

# Measuring the Effect of R-Peak Perturbations caused by Corruption using Heart Rate Complexity Metrics

Nicholas J. Napoli<sup>1,4</sup>, Matthew W. Demas<sup>1</sup>, Sanjana Mendu<sup>2</sup>, Chad L. Stephens<sup>3</sup>, Kellie D. Kennedy<sup>3</sup>, Angela R. Harrivel<sup>3</sup>, Randall E. Bailey<sup>3</sup>, and Laura E. Barnes<sup>1</sup>

**Abstract**—Heart rate complexity (HRC) is a proven metric for gaining insight into human stress and physiological deterioration. To calculate HRC, the detection of the exact instance of when the heart beats, the R-peak, is necessary. Electrocardiogram (ECG) signals can often be corrupted by environmental noise (e.g., from electromagnetic interference, movement artifacts), which can potentially alter the HRC measurement, producing erroneous inputs which feed into complex decision models. Current literature has only investigated how HRC is affected by noise when R-peak detection errors occur (false positives and false negatives). However, the numerical methods used to calculate HRC are also sensitive to the specific location of the fiducial point of the R-peak. This raises many questions regarding how this fiducial point is altered by noise, the resulting impact on the measured HRC, and how we can account for noisy HRC measures as inputs into our decision models. This work uses Monte Carlo simulations to systematically add white and pink noise at different permutations of signal-to-noise ratios (SNRs), time segments and HRC measurements to characterize the influence of noise on the HRC measure by altering the fiducial point of the R-peak. Using the generated information from these simulations provides improved decision processes for system design which address key concerns such as permutation entropy being a more precise, reliable, less biased, and more sensitive measurement for HRC than sample and approximate entropy.

## I. INTRODUCTION

The growing field of physiological telemetry systems is a continual source of research in numerous domains such as health care, aerospace, and nuclear power (e.g., [1], [2], [3]). These telemetry systems are designed to provide informative decision support and predictive analytics to gain physiological insight on health monitoring [1] and cognitive states identification (e.g., cognitive workload) [4], [5], [6]. One major physiological component that is leveraged in order to gain these insights is the sympathetic nervous system, which is the part of the autonomic nervous system responsible for monitoring and adapting to stress imposed on the body [7], [8], [9]. An extended increase in arousal of the sympathetic nervous system has been associated with a decrease in performance [10], increased cognitive workload [11], [12],

and deteriorating health for patients in a clinical setting [13], [14].

One of the most relevant and informative measures associated with sympathetic response is heart rate complexity (HRC) [13]. HRC algorithms have been developed as an integral latent physiological indicator being utilized in numerous studies assessing cognitive workload [15], [16], [17], a patient's physiological health and imminent risk [1], [18], [19], and analysis of depression [20]. HRC is often operationalized in the form of entropy measures which are calculated using a person's heart rate variability (HRV) [1]. These measures require accurate, precise detection of individual heart beats. While there are many methods to measure heart rate, such as electrocardiograms (ECGs), photoplethysmograms, and optical heart rate monitors, ECGs are the only method that provides information about the QRS waveform - the so-called combination of the three deflections typically seen in the ECG - which corresponds to the depolarization of the heart ventricles (a single heart beat). These QRS waveforms are critical to precisely calculating R-R intervals and thus HRV. HRV is then used to quantify the fluctuations in HRC. Thus, ECG telemetry stands as the only proven method for calculating HRC accurately. For these predictive technologies to provide insights, computational algorithms are required to examine variations from normal physiology [21]. Since many of these technologies require consistent monitoring, intermittent noise is inevitably introduced into the system. Thus noisy occurrences and detection of physiological anomalies generate ambiguities, leading to false detection, inaccurate decision support, and alarm fatigue [22], [23], [24]. However, much of the previous, original work surrounding HRV detection was developed on cleaned, retrospective datasets [1], [25], [26], [18]; thus, we cannot assume that complexity algorithms perform appropriately on noisy and corrupt data.

Understanding the robustness of HRC measurements under noisy conditions allows for corrective computational approaches and system design [24]. These corrective measurements are becoming a paramount objective as systems are expected to work in real time which leads to high risk of signal corruption. In this investigation, we aim to fill this research gap by evaluating complexity algorithms under noisy and corrupted conditions.

**Prior Work.** The majority of work on HRC was conducted by Costa et al. [27], based on Pincus's work with approximate entropy [28]. These ideas, concepts, and applications used to evaluate heart rate entropy spawned a multitude of measures [26], [29]. The two most utilized complexity mea-

\*This work was supported by the National Institute of Aerospace.

