algorithm during motion. Furthermore, the channel choosing algorithm is not always perfect, and will not always pick the channel that will give the lowest-error measurements, given that the highest correlation with the template is not always 100 % correlated to the lowest HR error. The MC estimate is not as good as the best channel during motion, but the improvements over the worst channels during motion make this approach valuable.

Furthermore, it is evident that the multichannel pulse oximetry approach investigated in this thesis has limitations when every channel is severely corrupted by motion artifacts, thus preventing highly accurate HR and SpO_2 determination from any of the PPG channels in the wearable MCPO. Nonetheless, we think that that signal reconstruction techniques developed by our group may take advantage of the varying frequency content present in the six independent PPG channels to improve HR and SpO_2 measurements during severe motion artifacts.

The benefit in performance index during motion for SpO_2 varied between data sets, and the multichannel estimate was not as good as the best channel for a majority of the data sets. Data sets 1, 3, 4, 5, 7, 8, 12, 14, 22, and 23 showed significant improvement in SpO_2 measurements—one or more channels were above tolerance in absolute relative SpO_2 error during motion while the multichannel estimate was at or below tolerance during motion. Of these data sets, only one had low accelerometer RMS values during motion and all data sets had high MCNL levels during motion. Since SpO_2 measurements depend highly on the amplitude of the signal, which is not necessarily correlated with the signal morphology, and both the RD and the IR PPG signal, it is more difficult to choose the correct channel at a given time point based on the IR PPG

morphology alone. Therefore, the multichannel-switching algorithm is currently not optimized to pick the best SpO_2 measurement currently, but still proved that multichannel SpO_2 estimates are statistically better than most of the single channel SpO_2 estimates. Based on 27 data sets, the multichannel SpO_2 estimates were overall 12.7 % worse than the best channel in performance index, but 40.7 % better than the worst channel in performance index. For accuracy, the multichannel SpO_2 estimate was 1.1 % worse than the best channel, but 4.3 % better than the worst channel. For precision, although the multichannel estimate was worse than any individual channel for SpO_2 measurements, the multichannel estimate was only worse by up to 1.3 %. As 1.3 % trade-off in precision for an overall higher percentage of low-error SpO_2 measurements during motion is considered reasonable, low-error SpO_2 measurements can be obtained during motion with a multichannel sensor for a larger range of motion compared to a single-channel pulse oximeter.

8. CONCLUSION

The work presented in this thesis proves the advantage of multichannel pulse over conventional single-channel pulse oximetry during a wide variety of random motion, much like the type of motion that can be expected from mobile patients wearing portable pulse oximetry sensors. We believe that multichannel pulse oximeters would yield more robust, motion-tolerant measurements than conventional single channel pulse oximeters. The overall accuracy of the measurements was increased by 5.6 bpm for HR and 4.3 % for SpO_2 when compared to the worst performing channel during motion. While this improvement is not a huge difference, since clinical accuracy is imperative for medical device measurements, and measurements must meet a narrow tolerance, even this small improvement can bring otherwise motion-corrupted measurements into the realm of clinically accurate measurements. The number of measurements within tolerance during motion was increased by 27.8 % for HR and 40.7 % for SpO_2 when compared to the worst performing channel during motion. This implies that the number of dropouts for HR and SpO_2 during motion would be decreased significantly when using a multichannel pulse oximeter, thus improving the performance of pulse oximetry in terms of combatting motion artifacts.

Furthermore, in addition to showing that HR and SpO_2 measurements are significantly different between channels during random motion, this study also showed that the motion corruption frequency present in the PPG waveforms varied significantly across channels. Although the multichannel pulse oximetry approach presented in this paper has limitations when every channel is severely corrupted by motion artifacts, thus preventing accurate HR or SpO_2 measurements from any of the PPG channels in the MCPO, we believe that the signal reconstruction techniques developed by our group may take advantage of the varying frequency content present in the six independent PPG channels to further improve HR and SpO_2 measurements during severe motion artifacts

9. FUTURE RECOMMENDATIONS

The type of motion introduced by our experimental protocol is consistent, aperiodic, and varying in intensity. The algorithms currently being implemented in software to combat motion artifacts have not been tested with this type of motion. Performance of these algorithms on our data in comparison with the multichannel-switching algorithm would be interesting to see, and allow a better analysis and comparison of our algorithm against the methods currently being implemented. Furthermore, the type of motion we introduced was constant over each of the five-minute motion segments and largely present in one-axis. Testing our sensor and multichannel-switching algorithm on random motion where the axis of motion would change occasionally, and where the motion segments were integrated more authentically, such as in someone running through an obstacle course, would be interesting.

Currently, a certain level of motion corruption limits the multichannel approach. Specifically, if all six channels are severely corrupted by motion, then the multichannel sensor would not be superior compared to a conventional single-channel pulse oximeter. For example, data sets 21 and 31 were eliminated from the statistical calculations because all 6 channels were severely corrupted by the dominant motion frequency. Initial work has been done on implementing a notch filter on the primary motion frequency (see: Appendix D). Using a notch filter on data set 21 brought previously corrupted channels back into the range of useable data with low-error measurements, likely making the multichannel-switching algorithm beneficial to a data set that was previously unusable. Keep in mind that the notch filter must be updated when the motion frequency changes, and will not be useful when the motion frequency overlaps the HR frequency, but development of a real-time notch filter in conjunction with the multichannel-switching algorithm would lead to more robust motion-tolerance, and lead to better results from the multichannel pulse oximeter.

Furthermore, the multichannel switching algorithm is limited to estimates from one channel every two seconds, ignoring data from the other five channels. Moreover, the multichannel-switching algorithm is not perfect all of the time, and on occasion chooses a channel that will produce large measurement errors. Investigation was started on implementation of a Kalman filter weighed by MCNL values (see: Appendix C). Preliminary results showed that the output of this filter yields smoother results on some data sets, and would likely make multichannel HR and SpO_2 estimates more accurate and precise than measurements based on the multichannel-

switching algorithm alone. Furthermore, for the purpose of this study, HR and SpO_2 measurements were calculated from all time points in each data set, even if the level of signal corruption was high. Elimination of HR or SpO_2 measurements when one or all of the channels fall above a specified threshold of MCNL could further improve the algorithm, leading to more accurate HR and SpO_2 measurements during motion.

Finally, since subjects in this study were not subjected to hypoxia during motion. The multichannel SpO_2 measurements analyzed in this study were limited to a constant SpO_2 level throughout the duration of the experiment. Introducing a drop in O_2 would allow more accurate calibration of our prototype device, and enable more accurate measurements, especially if SpO_2 measurements across channels require different empirical calibrations.

REFERENCES

1. Mannheimer, P.D. The light-tissue interaction of pulse oximetry. *Anesthesia & Analgesia* **2007**, *105*, S10-S17.

2. Jubran, A. Pulse oximetry. *Critical Care* **1999**, *3*, 11-17.

3. Dao, D.K.; Mendelson, Y.; Chon, K.H. Multi-channel pulse oximetry for wearable physiological monitoring. In *IEEE International Conference on Body Sensor Networks (BSN)*, IEEE: Cambridge, MA, 2013; pp 1-6.

4. Vetter, R.; Rossini, L.; Ridolfi, A.; Sola, J.; Chetelat, O.; Correvon, M.; Krauss, J. Frequency domain spo2 estimation based on multichannel photoplethysmographic measurements at the sternum. In *World Congress on Medical Physics and Biomedical Engineering*, Olaf Dossel, W.C.S., Ed. Spinger Berlin Heidelberg: Munich, Germany, 2009; Vol. 25, pp 326-329.

5. Wang, L.; Lo, B.P.L.; Guang-Zhong, Y. Multichannel reflective ppg earpiece sensor with passive motion cancellation. *IEEE Transactions on Biomedical Circuits and Systems* **2007**, *1*, 235-241.

6. Petterson, M.T.; Begnoche, V.L.; Graybeal, J.M. The effect of motion on pulse oximetry and its clinical significance. *Anesthesia & Analgesia* **2007**, *105*, S78-S84.

7. Dildy, G.A.; Clark, S.L.; Loucks, C.A. Intrapartum fetal pulse oximetry: Past, present, and future. *American Journal of Obstetrics and Gynecology* **1996**, *175*.

8. Dresher, R. Wearable forehead pulse oximetry: Minimization of motion and pressure artifacts. Worcester Polytechnic Institute, 2006.

9. Mendelson, Y. Pulse oximetry. In *Wiley Encyclopedia of Biomedical Engineering*, John Wiley & Sons, Inc.: 2006.

10. Asada, H.H.; Hong-Hui, J.; Gibbs, P. Active noise cancellation using mems acceerometers for motion-tolerant wearable bio-sensors. In *Engineering in Medicine and Biology Society*, IEEE: San Francisco, CA, 2004; Vol. 1, pp 2157-2160.

11. Comtois, G.; Mendelson, Y.; Ramuka, P. A comparative evaluation of adaptive noise cancellation algorithms for minimizing motion artifacts in a forehead-mounted wearable pulse oximeter. In *Engineering in Medicine and Biology Society*, IEEE: Lyon, 2007; pp 1528-1531.

12. Foo, J.Y.; Wilson, S.J. A computational system to optimise noise rejection in photoplethysmography signals during motion or poor perfusion states. *Medical and Biological Engineering and Computing* **2006**, *44*, 140-145.

13. Lee, B.; Han, J.; Baek, H.J.; Shin, J.H.; Park, K.S.; Yi, W.-J. Improved elimination of motion artifacts from a photoplethysmographic signal using a kalman smoother with simultaneous accelerometry. *Physiological Measurement* **2010**, *31*, 1585-1603.

14. Yousefi, R.; Nourani, M.; OStadabbas, S.; Panahi, I. A motion-tolerant adaptive algorithm for wearable photoplethysmographic biosensors. *IEEE Journal of Biomedical and Health Informatics* **2014**, *18*, 2168-2194.

15. Goldman, J.M.; Petterson, M.T.; Kopotic, R.J.; Barker, S.J. Masimo signal extraction pulse oximetry. *Jounal of Clinical Monitoring and Computing* **2000**, *16*, 475-483.

16. Ram, M.R.; Madhav, K.V.; Krishna, E.H.; Komalla, N.R.; Reddy, K.A. A novel approach for motion artifact reduction in ppg signals based on as-lms adaptive filter. *IEEE Transactions on Instrumentation and Measurement* **2012**, *61*, 1445-1457.

17. Patterson, J.A.C.; Yang, G.-Z. Ratiometric artifact reduction in low power reflective photoplethysmography. *IEEE Transactions on Biomedical Circuits and Systems* **2011**, *5*, 330-338.

18. Shimazaki, T.; Hara, S.; Okuhata, H.; Nakamura, H. Cancellation of motion artifact induced by exercise for ppg-based heart rate sensing. In *Engineering in Medicine and Biology Society* IEEE: Chicago, IL, 2014; pp 3216-3219.

19. Wijshoff, R.W.C.G.R.; Mischi, M.; Veen, J.; van der Lee, A.M.; Aarts, R.M. Reducing motion artifacts in photoplethysmograms by using relative sensor motion: Phantom study. *Journal of Biomedical Optics* **2012**, *17*, 1-15.

20. Enriquez, R.H.; Castellanos, M.S.; Rodriguez, J.F.; Caceres, J.L.H. Analysis of the photoplethysmographic signal by means of the decomposition in principle components. *Physiological Measurement* **2002**, *23*, N17-N29.

21. Kim, B.S.; Yoo, S.K. Motion artifact reduction in photoplethysmography using independent component analysis. *IEEE Transactions on Biomedical Engineering* **2006**, *53*, 566-568.

22. Krishnan, R.; Natarajan, B.; Warren, S. Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data. *IEEE Transactions on Biomedical Engineering* **2010**, *57*, 1867-1876.

23. Peng, F.; Zhang, Z.; Gou, X.; Liu, H.; Wang, W. Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent component analysis and adaptive filter. *BioMedical Engineering OnLine* **2014**, *13*, 1-14.

24. Salehizadeh, S.M.A.; Dao, D.K.; Chong, J.W.; McManus, D.; Darling, C.; Mendelson, Y.; Chon, K.H. Photoplethysmograph signal reconstruction based on a novel motion aritfact detection-reduction approach part ii: Motion and noise artifact removal. *Annals of Biomedical Engineering* **2014**, *42*, 2251-2263.

25. Chong, J.W.; Dao, D.K.; Salehizadeh, S.M.; McManus, D.D.; Darling, C.E.; Chon, K.H.; Mendelson, Y. Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection-reduction approach. Part i: Motion and noise artifact detection. *Annals of Biomedical Engineering* **2014**, *42*, 2238-2250.

26. Li, K.; Warren, S. A wireless reflectance pulse oximeter with digital baseline control for unfiltered photoplethysmograms. *IEEE Transactions on Biomedical Circuits and Systems* **2011**, *6*, 269-278.

27. Lee, J.; Jung, W.; Kang, I.; Kim, Y.; Lee, G. Design of filter to reject motion artifact of pulse oximetry. *Computer Standards & Interfaces* **2004**, *26*, 241-249.

28. Li, K.; Warren, S.; Natarajan, B. Onboard tagging for real-time quality assesment of photoplethysmograms acquired by a wireless reflectance pulse oximeter. *IEEE Transactions on Biomedical Circuits and Systems* **2011**, *6*, 54-63.

29. Lee, H.-W.; Lee, J.-W.; Jung, W.-G.; Lee, G.-K. The periodic moving average filter for removing motion artifacts from ppg signals. *International Journal of Control, Automation, and Systems* **2007**, *5*, 701-706.

30. Milanesi, M.; Martini, N.; Vanello, N.; Positano, V.; Santarelli, M.F.; Paradiso, R.; Rossi, D.D.; Landini, L. Multichannel techniques for motion artifacts removal from electrocardiographic signals. In *EMBS Annual International Conference*, IEEE: New York, NY, 2006; pp 3391-3394.

31. Baca, A.; Biagetti, G.; Camilletti, M.; Crippa, P.; Falaschetti, L.; Orcioni, S.; Rossini, L.; Tonelli, D.; Turchetti, C. Carma: A robust motion artifact reduction algorithm for heart

rate monitoring from ppg. In *European Signal Processing Conference*, IEEE: Nice, France, 2015.

32.   Salehizadeh, S.M.A.; Dao, D.; Bolkhovsky, J.; Cho, C.; Mendelson, Y.; Chon, K.H. A novel time-varying spectral filtering algorithm for reconstruction of motion artifact corrupted heart rate signals during intense physical activities using a wearable photoplethysmogram sensor. *Sensors* **2016**, *16*.

33.   Santos, S.A.; Venema, B.; Leonhardt, S. Accelerometer-assisted ppg measurement during physical exercise using the lavimo sensor system. *Acta Polytechnica* **2012**, *52*, 80-85.

34.   Mendelson, Y.; Ochs, B.D. Noninvasive pulse oximetry utilizing skin reflectance photoplethysmography. *IEEE Transactions on Biomedical Engineering* **1988**, *35*, 798-805.

35.   Orphanidou, C.; Bonnici, T.; Charlton, P.; Clifton, D.; Vallance, D.; Tarassenko, L. Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring. *IEEE Journal of Biomedical and Health Informatics* **2015**, *19*, 832-838.

36.   Silva, I.; Lee, J.; Mark, R.G. Signal quality estimation with multichannel adaptive filtering in intensive care settings. *IEEE Transactions on Biomedical Engineering* **2012**, *59*, 2476-2485.

37.   Li, Q.; Mark, R.G.; Clifford, G.D. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a kalman filter. *Physiological Measurement* **2008**, *29*, 15-32.

APPENDICES

A. FINGER AND EAR SENSORS

Originally, the scope of this project included testing two other transmittance-based multichannel pulse oximeters, utilizing a 3PD finger sensor and a 3PD ear sensor. After analyzing preliminary data, the two 3PD multichannel sensors were deemed less beneficial than the 6PD forehead sensor. First, the ear and the finger are both small surfaces, and the area available to place multiple PDs is much more constrained than the forehead. Since the PDs are built into a smaller area, the PPG waveforms are more likely to be much more similar during motion. Second, the differences in path length displacement between channels are more likely to differ in reflectance pulse oximetry than in transmittance pulse oximetry. Time series plots of the AC IR PPG waveforms from three channels in the 3PD ear sensor and the 3PD finger sensor are shown for comparison in Figure A1, confirming that motion artifacts affect all three channels similarly. In general, differences in channels were much greater in the 6PD forehead sensor than in either of the 3PD sensors. Based on these observations, it was determined that multichannel pulse oximetry works best in reflectance pulse oximetry, and future devices that make use of the multichannel hardware design should be limited to reflectance-based sensors.
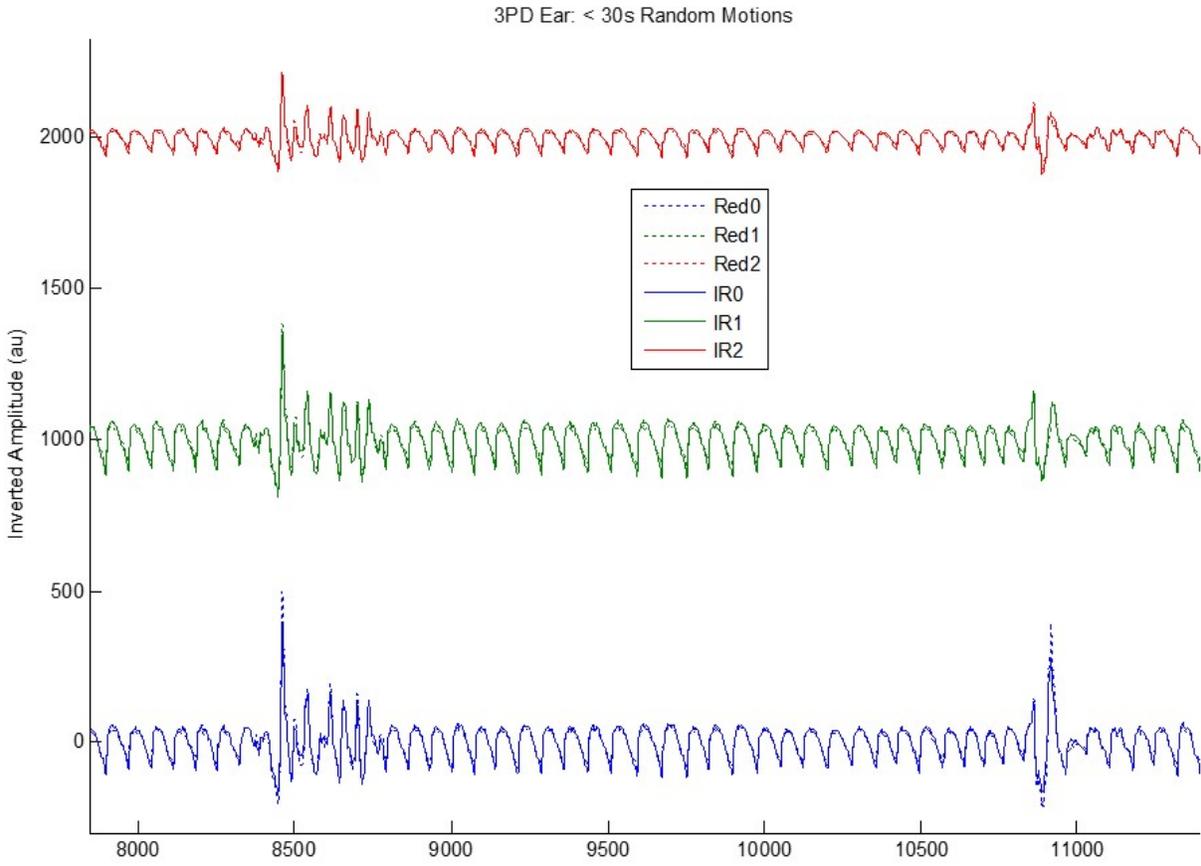
Figure A1. Time series plots of the 3 RD and 3 IR AC PPG waveforms from the 3PD ear sensor.
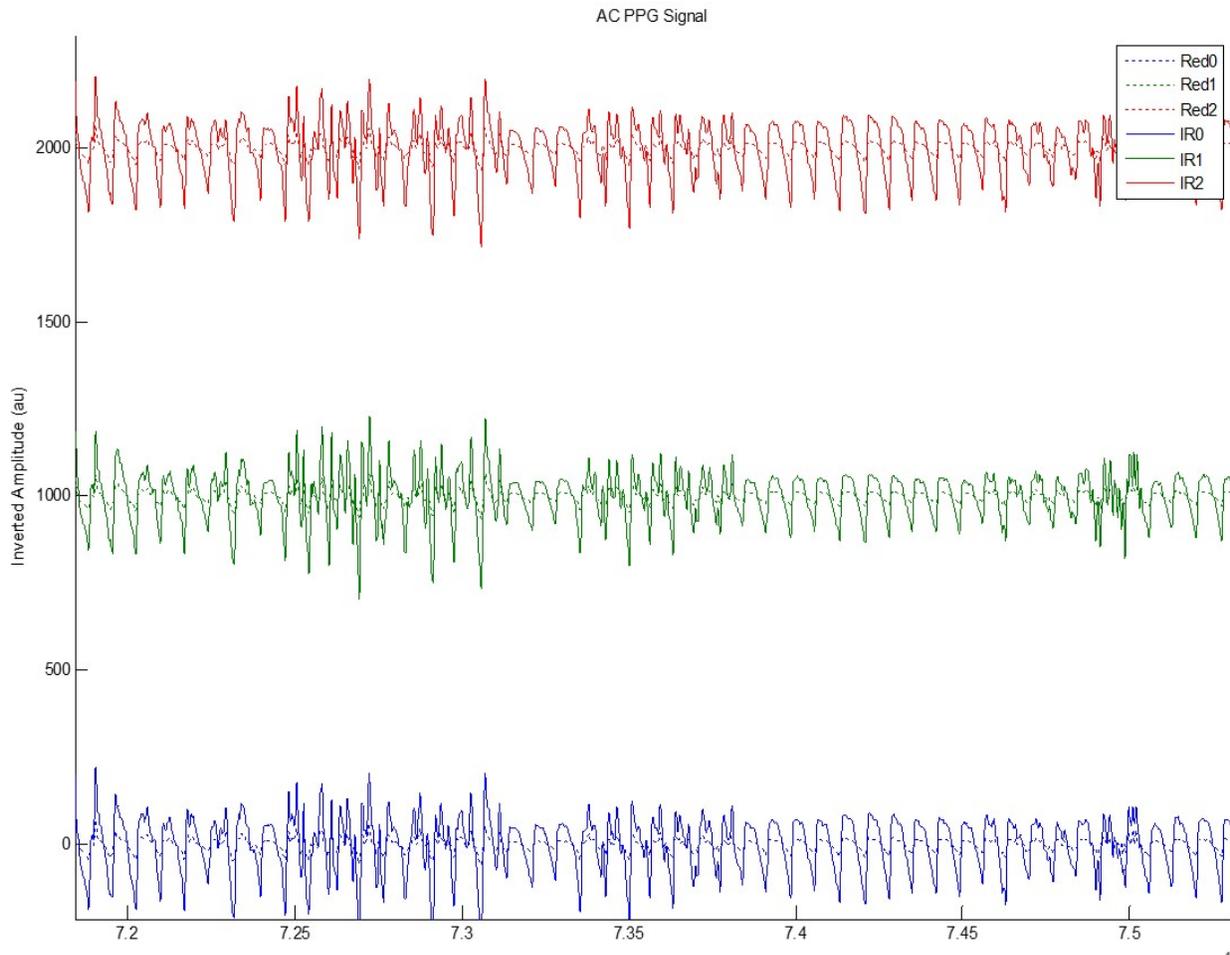
Figure A2. Time series plots of the 3 RD and 3 IR AC IR PPG waveforms from the 3PD finger sensor.

B. Raw Signal Fusion: MCAF/GALL Filters

Silva et. al. created a multichannel adaptive filter (MCAF) algorithm that was implemented by fusing together the raw signals using a series of adaptive filters [36]. One channel is chosen as the reference channel, and the rest of the channels are put through a GALL Filter and then fused together with a Kalman filter. The diagrams of these filters are shown below in Figure 29.
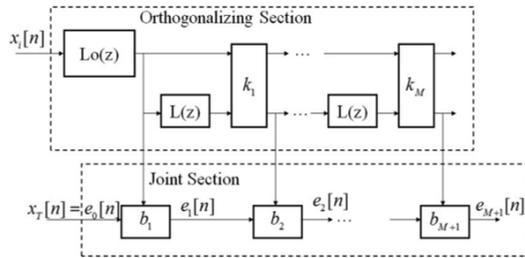
Fig. 3.   Overview of the GALL filter. The GALL filter is similar to a standard gradient adaptive lattice filter, but with delays replaced by Laguerre functions.
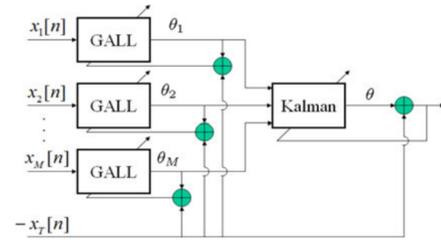
Fig. 2.   Overview of MCAF prediction. Note that the estimates θ_i are scalars on a sample-by-sample basis. The MCAF consists of a bank of GALL filters, in which individual channels attempt to predict the target channel. These individual predictions θ_i are then linearly combined through a Kalman filter for the final estimate.

(a)   (b)

Figure B1. Diagrams of the (a) GALL filter and (b) complete filter bank used in the MCAF algorithm.

This raw signal fusion was implemented on a few of the random motion data sets and then the HR was calculated from each of the six individual PPG waveforms and from the PPG output of the MCAF filter. The inputs and output of this algorithm on a section of random motion are shown in Figure 30.
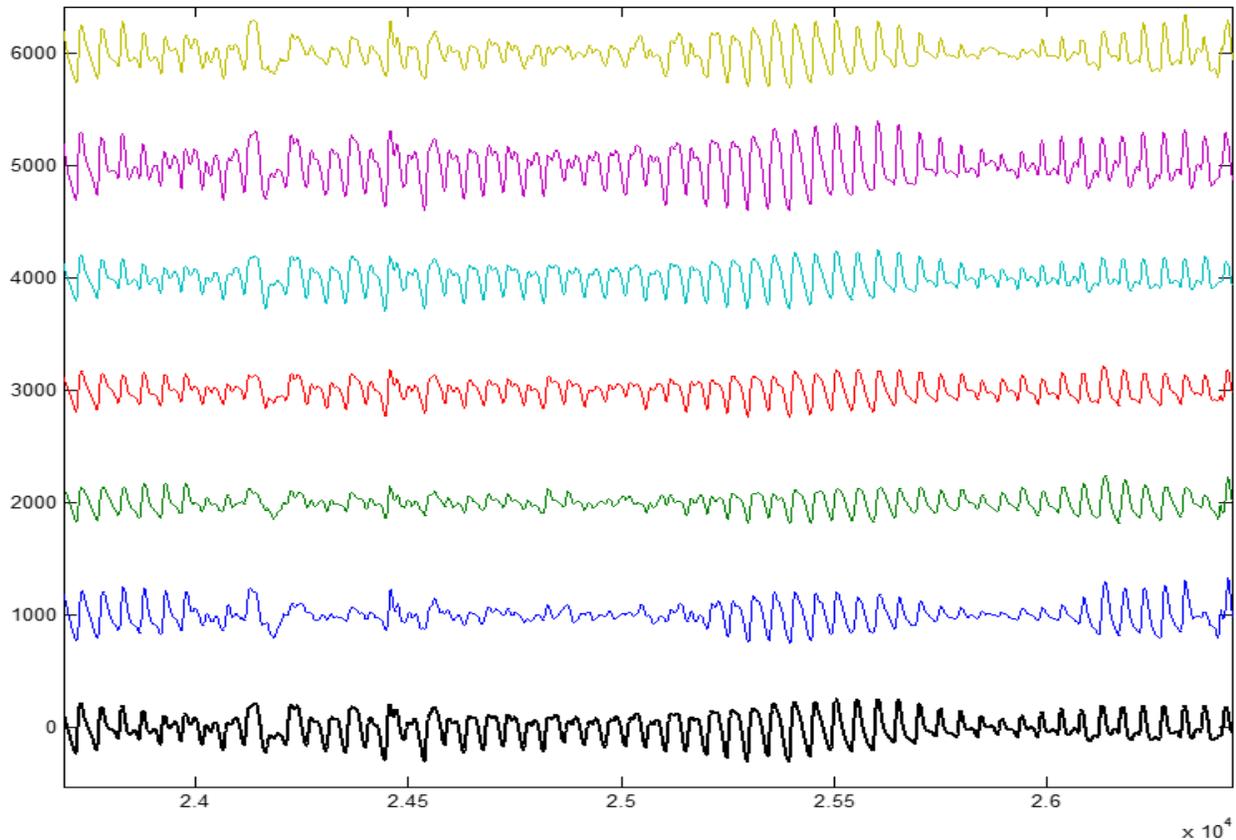
Figure B2. Time series of the AC IR signal from all 6 channels and the output of the MCAF algorithm.

We found that the output did not always reflect the best channel, and at best the output of the raw signal fusion was as good as the best channel. Furthermore, we found that using adaptive filters on multiple raw signals was time consuming, making this algorithm slow and computationally less efficient. In future development of data fusion algorithms for multichannel data, raw signal fusion is not recommended.

C. KALMAN FILTERING

Initial investigation was done on data fusion algorithms to try and fuse HR and SpO$_2$ estimates together from all six channels. This data fusion development was determined to be outside of the scope of this thesis, but initial development was done on a weighted Kalman filter based on the model used in [37]. Implementation of a weighted Kalman filter fuses together estimates from all six channels, and takes past measurements into consideration, theoretically making estimates more accurate. The Kalman update equations are shown below in Equations 17-23.

$$x_k = Ax_{k-1} + Bu_k + w_{k-1} \tag{17}$$

$$z_k = Hx_k + v_k \tag{18}$$

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k \tag{19}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{20}$$

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{21}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \tag{22}$$

$$P_k = (I - K_k H)P_k^- \tag{23}$$

The random variables w and v are independent noise normally distributed with p(w) ~N(0,Q) and p(v) ~N(0,R). In our particular implementation, the measurement noise covariance matrix R was updated based on the MCNL for each measurement. The HR and SpO$_2$ measurements from each channel are then weighted by a combination of the MCNL and the Kalman residual taken after each measurement. The residual, weight and final estimate equations are shown below.

$$R \rightarrow R \cdot \exp\left\{\frac{1}{MCNL^2} - 1\right\} \tag{24}$$

$$r_k = z_k - x_k^- \tag{25}$$

$$\sigma_k^2 = \left(\frac{r_k}{MCNL_k}\right)^2 \tag{26}$$

$$HR = \sum_{k=1}^{6} \left(\frac{\prod_{i=1,\, i \neq k}^{6} \sigma_i^2}{\sum_{i=1}^{6} \left(\prod_{j=1,\, j \neq i}^{6} \sigma_j^2 \right)} \cdot HR_k \right) \qquad \textbf{(27)}$$

$$SpO_2 = \sum_{k=1}^{6} \left(\frac{\prod_{i=1,\, i \neq k}^{6} \sigma_i^2}{\sum_{i=1}^{6} \left(\prod_{j=1,\, j \neq i}^{6} \sigma_j^2 \right)} \cdot SpO_{2_k} \right) \qquad \textbf{(28)}$$

While implementation of the Kalman filter did not change the final results significantly for all data sets, it was quite beneficial for other data sets, such as the data sets shown in Figures C1 and C2. Figure C1 shows HR estimations before and after the Kalman filter was used for data set 28 and Figure C2 shows SpO_2 estimations before and after the Kalman filter was used for data set 13. It is clear that the Kalman filter smoothed HR estimates for data set 28, making the output more stable and less sporadic. For the SpO_2 estimates in data set 13, the SpO_2 estimates from the multichannel Kalman filter were much better than the SpO_2 estimates calculated just by switching channels.
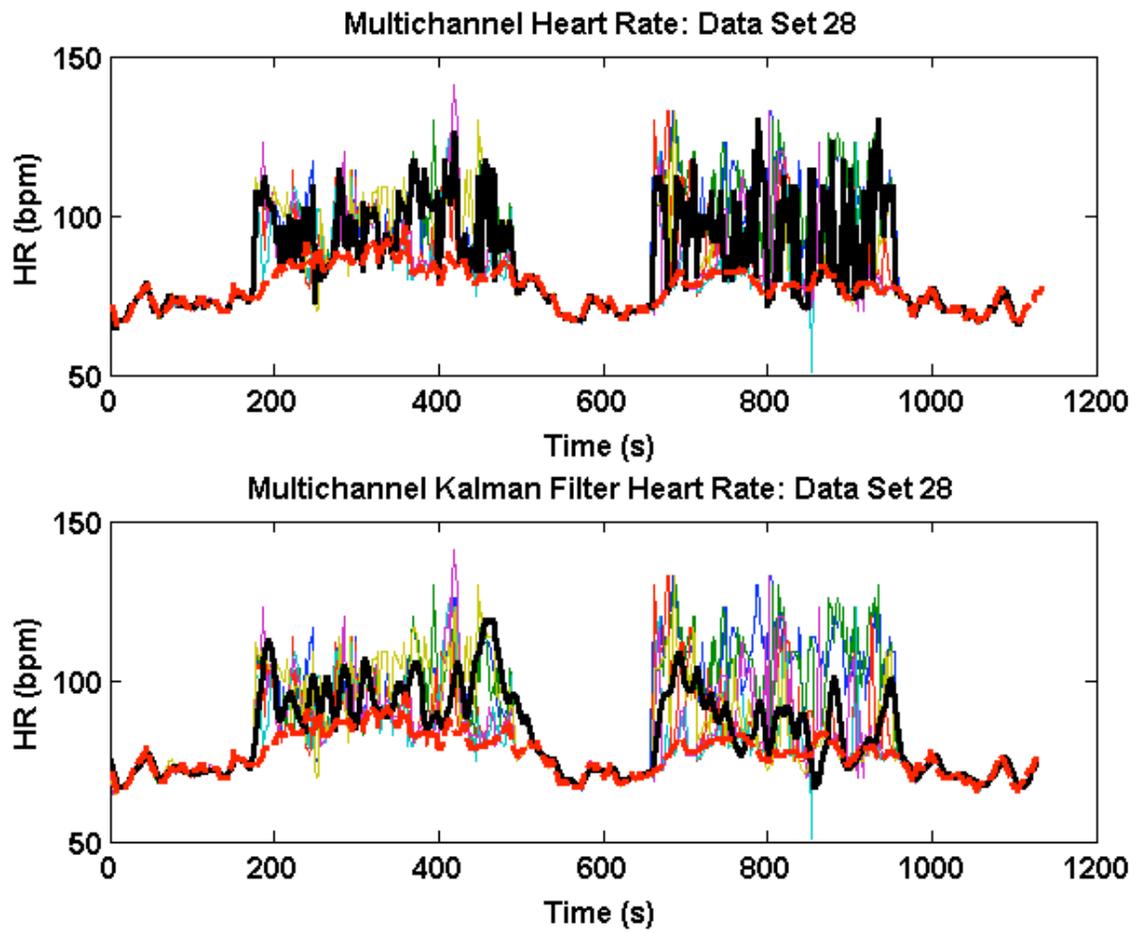
Figure C1. Multichannel estimates of before and after implementation of an MCNL-weighted Kalman Filter for HR measurements from data set 28.
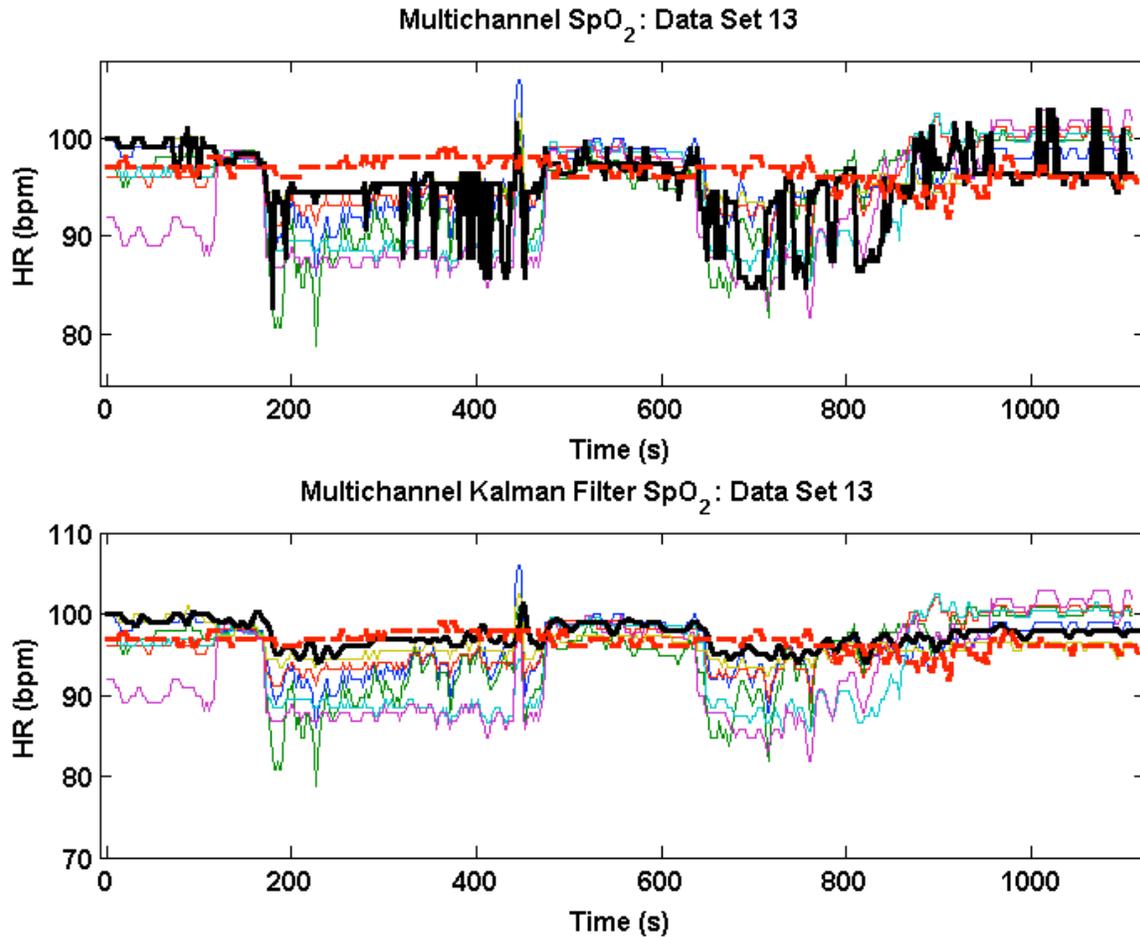
Figure C2. Multichannel estimates of before and after implementation of an MCNL-weighted Kalman Filter for SpO₂ estimates from data set 13.

The exact specifications of the multichannel MCNL-weighted Kalman filter were not fully explored in this thesis, but I believe that this algorithm would further improve multichannel estimates with a bit of further development. The MCNL weights ensures that the measurements are still based on the signal morphology of the PPG, but the fusion of all 6 channels' estimates and the consideration of previous measurements would prevent extraneous high or low measurements when the multichannel-switching algorithm incorrectly chooses a channel that produces high error measurements.