<sup>1</sup>Systems and Information Engineering, University of Virginia, Charlottesville, Va 22904 {njn5fg,mwd8jf,lb3dp}@virginia.edu

<sup>2</sup>Department of Computer Science, University Virginia, Charlottesville, Va 22904 sm7gc@virginia.edu

<sup>3</sup>NASA Langley Research Center, Hampton, Va 23681 {chad.l.stephens, kellie.d.kennedy, angela.r.harrivel}@nasa.gov

<sup>4</sup>National Institute of Aerospace, Hampton, Va 23681 {nicholas.j.napoli@nasa.gov}

asures since are approximate and sample entropy. In line with Pincus's and Costa's prior work, these complexity measures have been used on retrospective data to demonstrate predictive model feasibility and practicality in classifying patients in clinical settings [27]. While extensive studies have been conducted on various HRC methods and applications in the past few decades, few studies exist that evaluate the influence of noise and erroneous behavior on these measures. These works have demonstrated that missed detection of QRS complexes (heart beats) could drastically alter downstream entropy measurements [30], [24]. This was demonstrated by examining the effects of corruption through the false positive (FP) and false negative (FN) rates of the detection algorithms rather than looking at how the SNR itself alters entropy [30], [24]. This was done to control for the false positive and false negatives in the detection algorithms. These approaches manipulate the HRV signal through the concatenation of sequences or random down and up sampling, which changes the entropy measurement associated with the HRC [30], [24]. Although robust and accurate QRS complex detection algorithms are highly desirable and have been achieved through various types of machine learning approaches [18], QRS detection is not the only avenue through which erroneous entropy measurement can occur. ECG sampling frequency and R-peak interpolation have demonstrated importance as methodological considerations in obtaining consistent HRV signals [31]. These concerns underline the importance of the fiducial marker for the R-peak and its relationship to producing reliable HRV measurements, thus impacting the HRC. This can be depicted in Figure 1, an illustration where the ECG signal R-peak has an arbitrary confidence interval which therefore induces uncertainty in the HRV signal below. Noise plays an integral role in where the fiducial marker for the R-peak is placed, which affects downstream applications. However, neither the extent to which noise alters HRV and HRC measurement through changes in the fiducial marker nor the downstream implications of these fiducial markers on statistical models have been characterized. Ultimately, current literature's focus has been on the advancement of complexity algorithms and their utility for classification, rather than understanding how and when these algorithms fail. Developing this understanding is critical for appropriate system design in preventing alarm fatigue, erroneous prediction, and decision support systems [24], [22].

This work raises six relevant research questions (RQs): 1) How does the SNR and the color of the noise alter the fiducial marker of the R-peak? 2) Is any particular entropy measurement more precise under corruption than others? 3) Does signal length contribute to a more robust measurement in the presence of corruption? 4) Can a precise entropy measurement still be statistically sensitive in differing HRC dynamics expected in study populations? 5) If the fiducial point is altered by noise, how does this effect the direction of the measured entropy calculation? 6) How much impact does noise and the fiducial marker downstream have on the statistical implications of a study?

**Challenges.** Due to the multiple stages involved in the

process of calculating HRC, it is difficult to pinpoint exactly where noise alters an entropy measure's reliability. Noise can alter HRC reliability through numerous avenues such as False Positives for the QRS detection, false negatives of the QRS detection, or the QRS waveform fiducial point shifting. However, the robustness of current methods has only been evaluated by resampling HRV signals, thus only examining the false positive and false negative rate [30], [32], [33]. Therefore, the fundamental issue of how corruption in an ECG signal alters the HRV signal through slight shifts in fiducial points and QRS waveform timing has been neglected and is difficult to capture. Additionally, the numerous types of complexity measures produce different outputs with different scales, further complicating the validation process and comparative analysis.