D. NOTCH FILTERING

Spectrograms of all 31 data sets were plotted for the entire duration of the 19-minute segment of alternating rest and motion periods. We noticed that for the majority of data sets, the motion frequency was easily identifiable from the HR frequency, and that the primary motion frequency remained relatively constant during both motion segments. Figure D1 shows an example spectrogram for two different channels from the same data set. It can be clearly seen that the motion frequency appears during the motion segments around 2 Hz.
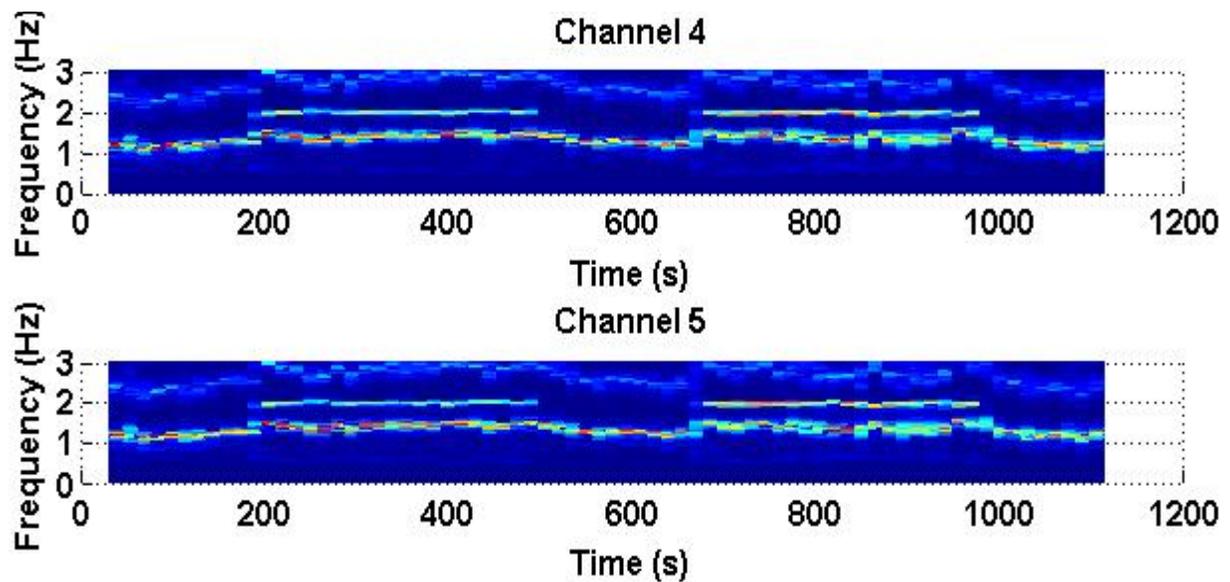
Figure D1. Spectrogram of channel 4 and channel 5 for the duration of one data set.

Since the motion frequency was visibly separate from the HR frequency, a notch filter at the primary motion frequency with a bandwidth of 0.4 Hz was used on the entire data set to decrease the effects of motion artifact on all six channels. PPG waveforms before and after implementation of the notch filter are shown in Figure D2. Post-filtering, the PPG looks much cleaner, and the peaks are more easily identifiable.

85
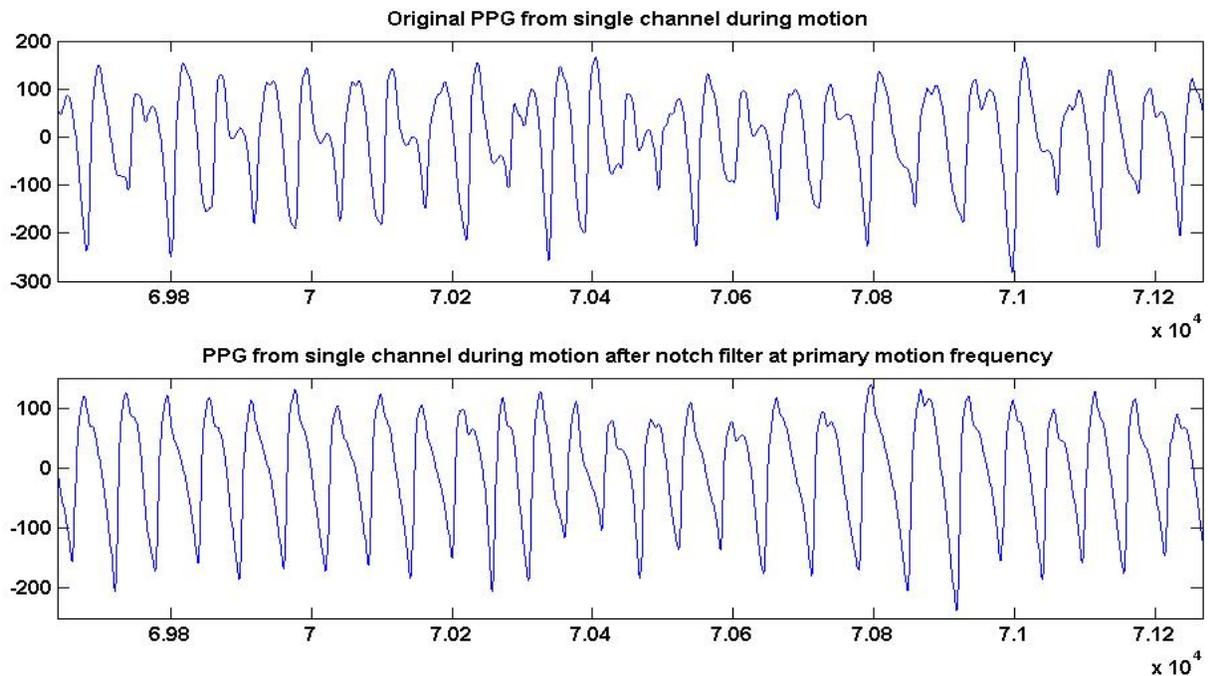
Figure D2. Time series of AC IR PPG before implementing a notch filter (top) and after implementing a notch filter (bottom) at the primary motion frequency.

Two examples of multichannel HR estimates taken from two different data sets before and after implementation of a notch filter at the primary motion frequency are shown in Figures D3 and D4. For data set 16, before implementation of the notch filter, some channels had high error in HR and some had low error in HR, making the multichannel estimate better than any individual channel in its mean relative HR error. After the notch filter was applied, all channels had a relatively low error. For data set 21, which was deleted from the statistic calculations because all six channels had a high HR error, after the notch filter was applied, the HR estimations across channels looked more similar to data sets where the multichannel estimate significantly outperformed the individual channel estimates and the signal was no longer overpowered by the motion frequency in such a way where the data were unusable.
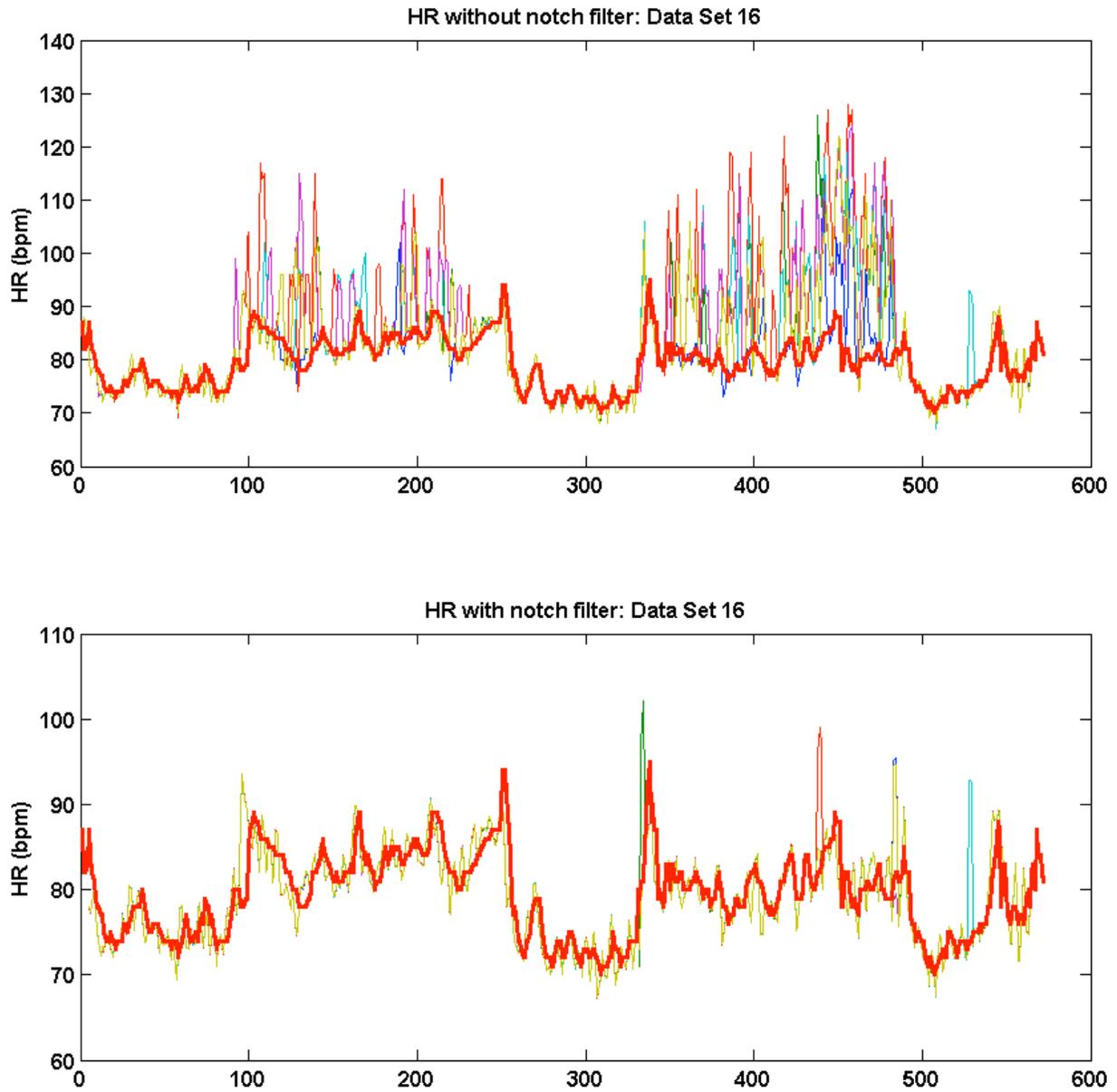
Figure D3. Time-series heart rate plots before (top) and after (bottom) implementation of a notch filter at the primary motion frequency for data set 16. Masimo reference HR is plotted in red.
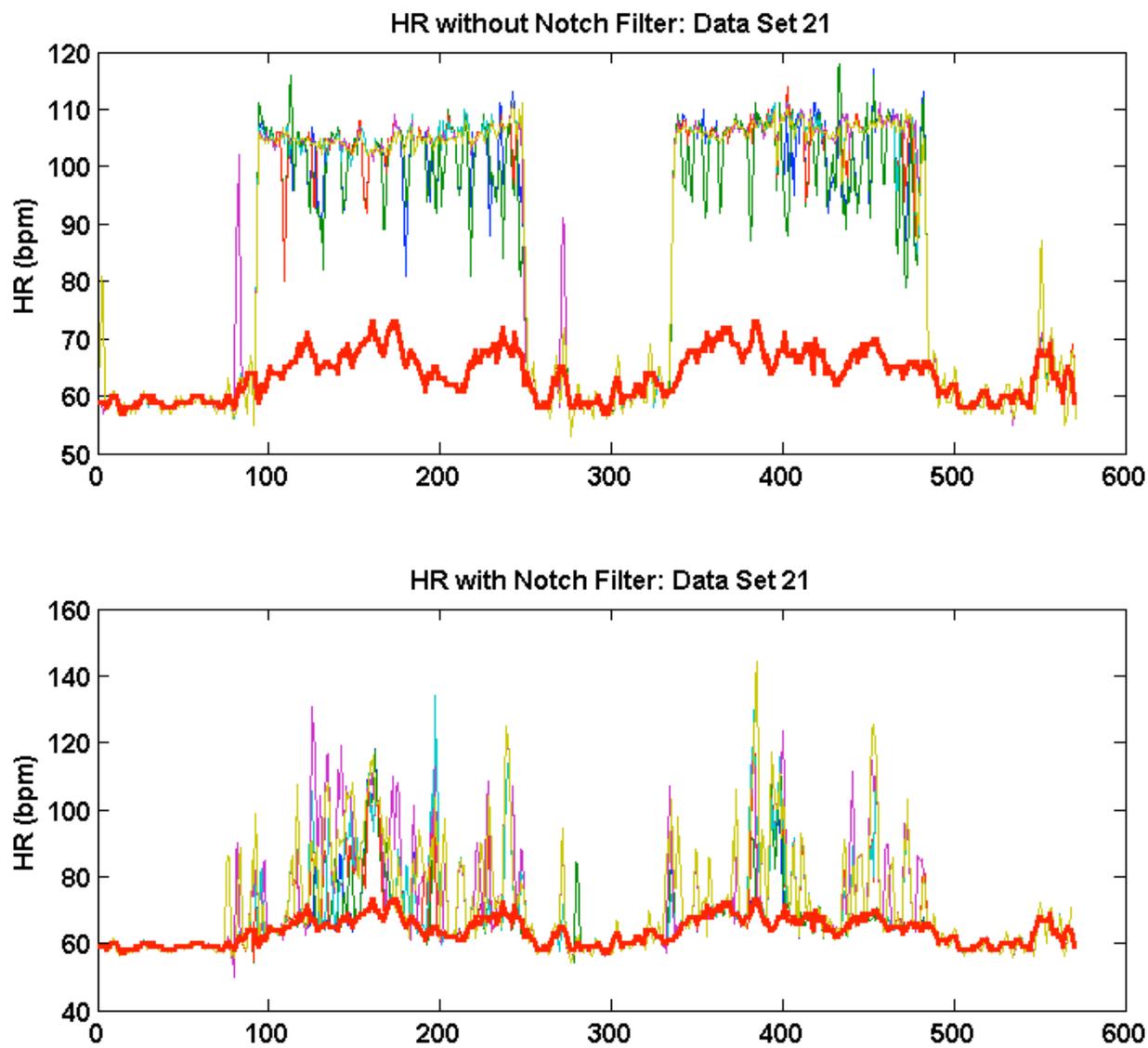
Figure D4. Time-series heart rate plots before (top) and after (bottom) implementation of a notch filter at the primary motion frequency for data set 21. Masimo reference HR is plotted in red.

Based on this initial observation, we believe that implementation of a real-time notch filter based on the motion frequency as seen in the accelerometer data could further improve the benefits of multichannel estimates during motion and allow the multichannel switching to cover a larger range of motion artifacts.

88

E. CONTROLLED MOTION

After initial multichannel analysis was performed, we thought it was possible that the channel that remained low in HR or SpO_2 error was dependent on the axis of motion, and/or the type of motion that was introduced during the recordings. To test this hypothesis, five different trials of left-right, up-down, and circular head motion were performed for one minute with the 6PD forehead sensor. The Masimo finger sensor was used for HR reference. SpO_2 measurements were compared against the SpO_2 estimates from the multichannel device calculated during rest for each channel. For HR, the channel that remained low in error changed between trials, and was not consistent for the same motion between all five trials. This led to the conclusion that the channel with the lowest error was not dependent on the type or axis of motion, but on other factors. Figure E1 depicts the percent HR error for all five trials across all 6 channels for each of the three head movements and during rest.
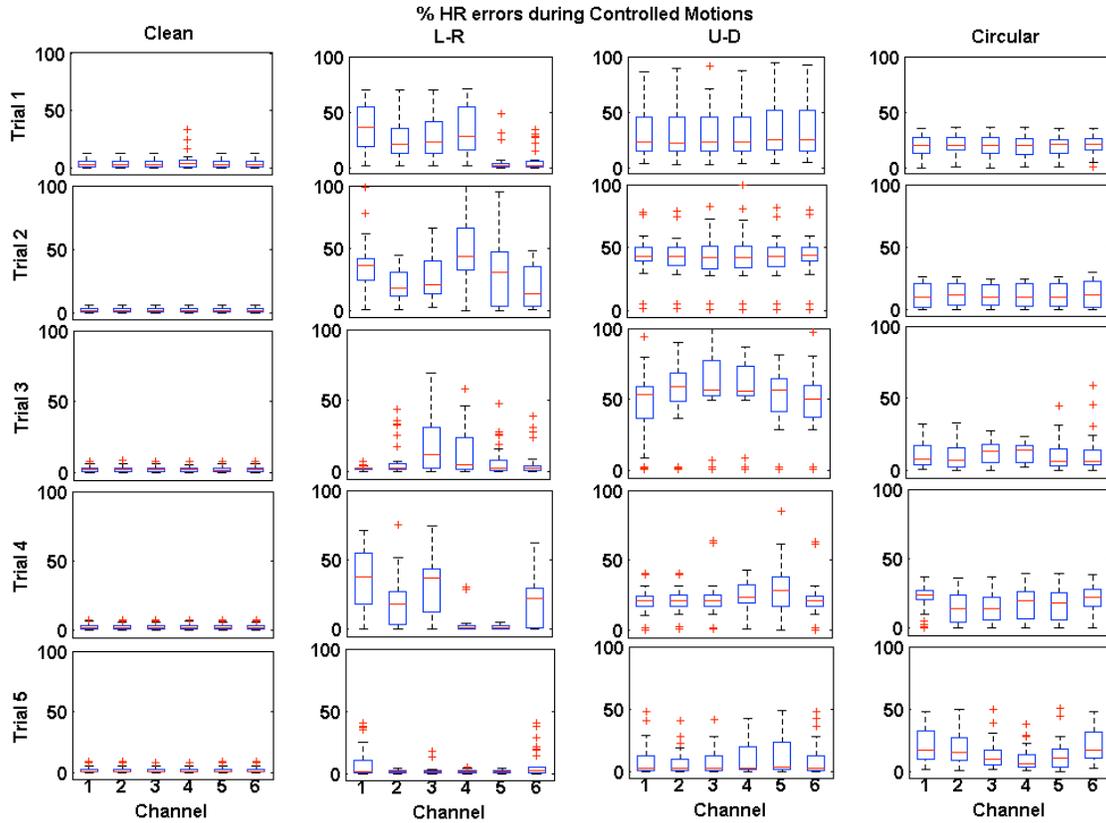
Figure E1. Percent HR errors for 5 trials of controlled motions across 6 channels recorded by the 6PD forehead sensor.