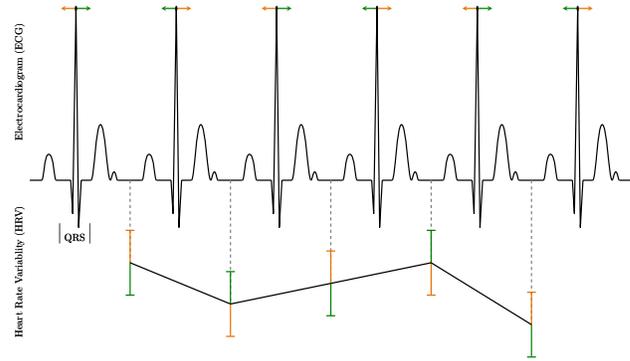


Fig. 1. This illustration demonstrates the variation in the identified location of R-wave peaks results in uncertainty in the heart rate variability (HRV). Variation in the identified location of the R-peak from the QRS complex results in uncertainty in the heart rate variability (HRV) signal.

**Insight.** A systematic process of validation in which noise can be exactly replicated through the seeding of a random number generator and introduced to a signal provides control and reproducibility when applied to testing datasets. Monte Carlo simulations allow for an increased number of controlled iterations enabling the evaluation of the aleatory uncertainty from a set of selected parameters (e.g, SNR, type of entropy measurement) to form distributions that characterize errors caused by the fiducial R-peak. This controlled simulation enables rigorous analysis on how the fiducial points impose errors in the system. Here we strictly analyze shifts in the mean and variance associated with the fiducial point at which noise enters the system. Using these approaches, we can overcome scaling differences from different entropy measures by evaluating the percent error (PE) in the various reported entropy distributions with respect to that found without the introduced noise.

**Contributions.** We designed and implemented a Monte Carlo simulation to evaluate the effects of time window size and SNR level of ECG signals on the percent error of three entropy calculation methods applied to those ECG signals. This method provided a controlled framework for evaluating the effects of ECG corruption on the reliability

of three entropy measures.

The contributions of this work are as follows:

- 1) We characterize how SNR alters the locations of the fiducial point of the R-peak via a probability mass function (PMF), and demonstrate that the fiducial point for R-peak is most affected by white noise at varying SNR levels, as compared to pink noise.
- 2) We demonstrate that permutation entropy is more precise than both approximate and sample entropy (i.e., it has the lowest standard deviation).
- 3) We demonstrate that, regardless of the entropy measurement or type of noise, as the time window segment increases, the precision of the measured entropy improves.
- 4) We demonstrate that increased precision does not imply a lack of sensitivity for demonstrating significant differences in HRC.
- 5) We demonstrate that as the SNR increases (altering the fiducial marker), there is a directional change in the measured entropy.
- 6) We show the level of SNR that affects the likelihood of making a distinguishable distribution statistically indistinguishable.

All of these findings ultimately aid us in future design methodologies to reduce predictive modeling error downstream from upstream feature extraction methodologies.

## II. METHODS

The methods section is split into two sections. The first method section addresses research questions RQ1 to RQ3 and the second section addresses RQ4 to RQ6.

### A. Monte Carlo Simulation: Addressing RQ1 to RQ3

This section outlines the Monte Carlo simulation implementation method shown in Figure 2. These simulations were designed to study the effects of time window segmentation length and signal corruption resulting from multiple “colors” of noise on the prevalence of fiducial marker shifts on approximate, permutation, and sample entropy calculations.

1) *Normal Sinus Rhythm Data:* Single lead ECG data from 10 subjects (130 min, 128 Hz) selected from the MIT-BIH Normal Sinus Rhythm Database on PhysioNet [25] were utilized. The Normal Sinus Rhythm database was selected to reduce measurement variability that could be attributed to physiological dysfunction such as atrial fibrillation, ectopy, and other disorders [27]. Each subject’s ECG signal was divided into 65, 26 and 13 intervals for 2, 5, and 10 minute time windows, respectively. Following signal preprocessing, noise was added, and the prevalence of shifts in the fiducial marker (i.e., the location of the R-peak) was collected to form a PMF.

2) *Signal Preprocessing:* After segmenting the ECG signals into intervals, a high pass and low pass finite impulse response Butterworth zero phase filter was applied to each time window. This produced our cleaned ECG signal,  $X_S$ . The zero phase filter avoids distortion in the phase of signal [34]. A key component for also avoiding fiducial marker shifting.

3) *Simulation A: Determine Effect of Noise on Prevalence of Fiducial Shifts:* The following section describes the procedure for creating a PMF of the prevalence of fiducial marker shifts caused by the addition of either white or pink noise to ECG signals.

a) *Step 1: Add Noise:* In Step 1, a white or pink noise signal,  $X_N$ , was added to the cleaned signal  $X_S$  comprising signal-to-noise ratios (SNRs) from 2 to 20 in increments of 2. The SNR level  $X_N$  added to  $X_S$  is defined as

$$SNR = 10 \log_{10} \sqrt{\frac{\sum (X_S \cdot \bar{X}_S)^2}{\sum (X_N \cdot \bar{X}_N)^2}}, \quad (1)$$

where  $\bar{X}_N$  and  $\bar{X}_S$  are the complex conjugates of  $X_N$  and  $X_S$ . A randomly seeded  $X_N$  of each SNR was added to each subject’s cleaned ECG signal interval 100, 250, or 500 times for each 2, 5, and 10 minute interval, respectively.

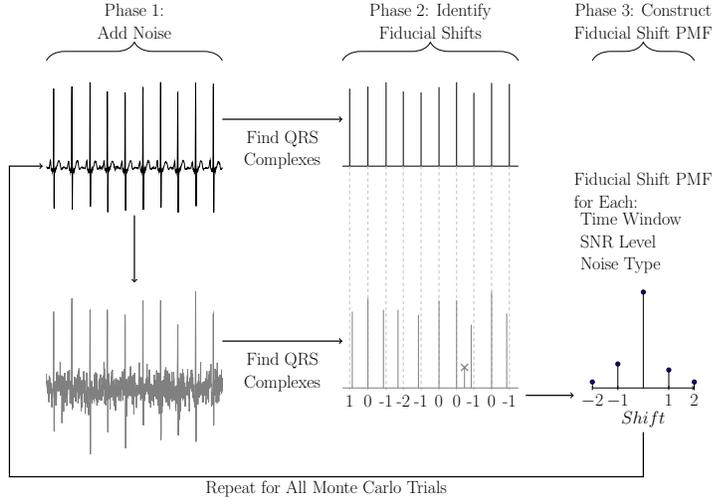
*White Noise:* White noise is designed into the implementation of the Monte Carlo Simulation to evaluate the effects of electromagnetic interference, problematic sensors, or issues with wireless devices [35], [36]. A uniform distribution was used to model white noise in the frequency domain by sampling from a random Gaussian distribution in the time domain sequence. The SNR is altered by adjusting the variance of the Gaussian distribution.

*Pink Noise:* Pink noise was selected to evaluate the effect of correlated noise typically associated with observation noise on the ECG [37]. Pink noise was modeled as a decreasing function ( $1/f$ ) in the frequency domain and has close similarity to brownian motion-like noise which is modeled as a decreasing function ( $1/f^2$ ) and is related to electrode movement noise [37]. Pink noise was implemented using a noise generator package on MATLAB’s file exchange service based on the theory for discrete simulations of colored noise by Kasdin [38].

b) *Steps 2 and 3: Identify Fiducial Shifts and Construct Fiducial Shift PMFs:* In Step 2, a QRS complex detection algorithm was applied to both the cleaned and noisy data to identify individual heart beats, R-R intervals, and heart rate variability (HRV) [39]. Using the Pan-Thompkins algorithm, the fiducial marker’s location for the R-peak is determined from the rising edge of the waveform [39]. The locations of each R-peak extracted from the noisy signal were compared to the locations of the R-peaks extracted from the filtered signal. In Step 3, the prevalence of noisy R-peaks within  $\pm 39.1$  ms of their filtered counterparts were collected to form a PMF. This particular time increment was due to our sampling frequency of 128 Hz, thus  $F_s^{-1} = 128^{-1} = 39.1$  ms. False positives (i.e., noisy R-peaks detected outside of this window) and false negatives were not included in the construction of these PMFs. This process of creating PMFs was repeated for all subjects’ time intervals (2, 5, 10 mins), resulting in PMFs for each noise type (white or pink), and for each SNR level (2,4,...,20) equaling a total of 60 PMFs (i.e., 3 Time Intervals x 10 SNR Levels x 2 Types of Noise).

4) *RQ1: Effects of Noise Level and Type on Fiducial Shift PMFs:* We used visual inspection of PMFs and the

Simulation A



Simulation B

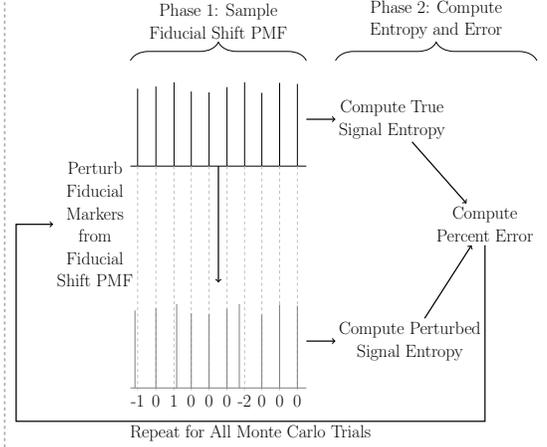


Fig. 2. Flow Chart Process of Simulation

corresponding tabulated shift frequency values to evaluate the effects of different SNRs and noise types on the prevalence of fiducial marker shifts.