For SpO$_2$ error, we found that the channel with the highest and lowest errors did not fluctuate very much at all between channels across trials, but the low and high error channel did not fluctuate between head motions either. From this, we concluded that the channel with the lowest or highest error for SpO$_2$ is also not dependent on the type or axes of motion, but upon other factors. Figure E2 shows the percent SpO$_2$ error for all five trials for all six channels in during the three types of head movements and during rest.
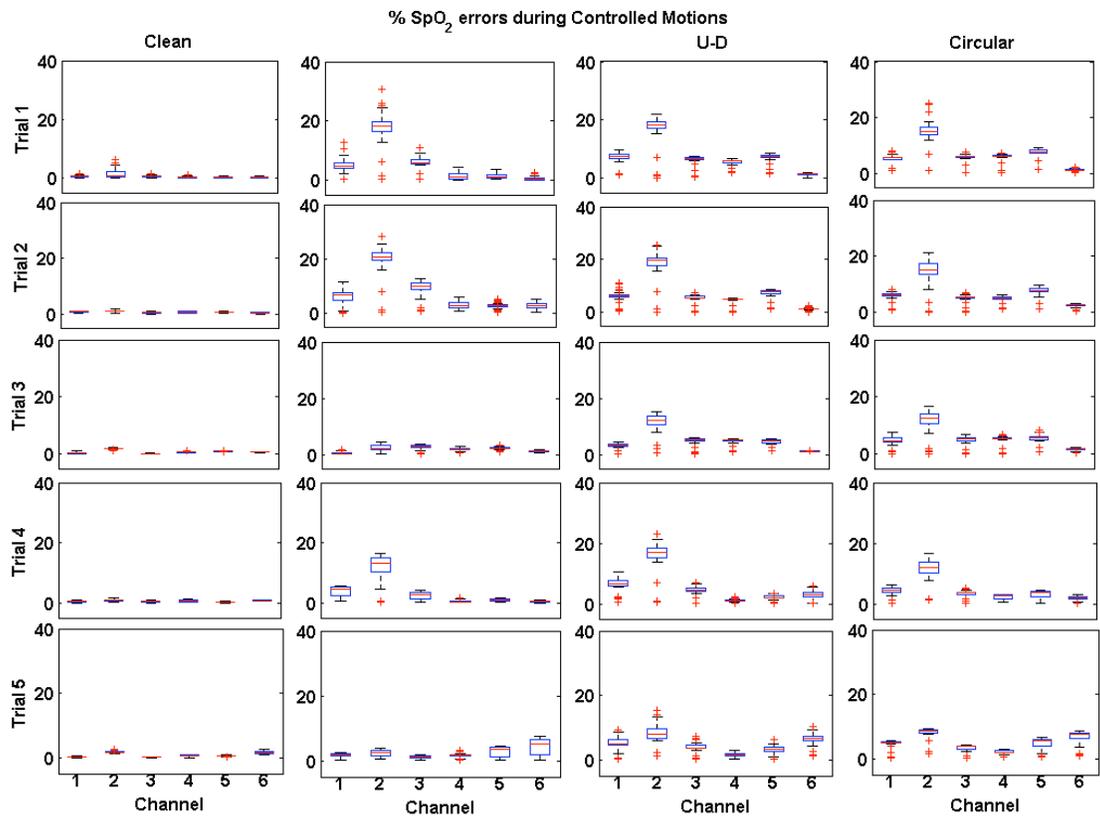
Figure E2. Percent SpO₂ errors for 5 trials of controlled motions across 6 channels of the 6PD forehead sensor.

F. Treadmill Data

Our colleagues at the University of Connecticut took data from 10 different subjects walking at three different speeds on a treadmill while wearing the 6PD forehead sensor, a Masimo finger sensor, and an ECG monitor. In order to see how the multichannel data and algorithm responded to repetitive treadmill motion, the multichannel template-matching algorithm was applied to the data. For this data, the periodic motion induced by the walking or running on the treadmill caused the template to start taking the shape of the treadmill motion, leading the algorithm to incorrectly pick the channel most affected by the motion frequency instead of the channel with the correct HR frequency. In order to combat this shortcoming, for the treadmill data only, the algorithm was updated such that the template could not be updated if the RMS of the accelerometer exceeded 50 au. This allowed the algorithm to choose the "best" channel based on the template created during rest.

We found that in data sets 5 and 8, the multichannel estimate outperformed the single channel estimates while the subject was running on the treadmill; some channels remained in the low-error HR range during treadmill motion while other channels remained high. The results were plotted in Figures F1 and F2. The multichannel estimate is shown in black, the Masimo reference is shown in dashed red, and the ECG HR estimate is shown in blue. The accelerometer data is shown below the HR estimates as a motion reference.
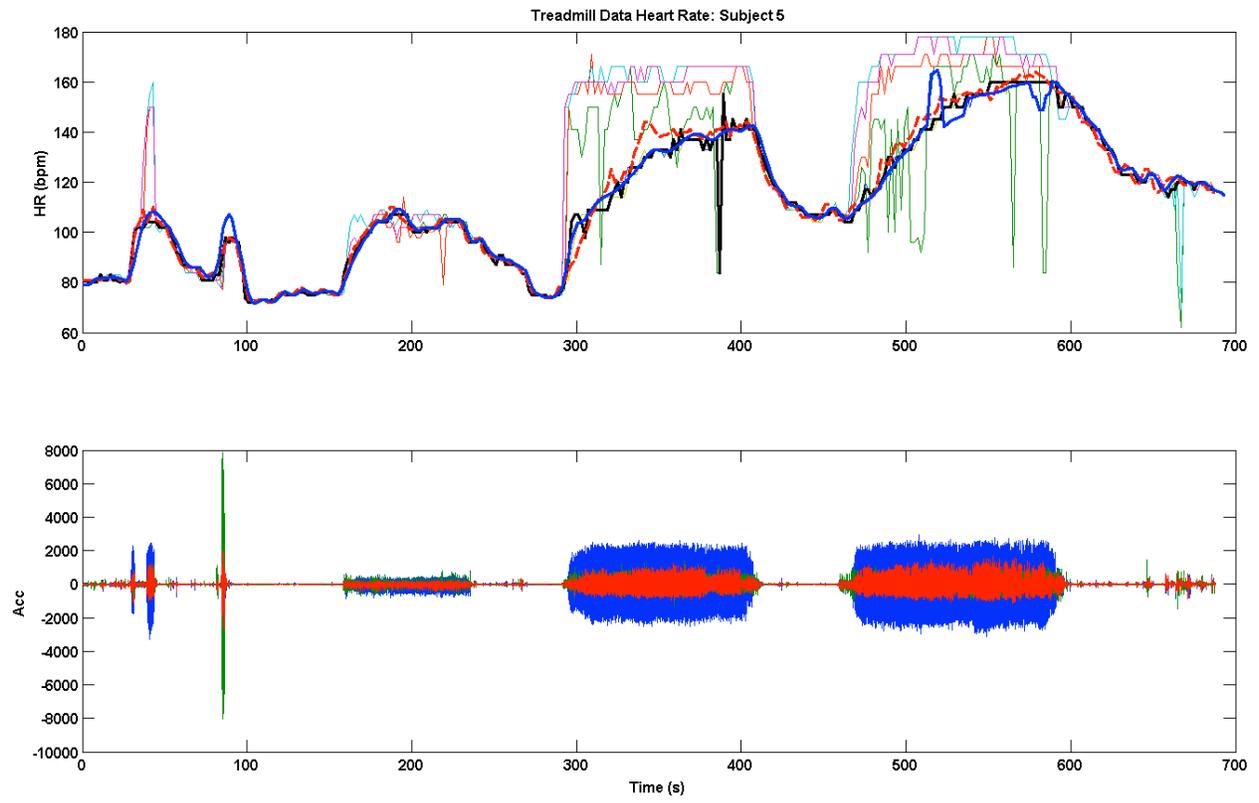
Figure F1. Treadmill data from subject 5 taken from the 6PD forehead sensor.
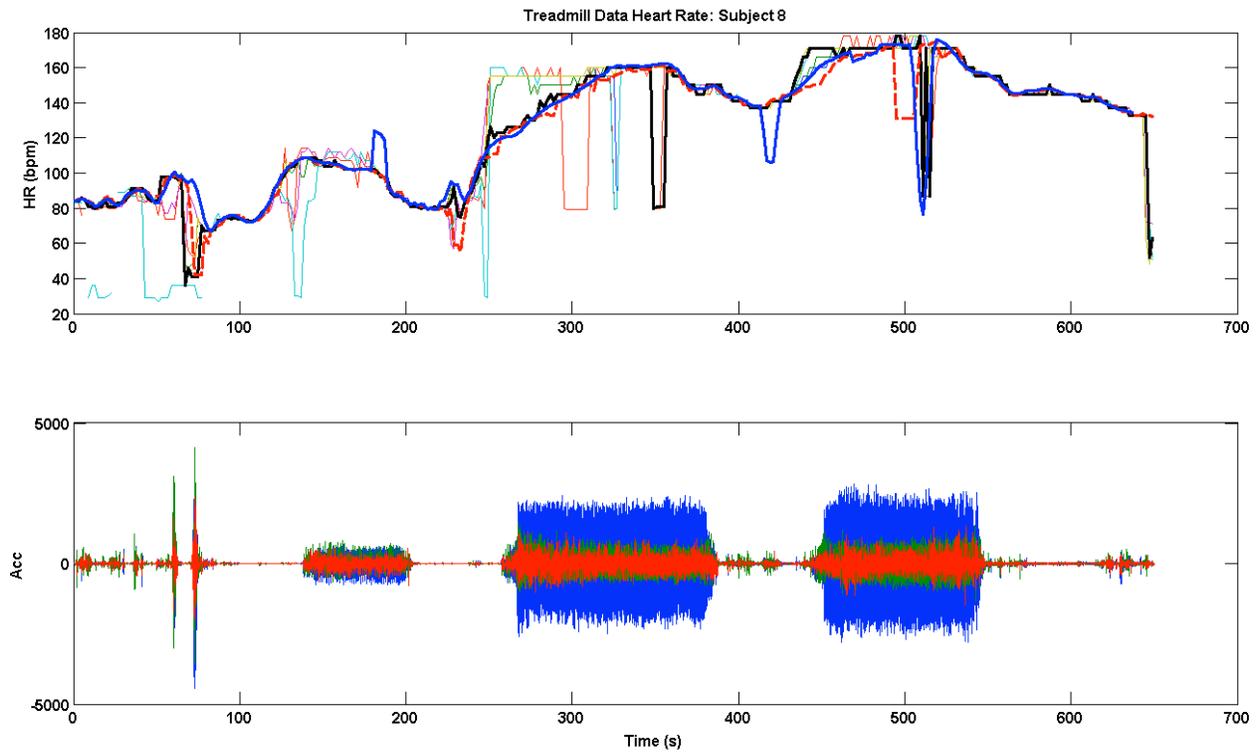
Figure F2. Treadmill data from subject 8 taken from the 6PD forehead sensor.

Similar to the random motion data, the multichannel data did not always vary in HR error during motion, and sometimes all six channels were overwhelmed with excessive motion frequency, thus preventing any channel from measuring a correct HR estimate during motion. These types of data sets in the treadmill data are shown in Figures F3 and F4.
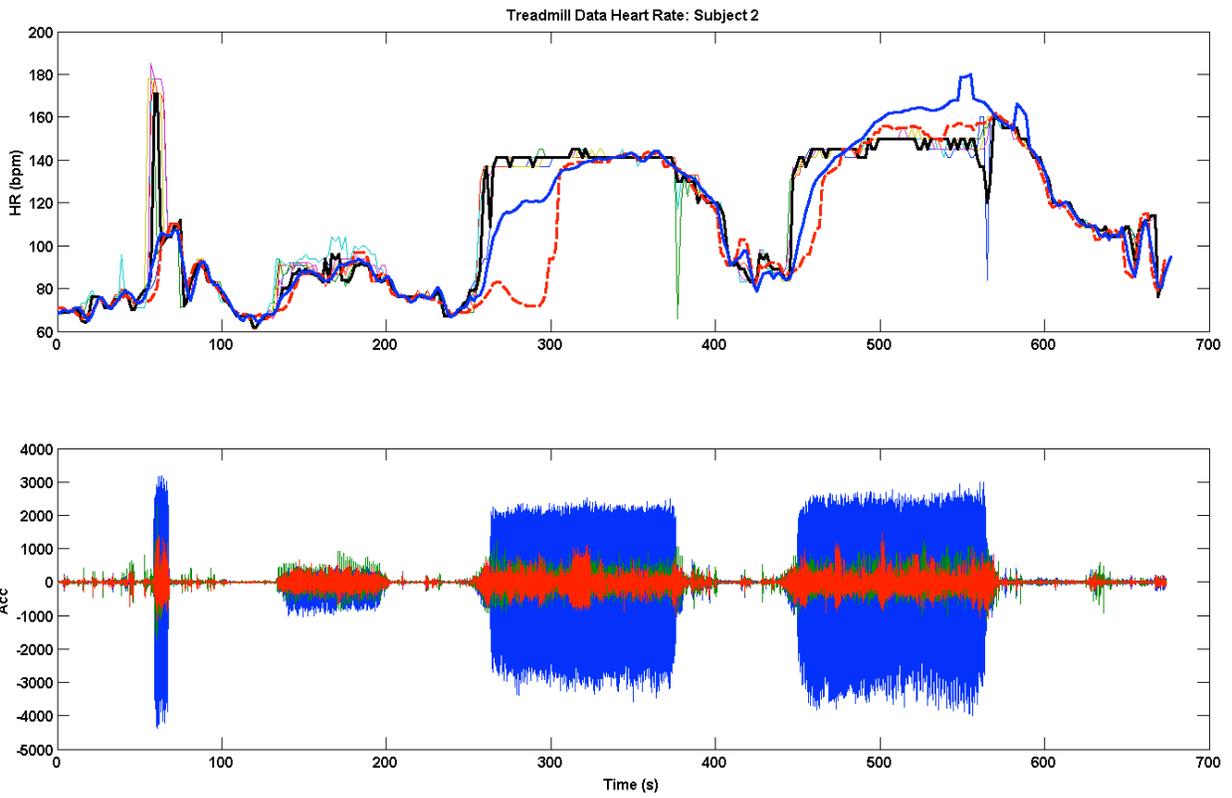
Figure F3. Treadmill data from subject 2 taken from the 6PD forehead sensor.
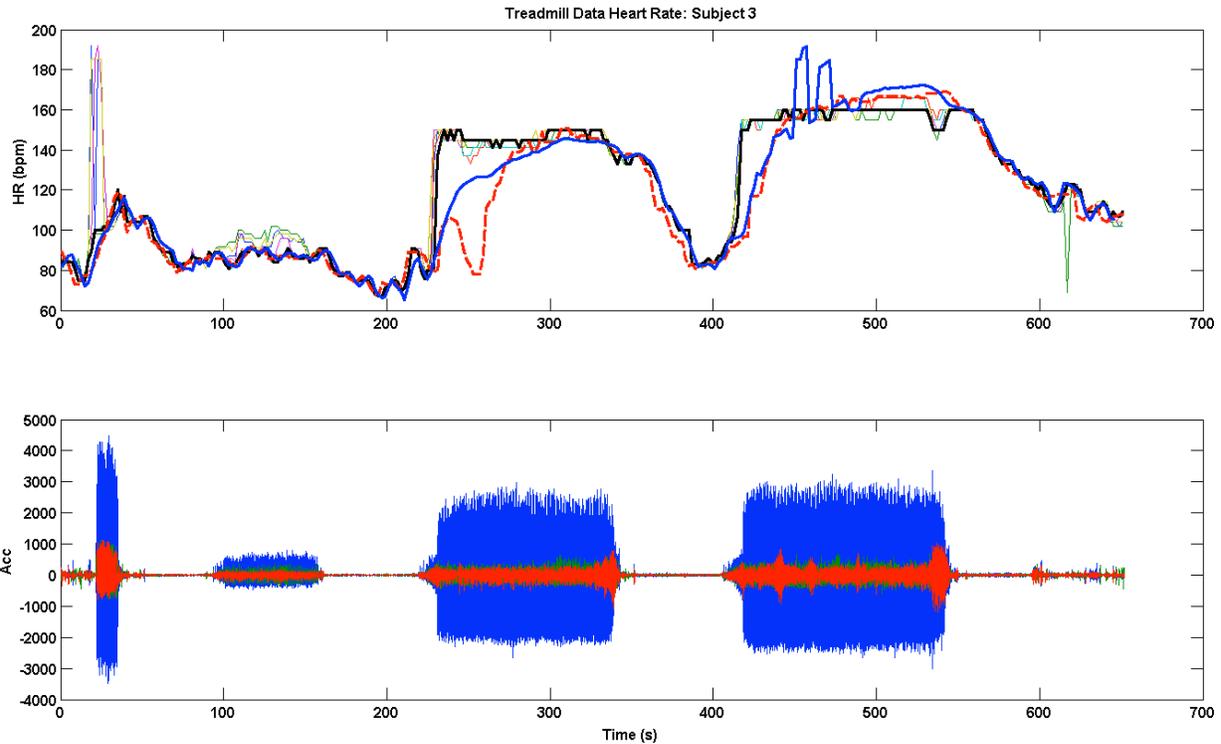
Figure F4. Treadmill data from subject 3 taken from the 6PD forehead sensor.

We concluded that the benefits of the multichannel data were less apparent in data sets similar to the sample data shown above. The HR errors calculated for all 10 subjects from the treadmill data are summarized in Tables 8 and 9.

Table F1. Performance Index for each individual channel and for the multichannel estimate (MC) taken from the treadmill data.

Performance Index: Percentage of HR error less than or equal to 5 bpm [%]

Subject	Ch 1	Ch 2	Ch 3	Ch 4	Ch 5	Ch 6	MC
1	64.6%	63.1%	66.2%	69.2%	67.5%	71.7%	69.4%
2	68.8%	67.4%	70.0%	63.2%	69.1%	71.2%	66.8%
3	59.5%	54.9%	71.5%	69.6%	62.0%	57.7%	66.6%
4	84.2%	84.0%	73.8%	75.6%	76.2%	81.8%	80.0%
5	86.9%	70.3%	64.0%	57.0%	59.6%	86.9%	86.6%
6	55.6%	64.3%	54.0%	29.9%	42.4%	52.7%	46.3%
7	68.8%	67.9%	64.5%	57.6%	63.2%	63.9%	69.2%
8	85.8%	74.8%	66.5%	62.5%	69.5%	74.5%	83.1%
9	63.5%	64.3%	62.1%	57.3%	60.5%	60.3%	64.3%
10	53.0%	62.8%	82.2%	73.1%	80.2%	69.2%	79.4%
Mean	**69.1%**	**67.4%**	**67.5%**	**61.5%**	**65.0%**	**69.0%**	**71.2%**
Mean Diff	2.1%	3.8%	3.7%	9.7%	6.1%	2.2%	
Std Err Diff	3.1%	3.3%	3.0%	3.2%	2.7%	1.9%	
t-stat	0.677	1.142	1.211	3.050	2.279	1.148	
t-crit	1.812						

96

The Performance Index for the multichannel treadmill data was higher than any individual channel, but only statistically significantly higher than channels 4 and 5.

Table F2. Mean HR error for each individual channel and for the multichannel estimate for ten subjects treadmill data compared against the Masimo finger sensor.

	Mean HR error across entire data set [bpm]						
Subject	**Ch 1**	**Ch 2**	**Ch 3**	**Ch 4**	**Ch 5**	**Ch 6**	**MC**
1	15.9	15.0	11.2	10.3	12.1	8.8	10.8
2	9.0	9.7	10.1	11.3	10.1	9.7	9.6
3	9.5	9.1	7.1	7.1	9.3	9.9	7.7
4	3.3	3.5	5.2	5.0	4.5	3.9	4.2
5	2.6	8.0	9.4	12.7	11.5	2.4	2.7
6	16.4	10.8	16.5	32.6	16.4	14.3	21.9
7	12.6	12.7	13.2	14.7	13.5	13.6	12.7
8	5.2	7.5	11.3	12.5	7.7	7.3	6.1
9	14.6	12.5	13.7	16.7	14.1	16.7	13.4
10	24.9	21.8	4.0	6.7	5.7	15.1	6.0
Mean	**11.4**	**11.1**	**10.2**	**13.0**	**10.5**	**10.2**	**9.5**
Mean Diff	1.9	1.6	0.7	3.5	1.0	0.7	
Std Err Diff	2.1	2.1	1.1	1.3	1.1	1.3	
t-stat	0.921	0.735	0.635	2.632	0.912	0.507	
t-crit	1.812						

The mean relative error, or accuracy, of the multichannel estimate for the treadmill data was lower than any individual channel across 10 subjects, however, the multichannel estimate was only statistically lower than channel 3.

G. Masimo HR vs ECG

ECG data were taken for the last 7 of the 31 random motion data sets used for the HR reference during motion, data sets 25 through 31. Since the first 24 data sets were taken with just the Masimo finger sensor for HR reference, and it was important to include these 24 data sets in the analysis done in this thesis, the correlation between the ECG HR measurements and Masimo HR measurements for our motion protocol was measured and plotted. Three example correlations between the Masimo and the ECG HR measurements are shown in Figure G1.