5) *Simulation B: Determine Effect of Fiducial Shift Prevalence on Heart Rate Complexity:* The following section discusses the process of determining the effects on entropy measures resulting from perturbing the location of R-peaks extracted from filtered (cleaned) signals by an amount sampled from the PMFs constructed during Simulation A.

a) *Step 1: Sample Fiducial Shift PMF:* In Step 1, the location of R-peaks of subjects' filtered ECG signal intervals were perturbed with a probability determined from the fiducial shift PMFs constructed in Simulation A. A HRV signal was then computed for both the filtered and perturbed signals.

b) *Step 2: Compute Entropy and Percent Error:* In Step 2, three entropy measures were applied to each of the HRV signals—approximate entropy [28], permutation entropy [40], [29], and sample entropy [26].

*Approximate entropy:* is defined as

$$ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r), \quad (2)$$

where  $\Phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \log C_i^m(r)$ .  $C_i^m(r)$  is the number of matches to template  $i$  of length  $m$  within a tolerance of  $r$  (including self matches) [28].

*Sample Entropy:* is defined as

$$SampEn(m, r, N) = -\ln \left[ \frac{A^m(r)}{B^m(r)} \right] \quad (3)$$

where  $A^m(r)$  and  $B^m(r)$  are the probabilities of two sequences matching for  $m + 1$  and  $m$  data points respectively within a tolerance of  $r$  (excluding self matches) [26].

*Permutation Entropy:* employs rank order (i.e., not exact distance) to quantify time series similarity. This is defined as

$$H_n = -\sum_{j=1}^{n!} p'_j \log_2 p'_j \quad (4)$$

where  $p'_j$  is the proportion of the occurrence of the  $j^{\text{th}}$  template (of length  $m$ ) in the signal [40], [29]. Approximate [28] and sample entropy [26] both use distance as the measurement to examine similarity in time series to quantify entropy. The main difference between the two is that unlike approximate entropy, sample entropy excludes self matches. On the other hand, permutation entropy employs rank order (i.e., not exact distance) to quantify time series similarity [40], [29]. In Step 2, approximate, permutation, and sample entropy values were computed for both the filtered and perturbed HRV signals and compared to calculate percent error 100, 250, or 500 times for each 2, 5, or 10 minute interval, respectively; each noise type (white or pink); and each SNR level (2, 4, ..., 20).

6) *RQ2: Comparison of Entropy Method Reliability:* We computed the standard deviation of each time interval and then applied the natural logarithm to the data to obtain relatively equal variances and reasonably symmetric distributions. We then tested for the differences in the mean log standard deviation of the percent error using a one-way ANOVA.

7) *RQ3: Effects of Window Size on Reliability:* Using the data from Simulation B, conditional distributions were constructed of the percent error in entropy for time window sizes of 2, 5, and 10 minutes and for both pink and white noise by collapsing over SNR, subject, and time interval number. Using visual comparison and descriptive statistics, the effects of window size on entropy reliability were addressed for both white and pink noise.

### B. Sympathetic Response Dataset: Addressing RQ4 to RQ6

The experience of hypoxia is known to result in autonomic-nervous-system-driven changes (sympathetic arousal) in the cardiac and respiratory systems [41]. We utilize a dataset which characterizes hypoxic responses because of these known physiological changes in

sympathetic arousal to address research questions RQ4 to RQ6.

1) *NASA Hypoxia Data*: The dataset was collected by a research team at NASA Langley Research Center (LaRC) who subjected 49 volunteers (all with current hypoxia training certificates) to normobaric hypoxia to study the impact on aircraft pilot performance [42], [43].