(a)   **(b)**   **(c)**
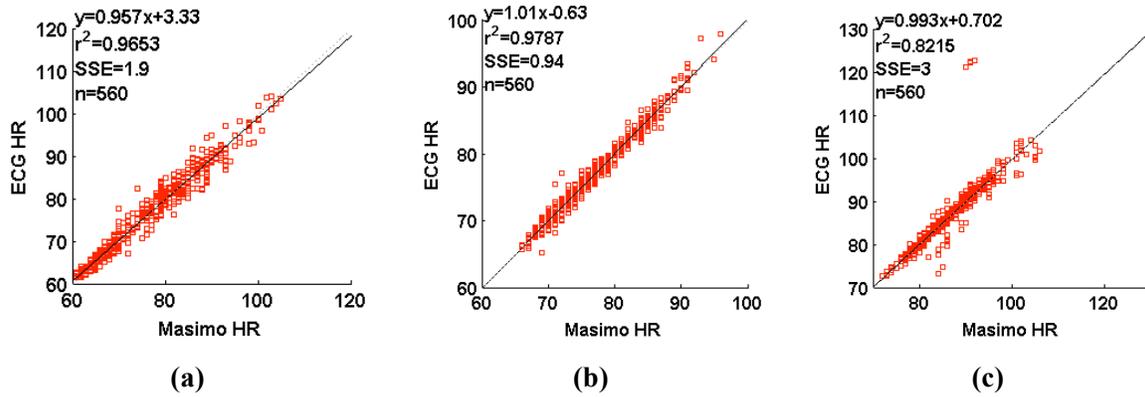
Figure G1. Correlation between Masimo HR estimates and ECG HR estimates throughout the duration of three different data sets: (a) data set 26 (b) data set 28 and (c) data set 30.

The remaining four data sets where ECG data were collected have similar correlations between the Masimo HR estimates and the ECG HR estimates. To show this more clearly, the time series of both HR estimations are shown in Figure G2.
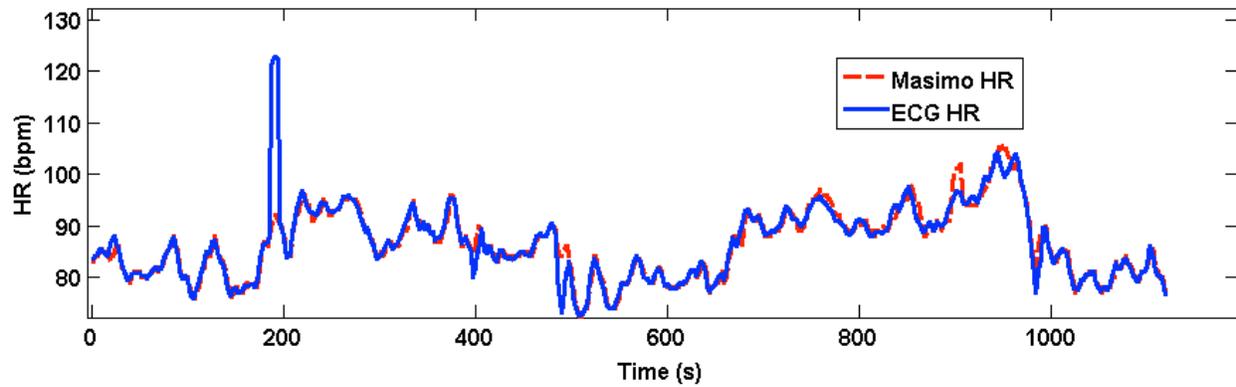
Figure G2. Time series HR estimates from the Masimo sensor and the ECG data during the duration of one complete data set.

Although there are occasional differences between the Masimo HR and the ECG estimation for these seven data sets, because of the high correlation between the two HR estimates and how closely the Masimo HR estimation generally follows the ECG HR estimation, we determined that it was sufficient to calculate HR errors by comparing the multichannel HR estimates to the Masimo pulse oximeter as a gold standard reference instead of the ECG derived HR estimates.

H. MULTICHANNEL HR BLAND-ALTMAN PLOTS

To compare how HR measurements differ across channels during motion, for the 24 data sets with only the Masimo HR reference, Bland-Altman and regression plots were created for each data set for each channel, and for the multichannel estimate. These Bland-Altman plots measure the bias and precision of the HR measurements from our device compared to the HR measurements from the Masimo reference sensor.
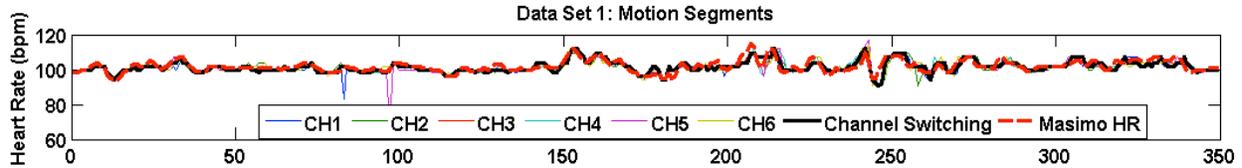
Figure H1. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 1.
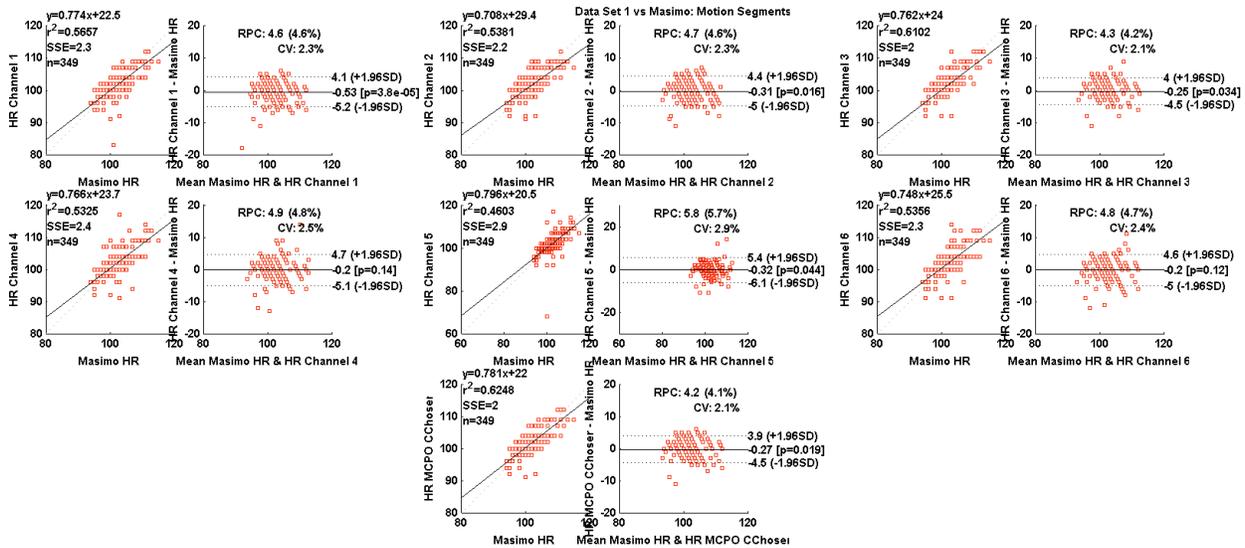
Figure H2. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 1.
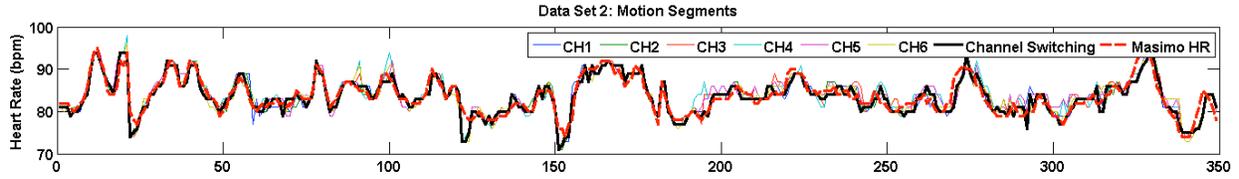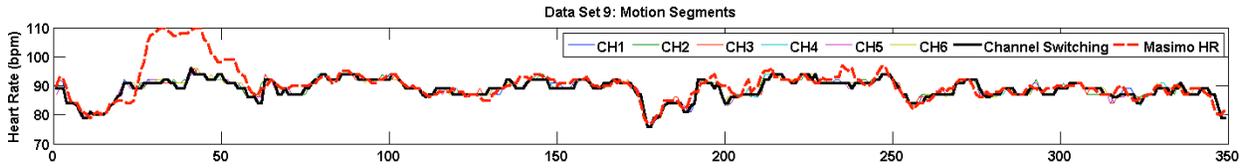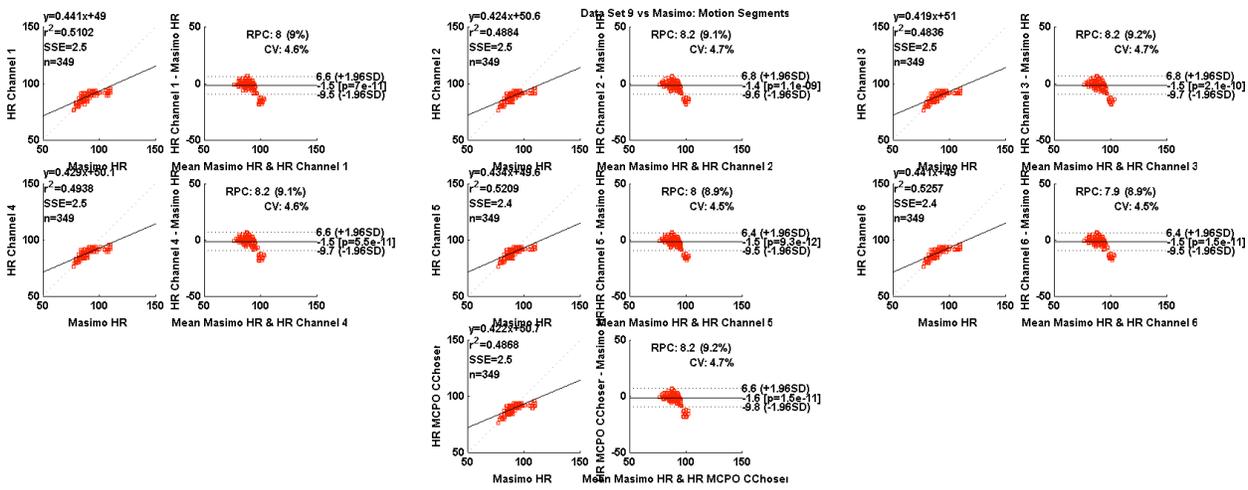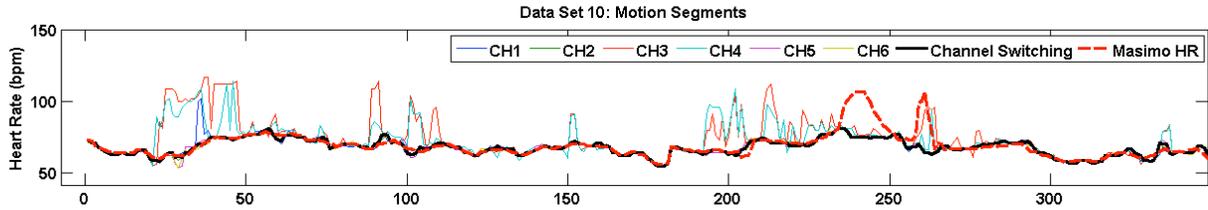
Figure H3. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 2.
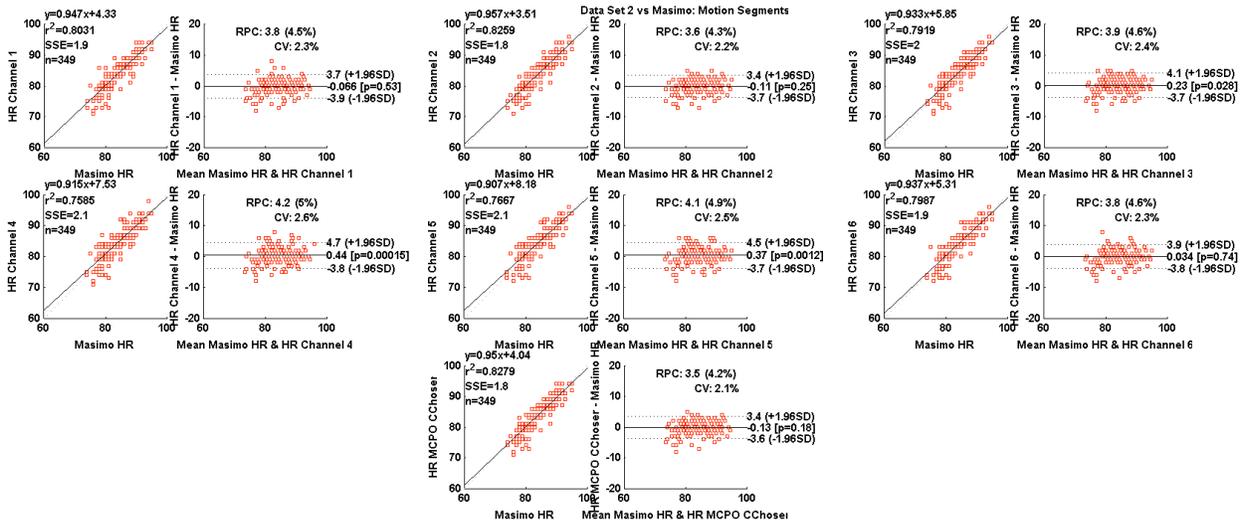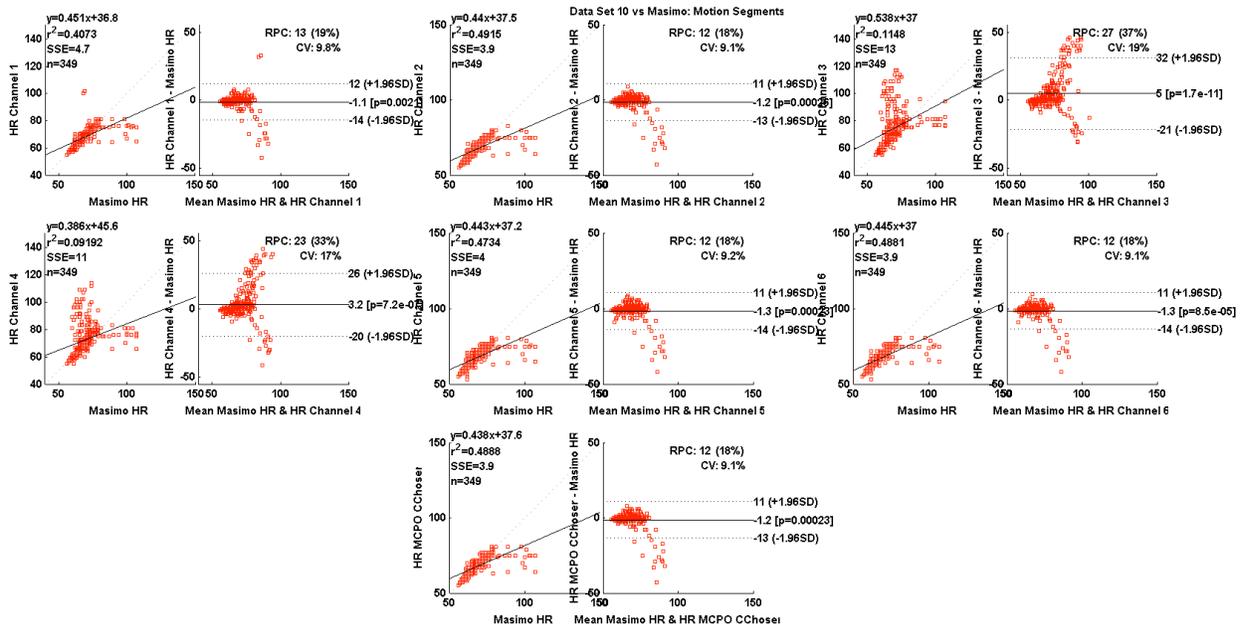

Figure H4. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 2.
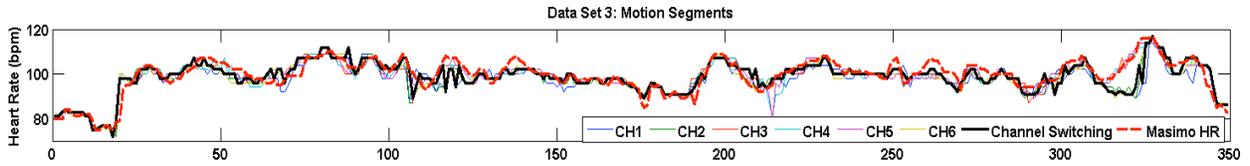
Figure H5. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 3.
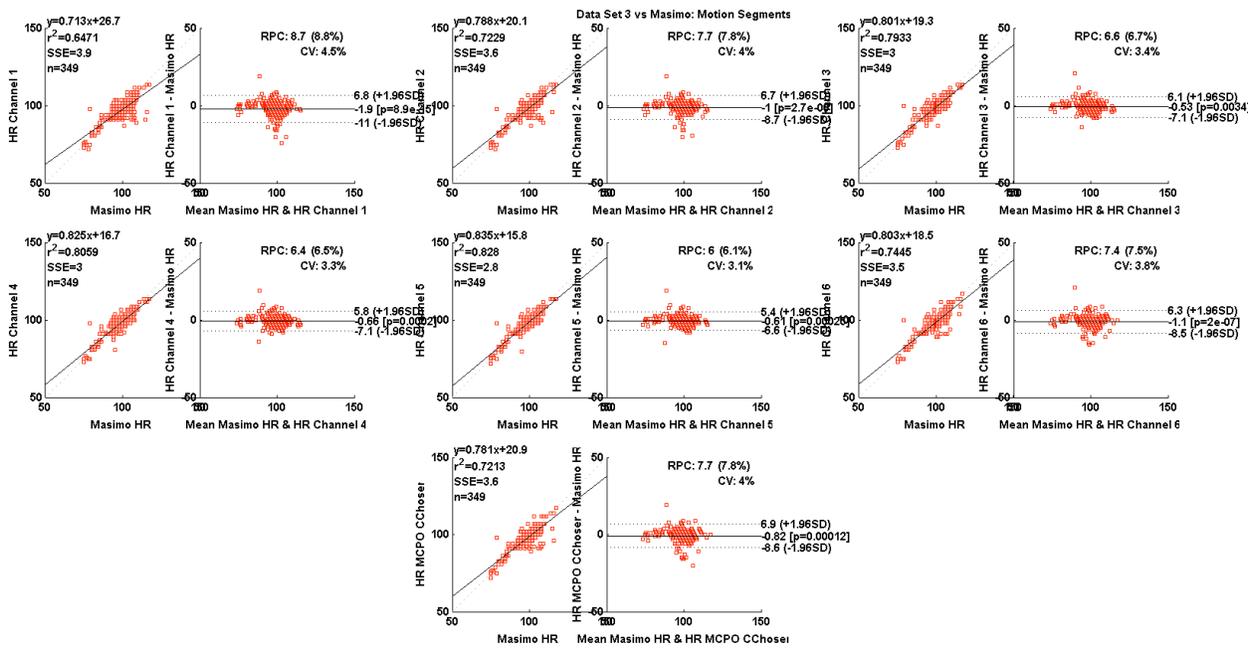
Figure H6. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 3.
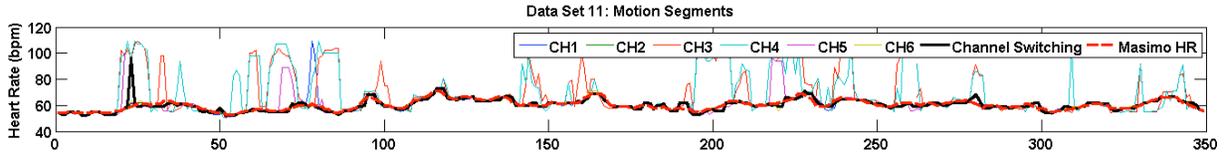
102
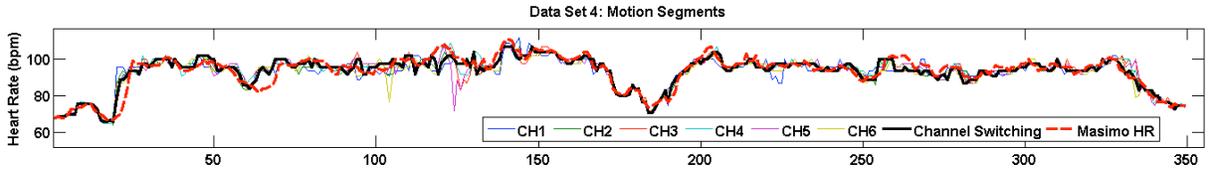
Figure H7. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 4.
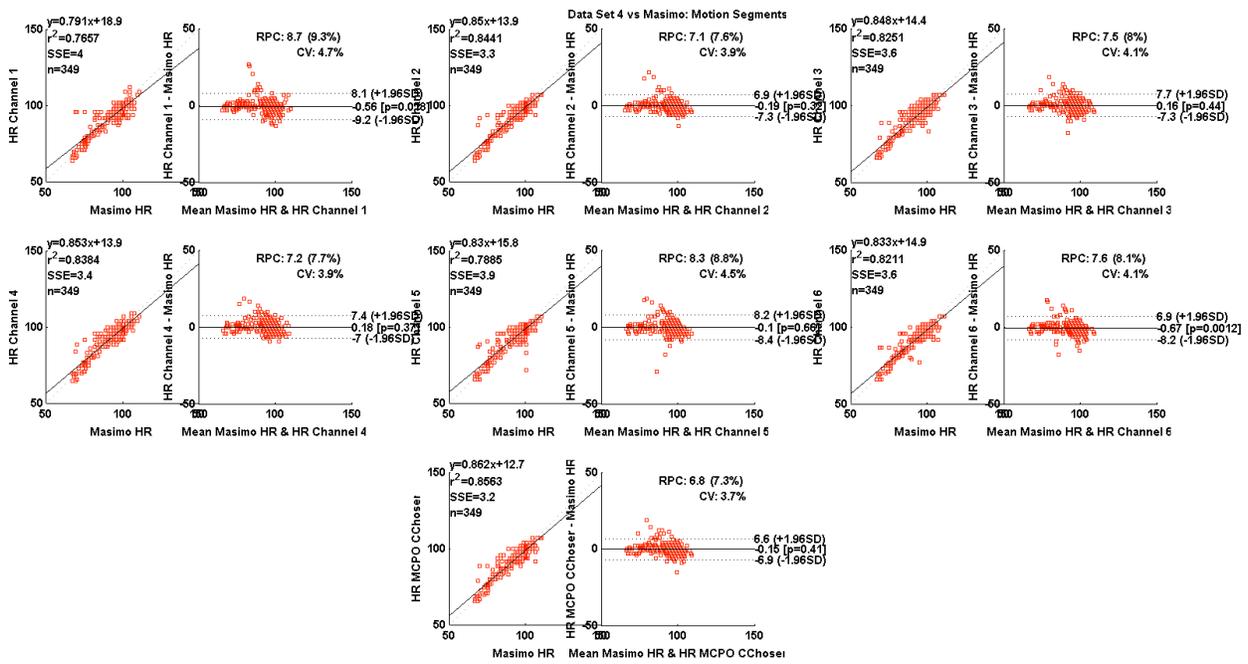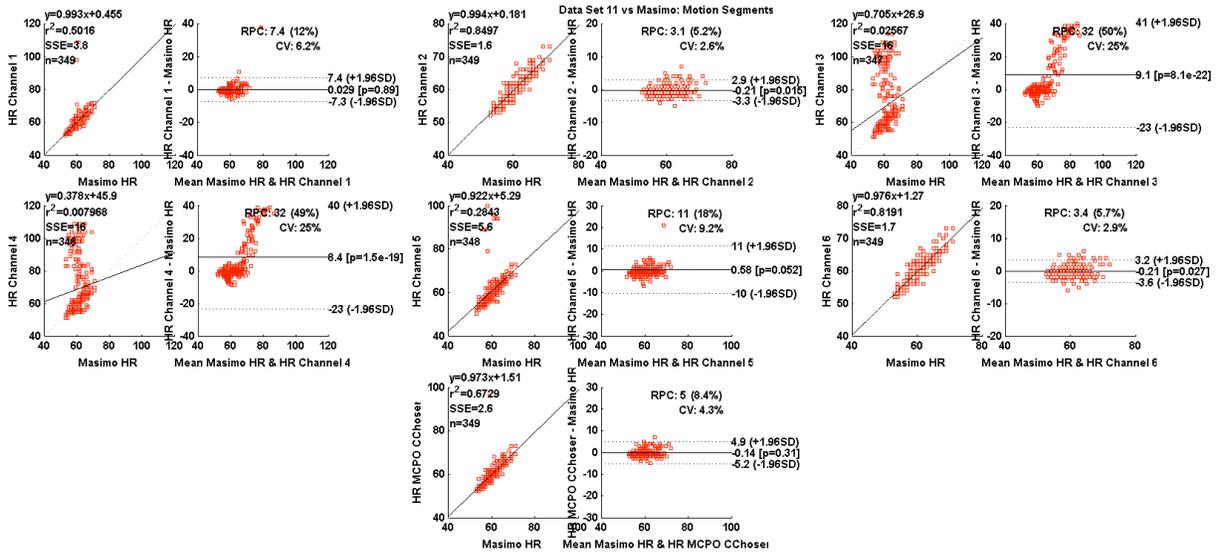

Figure H8. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 4.
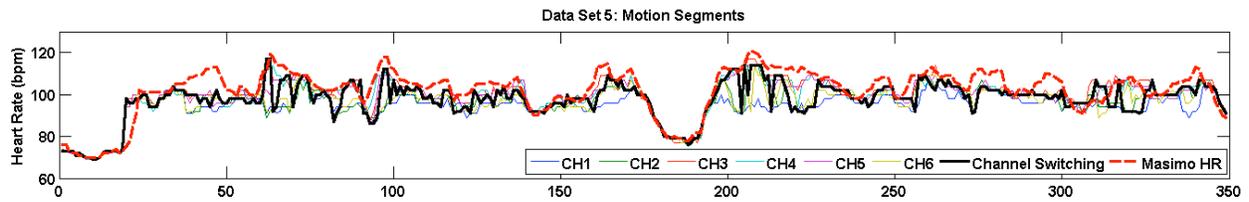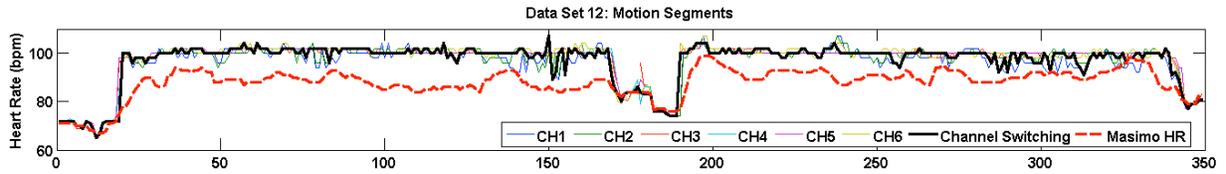
Figure H9. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 5.
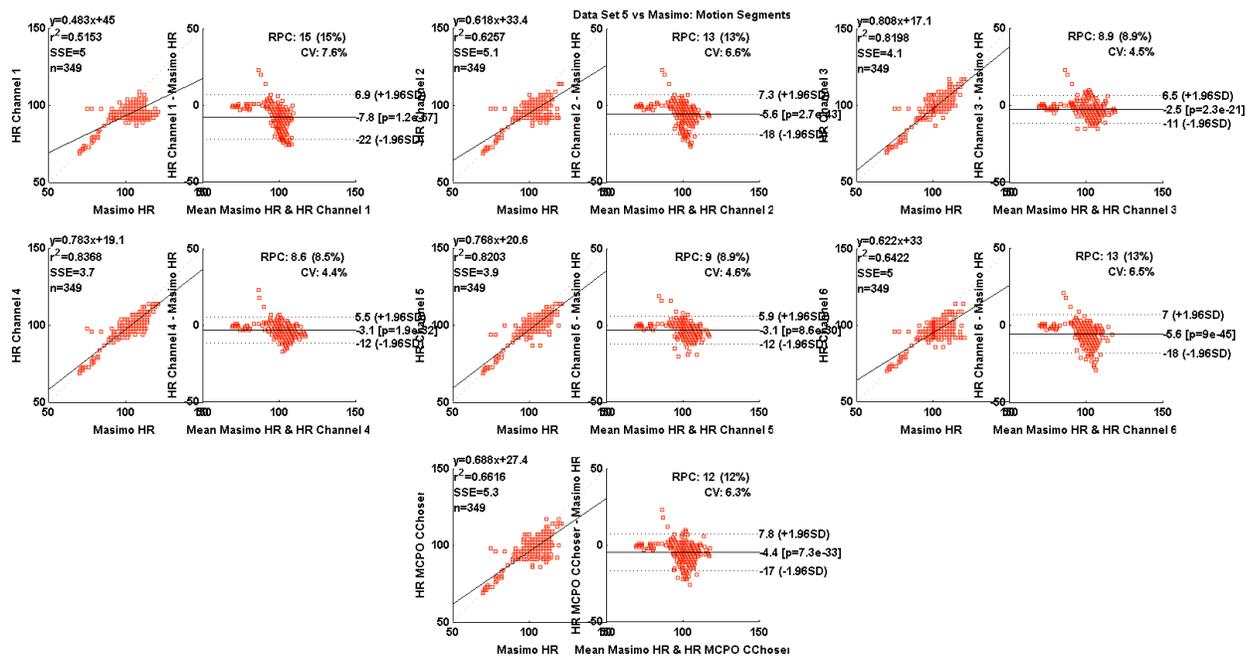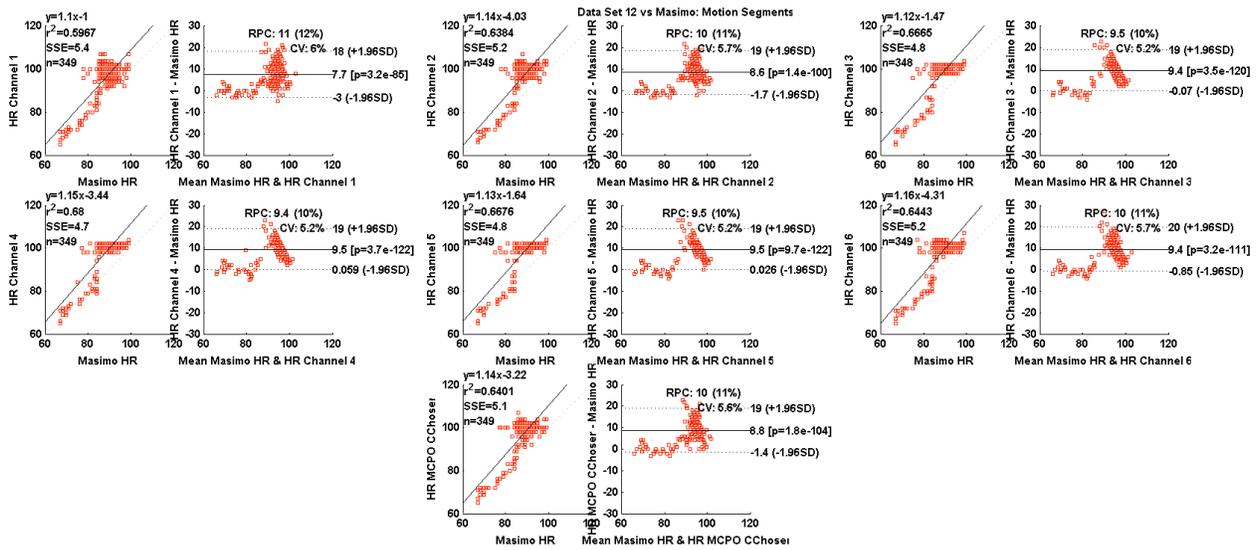

Figure H10. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 5.
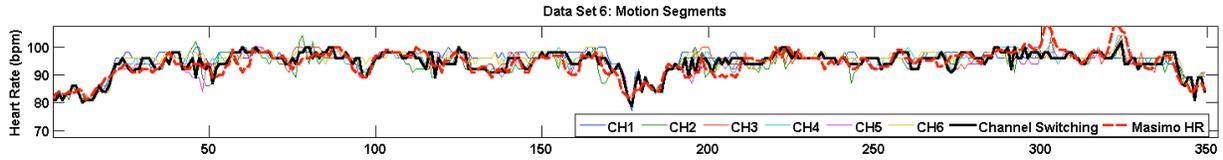
Figure H11. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 6.
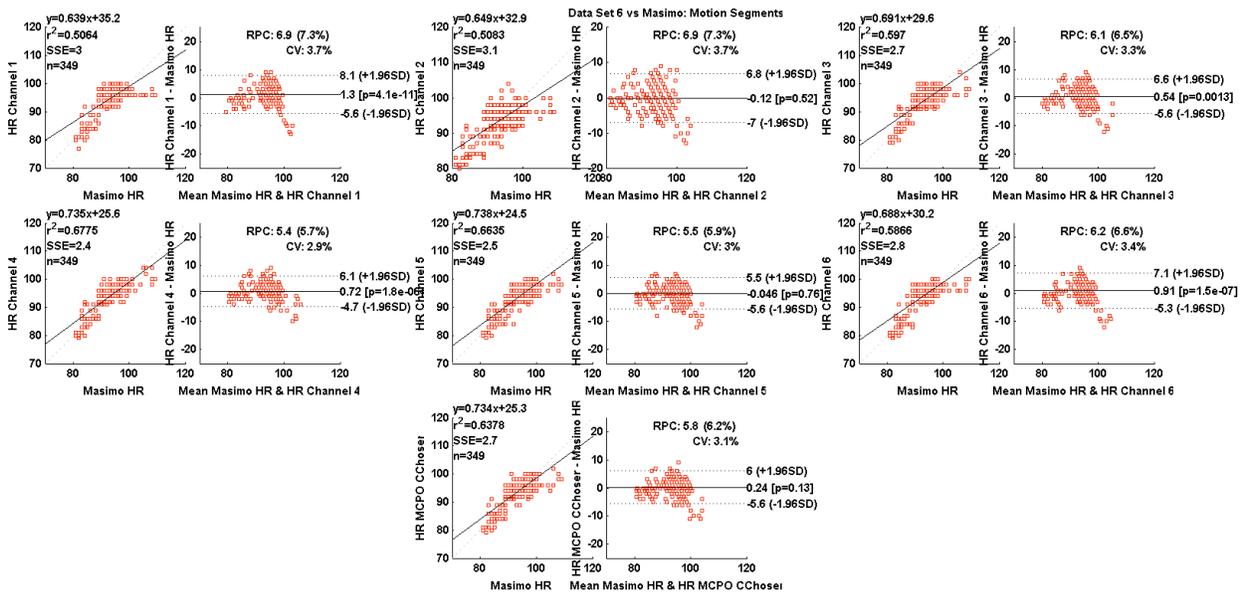
Figure H12. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 6.
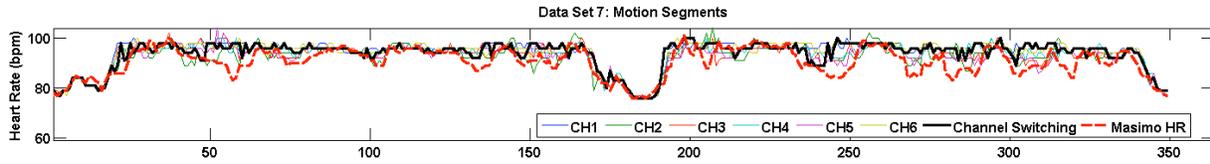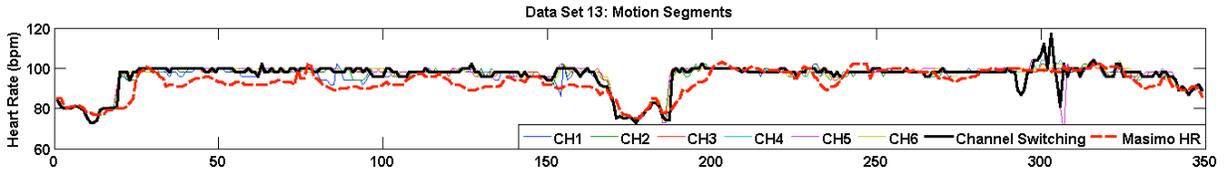
Figure H13. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 7.
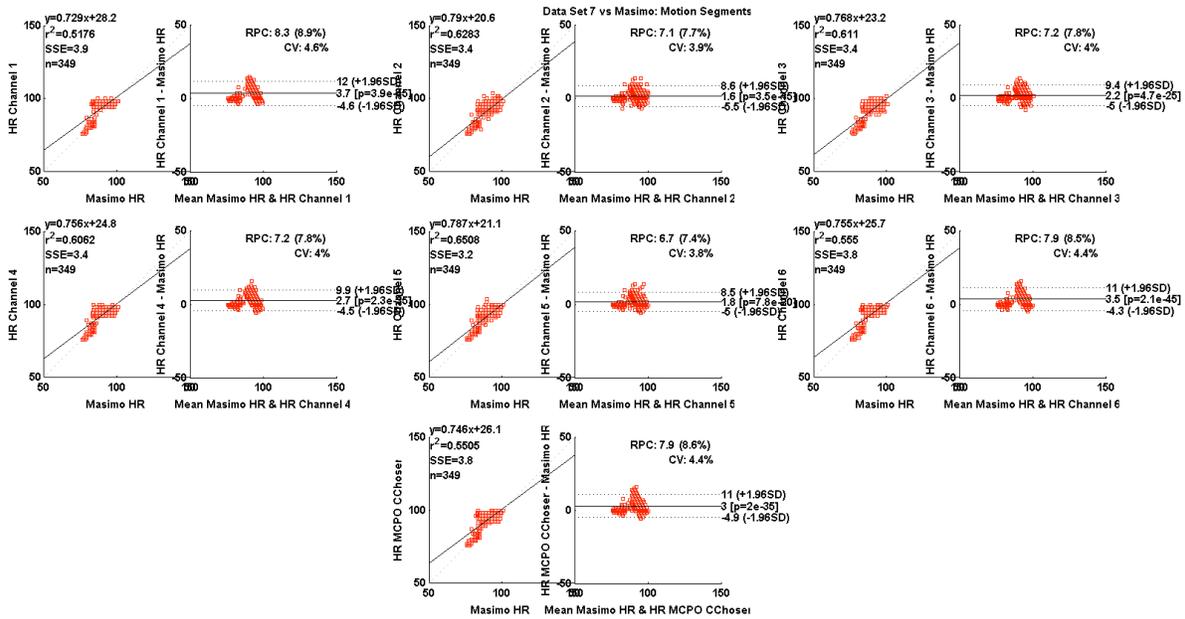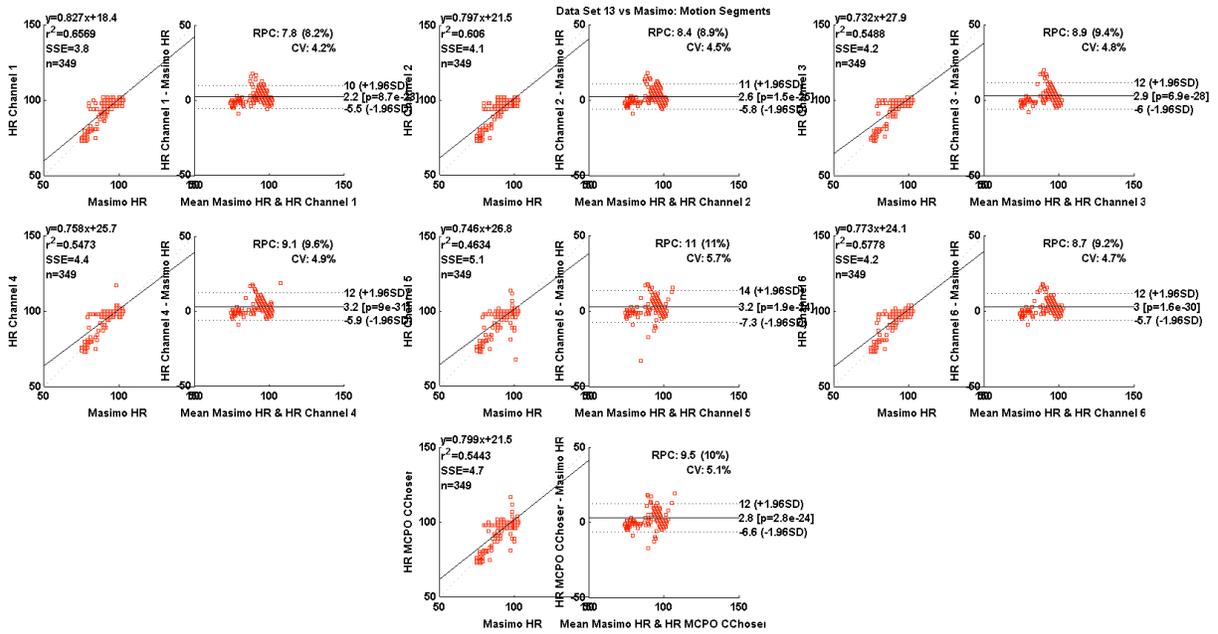

Figure H14. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 7.
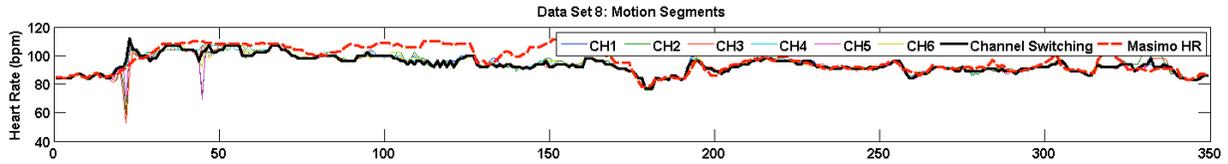
Figure H15. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 8.
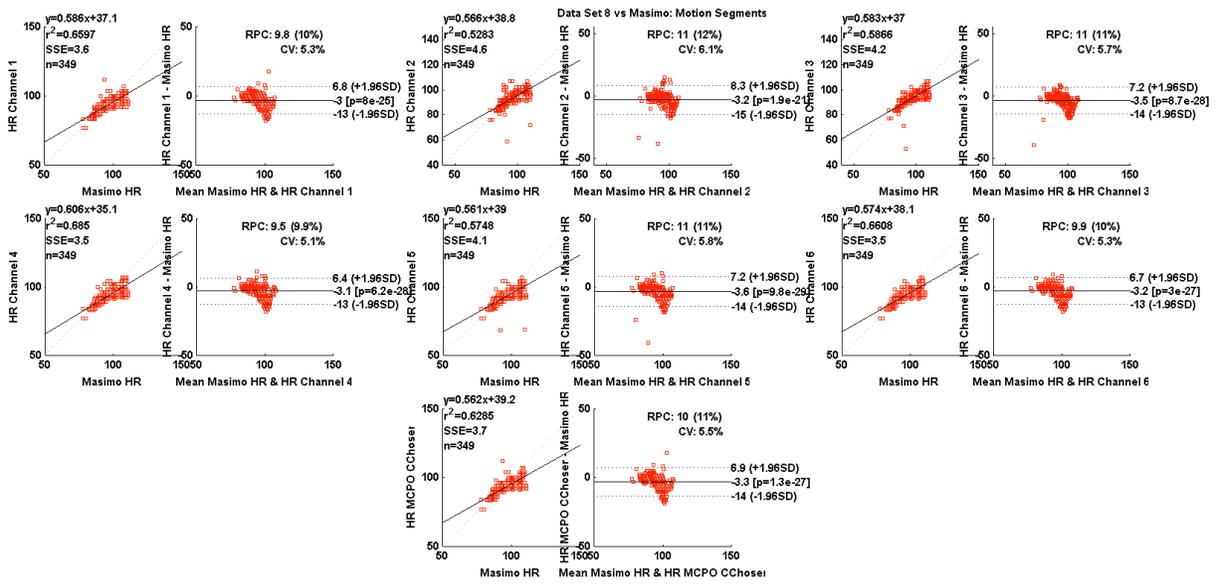
Figure H16. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 8.
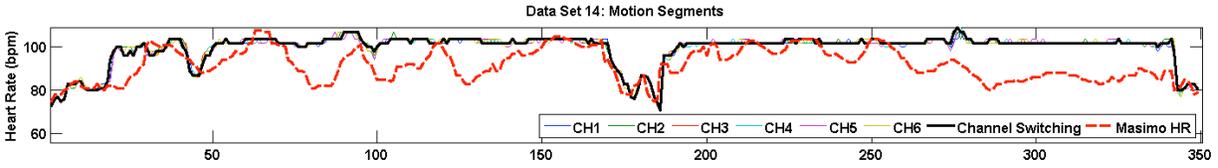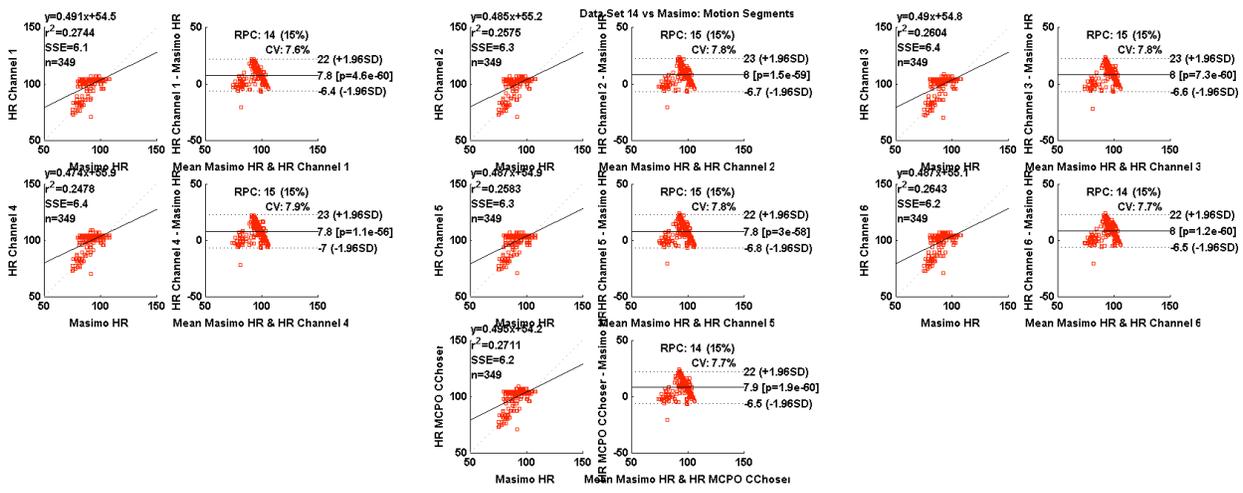
Figure H17. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 9.

Figure H18. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 9.
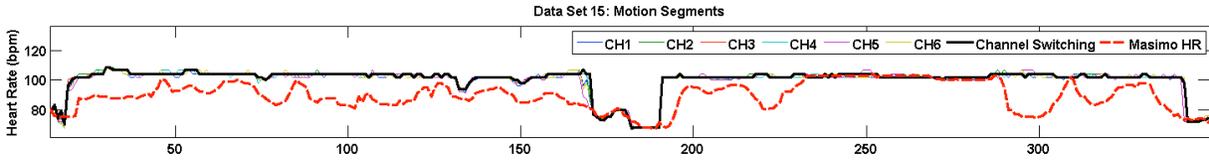
108

Figure H19. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 10.
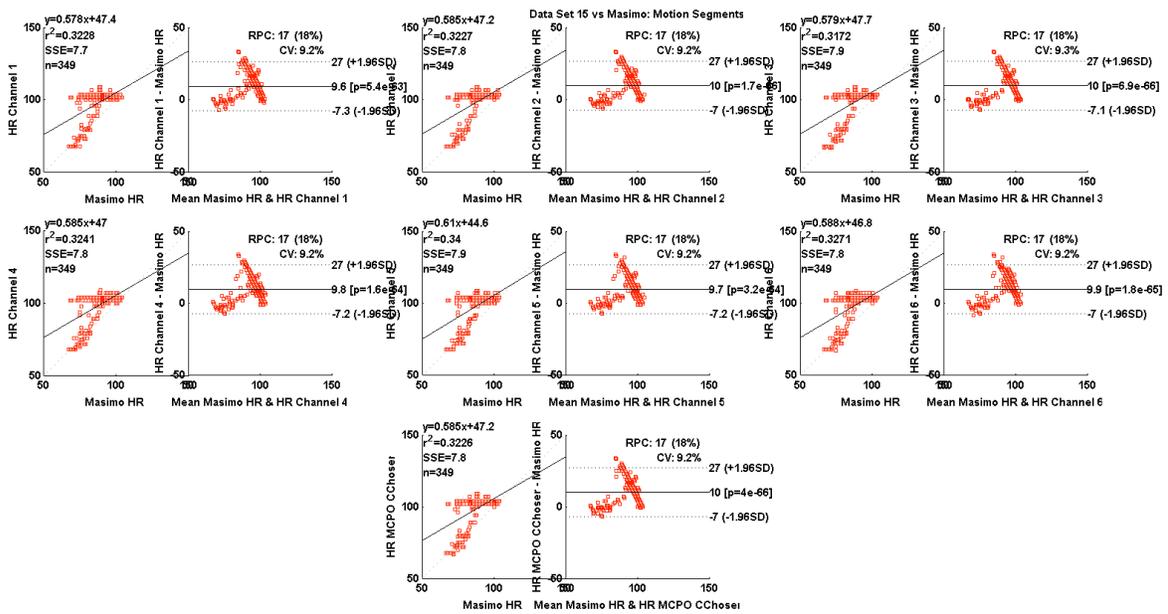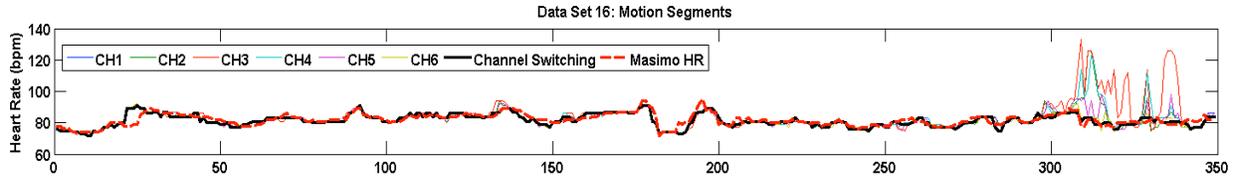

Figure H20. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 10.

Figure H21. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 11.
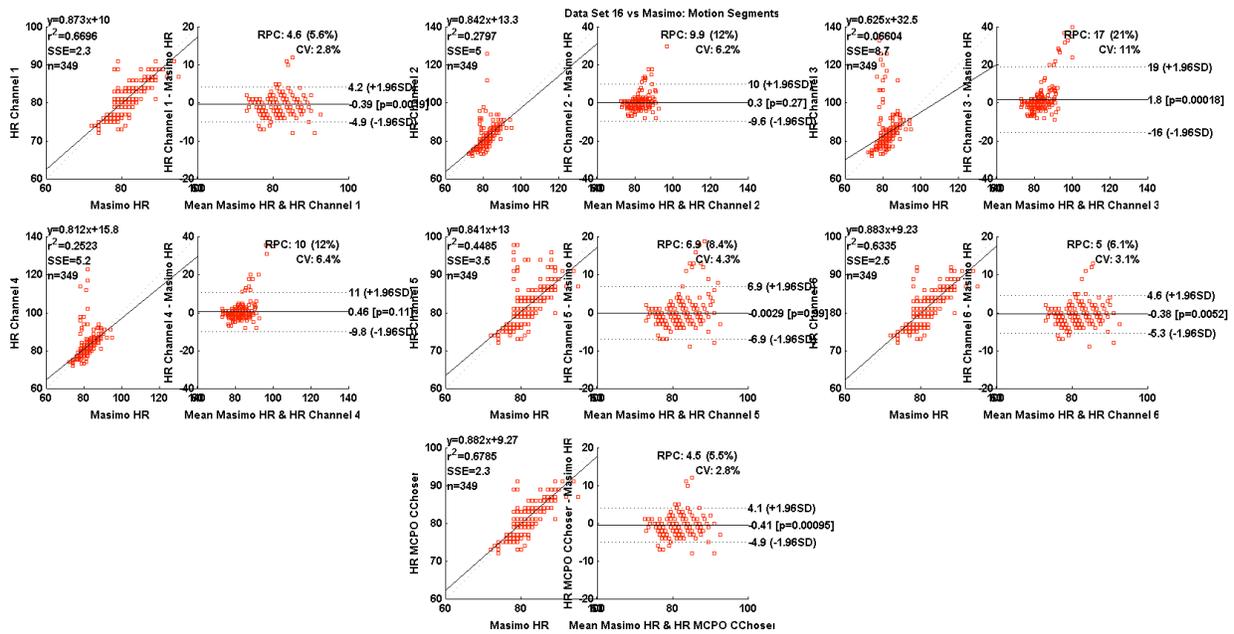

Figure H22. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 11.

Figure H23. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 12.

Figure H24. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 12.
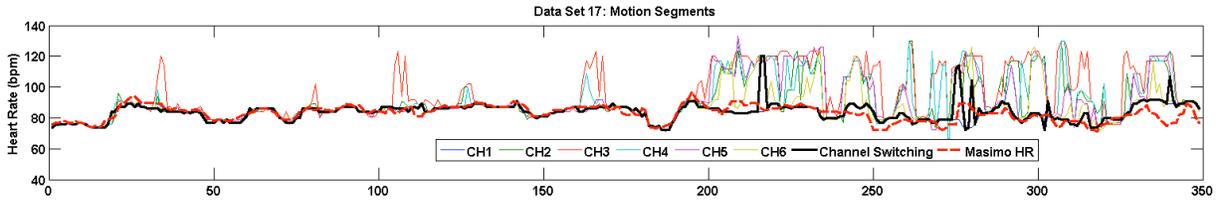
Figure H25. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 13.
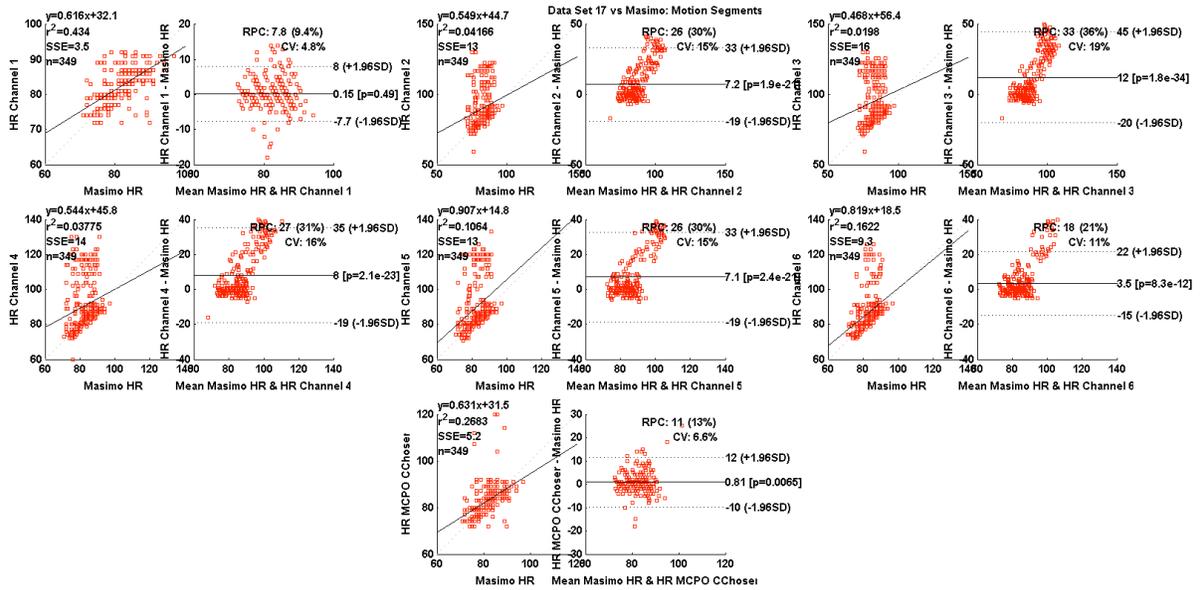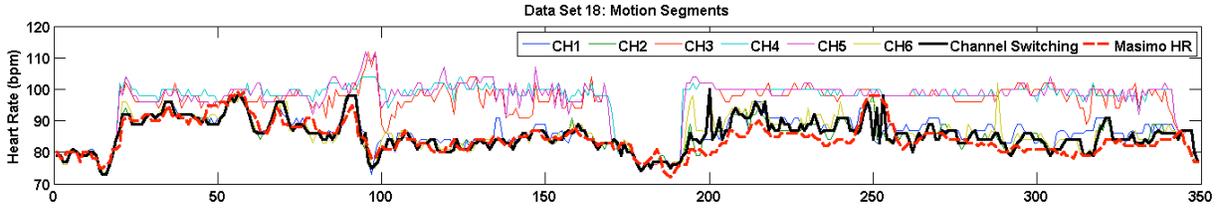

Figure H26. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 13.

112

Figure H27. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 14.
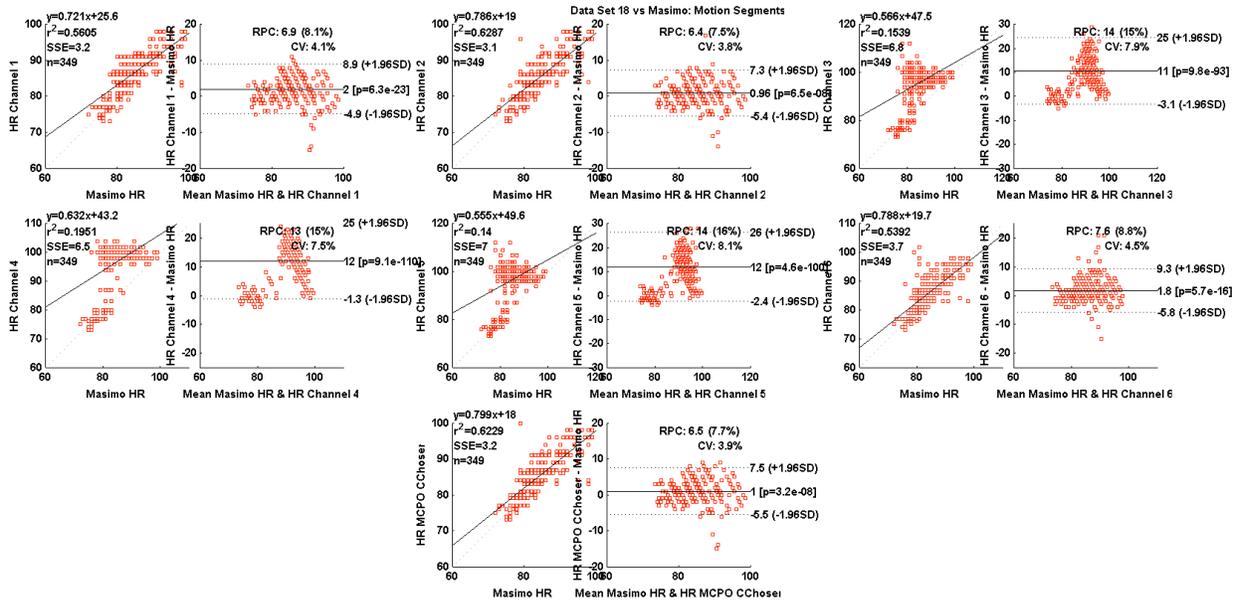

Figure H28. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 14.

113

Figure H29. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 15.

Figure H30. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 15.
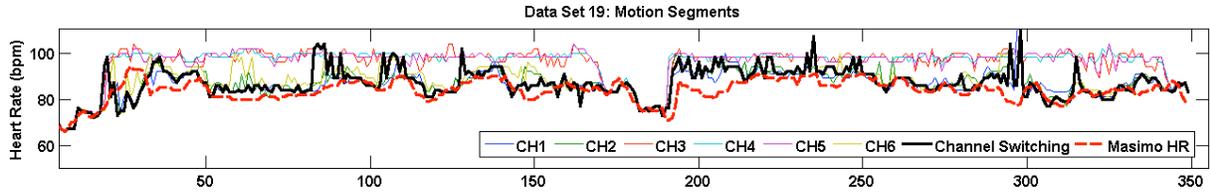
114

Figure H31. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 16.
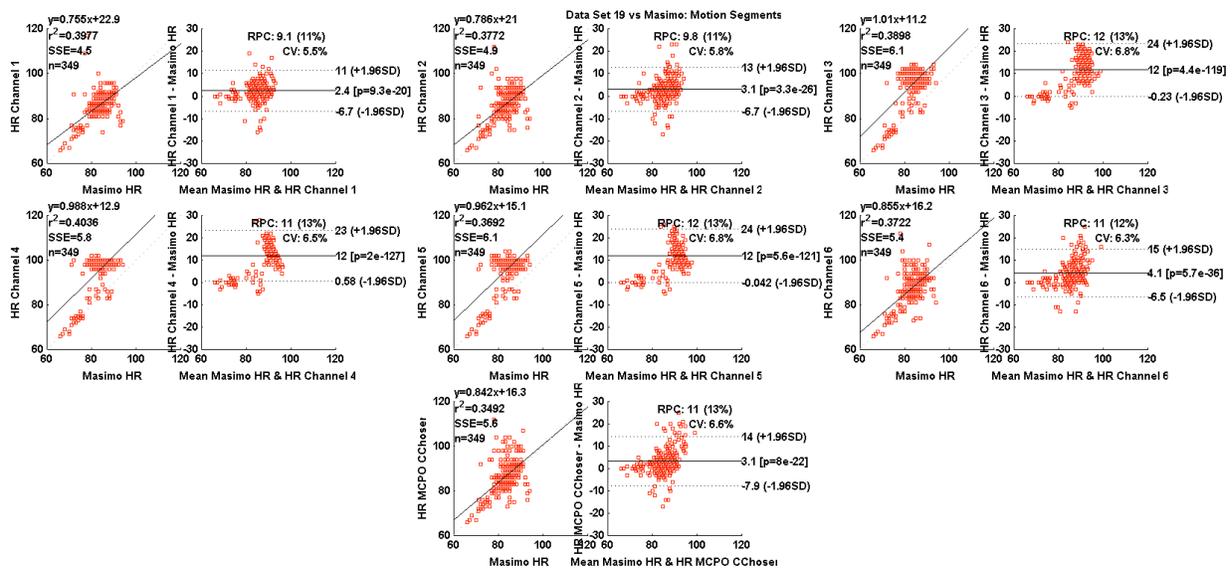

Figure H32. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 16.

115

Figure H33. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 17.

Figure H34. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 17.
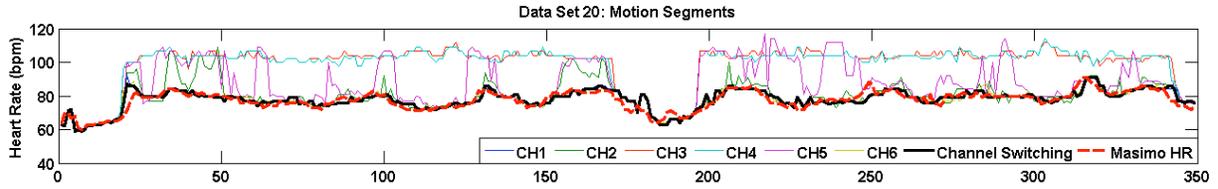
116

Figure H35. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 18.
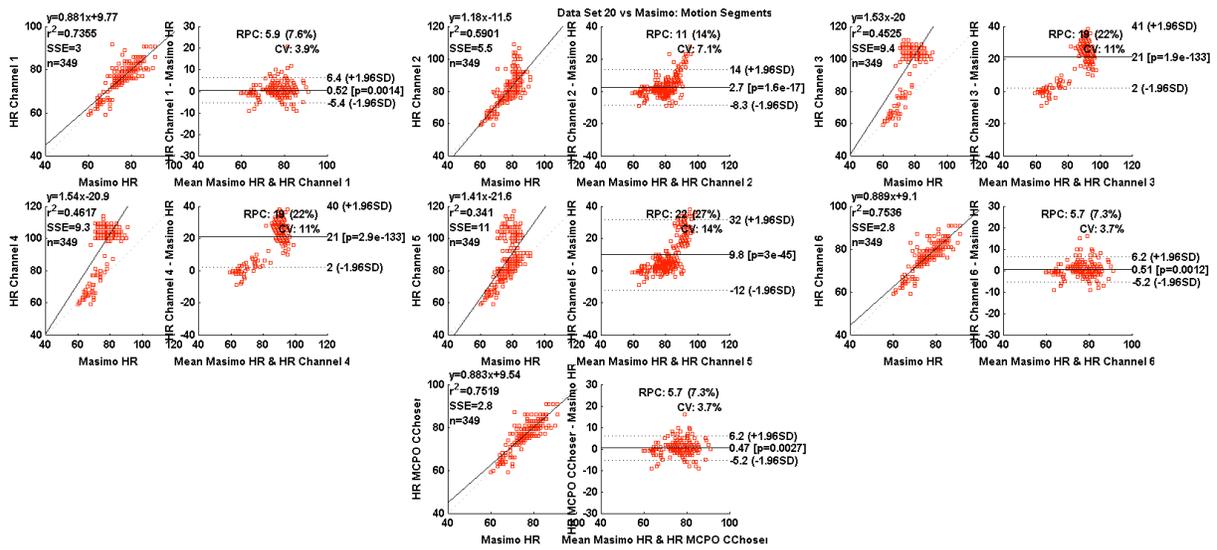

Figure H36. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 18.
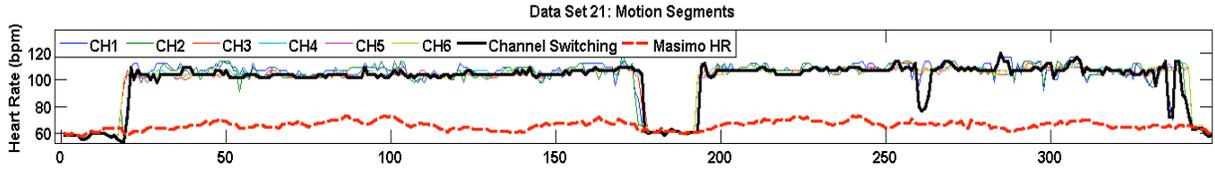
Figure H37. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 19.
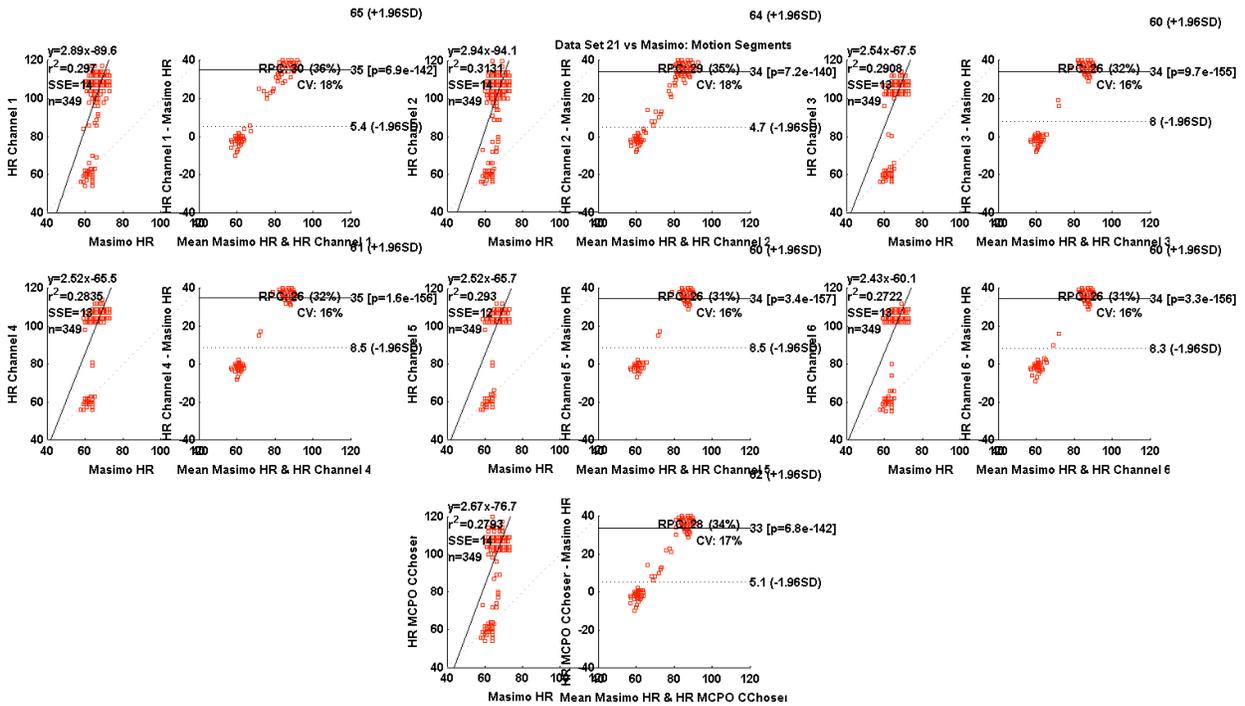

Figure H38. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 19.

Figure H39. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 20.

Figure H40. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 20.
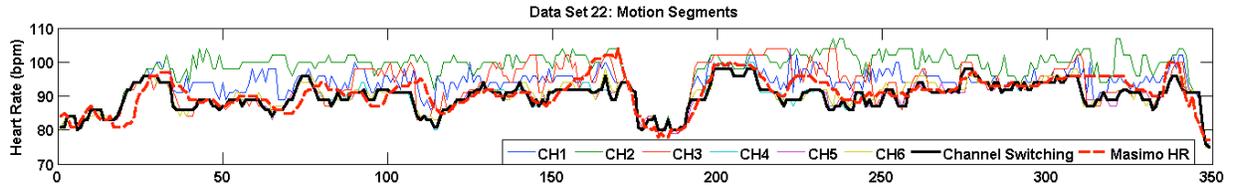
119

Figure H41. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 21.
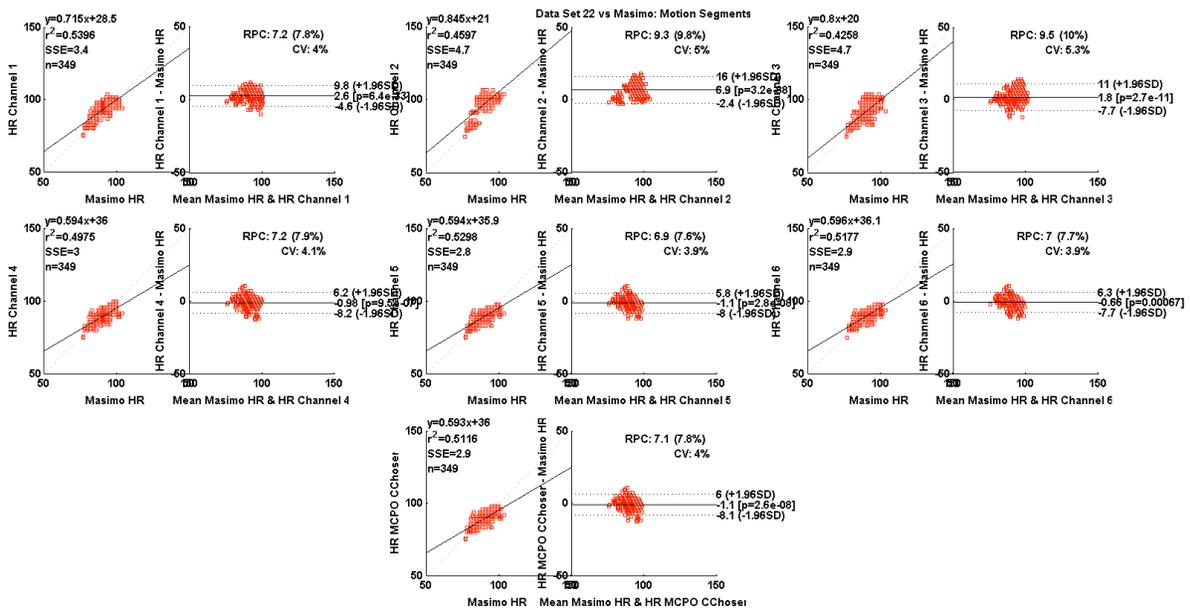

Figure H42. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 21.
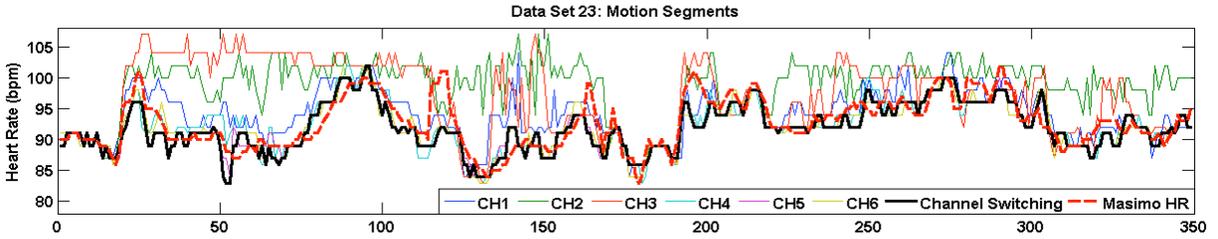
120

Figure H43. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 22.

Figure H44. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 22.

Figure H45. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 23.
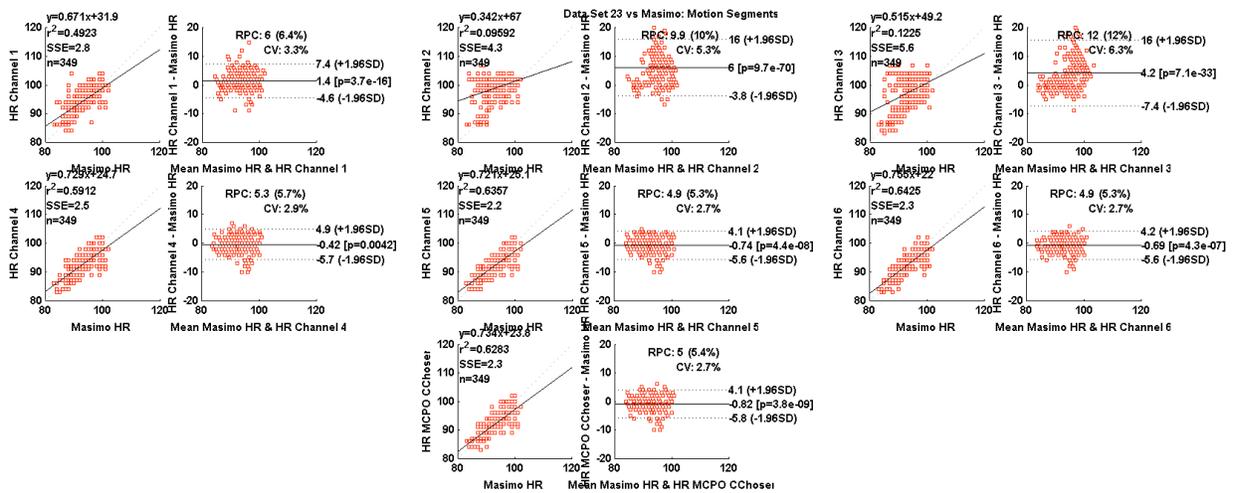

Figure H46. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 23.
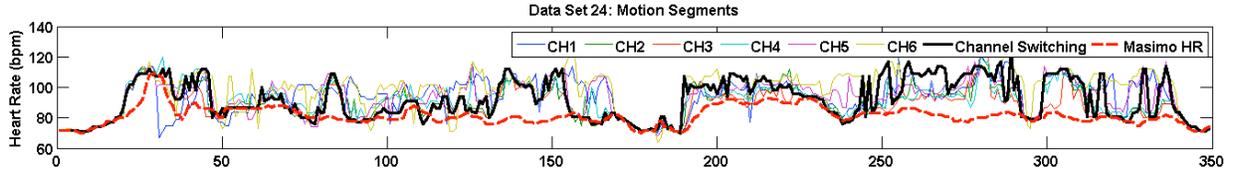
Figure H47. Time series plots of HR measurements from our multichannel device and from the Masimo reference sensor only during motion for data set 24.
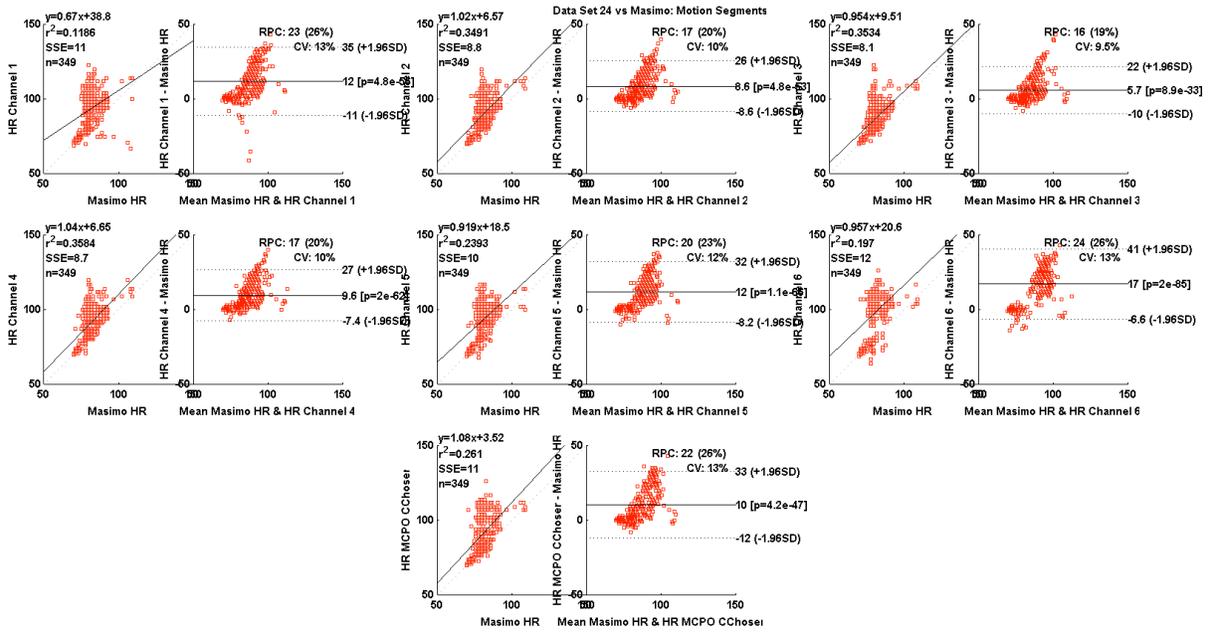

Figure H48. Correlation and Bland-Altman plots for all six individual channels and for the multichannel switching algorithm compared to the Masimo reference HR measurements during motion for data set 24.