The goal of the study was to understand cognitive impairment due to exposure to mild hypoxia in order to improve the safety of psychophysiological-based automation interfaces. Subjects in the study experienced simulated altitudes of Sea Level (21% O<sub>2</sub>) and 15,000 feet (11.2% O<sub>2</sub>) induced by an Environics, Inc. Reduced Oxygen Breathing Device (ROBD-2). During non-hypoxic and hypoxic exposures each subject performed a battery of written, computer-based, and flight simulation tasks each lasting 10 minutes. In each exposure, the research team collected task performance measures, a subjective self-report of workload (NASA Task Load Index), and multiple physiological responses (including ECG). This article discusses only the ECG data collected during hypoxic and non-hypoxic exposures.

2) *RQ4: Evaluating Entropy Method Sensitivity*: To evaluate the sensitivity of entropy measures in detecting mild hypoxia, approximate, permutation, and sample entropy were calculated for ECG data collected during the final 2 min and final 5 min of non-hypoxic and hypoxic exposures. Ten minute time segments were excluded from this analysis because hypoxic exposures were only ten minutes long and thus not all subjects demonstrated indicators of mild hypoxia (e.g., SpO<sub>2</sub> < 80%) until several minutes into the 15,000 ft exposures. Due to small sample size and non-normality, the non-parametric Wilcoxon rank-sum test was employed to test for differences between the hypoxic and non-hypoxic cohorts with smaller p-values indicating a greater ability for an entropy calculation method to discriminate between hypoxic and non-hypoxic states. The initialized parameters used in calculating the entropy measurements were the same as in Simulation B ( $m = 2$  &  $r = 0.15\sigma$  for sample and approximate entropy and  $m = 3$  for permutation entropy order).

3) *Simulation C: Effect of Fiducial Shifts on Entropy Data from Hypoxic Subjects*: Using the NASA hypoxia ECG dataset and the PMFs of the fiducial marker shifting found in Simulation A, we corrupt the HRV signals by altering the locations of the R-peaks for 2 and 5 minute time segments at various SNR levels. Each subject and their respective cohort's (hypoxia and non-hypoxia) fiducial markers are corrupted 250 times to examine the raw change in the entropy distribution.

4) *RQ5: Effects of Fiducial Shifts on Entropy Skew and Bias in Hypoxic Subjects*: We pooled the approximate, permutation, and sample entropy values for all subjects' Monte Carlo trials for both hypoxic and non-hypoxic exposures at each SNR level and time window size (2 and 5 min). Kernel density plots of each pooled distribution were obtained and visually inspected for skewness. For each conditional pooled distribution, the mean entropy value was computed and

TABLE I  
PINK NOISE MONTE CARLO SIMULATION SUMMARY OF PROBABILITY DISTRIBUTION OF THE FIDUCIAL SHIFT (TRIALS= $\sim$  65,000).

Target SNR	Distribution of Fiducial Shift by Milliseconds				
	-15.62ms	-7.81ms	0ms	7.81ms	15.62ms
2	0.0	4.63	90.39	4.95	0.0
4	0.0	3.68	92.41	3.89	0.0
6	0.0	2.94	93.97	3.07	0.0
8	0.0	2.33	95.22	2.42	0.0
10	0.0	1.85	96.21	1.92	0.0
12	0.0	1.47	96.99	1.53	0.0
14	0.0	1.18	97.60	1.21	0.0
16	0.0	0.94	98.10	0.96	0.0
18	0.0	0.75	98.48	0.76	0.0
20	0.0	0.60	98.79	0.60	0.0

qualitatively assessed for general trends in the data.

5) *RQ6: Effects of Fiducial Shifting on Cohort Discriminability*: We demonstrate this through an approach similar to that of RQ3. Using the NASA hypoxia dataset and the probability distributions associated with the fiducial shifts, we alter the fiducial markers of the R-peaks in the ECG signals to simulate corruption for each subject. This process is simulated 250 times for each time window segment length, cohort, and entropy measurement at each SNR level.

During each simulation trial the non-parametric Wilcoxon rank-sum test evaluated differences between the hypoxic and non-hypoxic distributions with the null-hypothesis significance level set at 0.05. After all simulation trials were complete, the percentage of tests rejecting the null-hypothesis (i.e., having  $p < 0.05$ ) were collected. A greater percentage of null-hypothesis rejections indicates a greater ability to resolve hypoxic from non-hypoxic states.

### III. EVALUATION AND DISCUSSION

In this section, we address the following research questions regarding the effects of ECG corruption on entropy dynamics which were introduced in Section I:

- RQ1 How does SNR and various colored noise effect the fiducial point of the R-peak?
- RQ2 If the fiducial point of the R-peak shifts, does a single entropy method demonstrate superior precision from the proposed three types of entropy calculations that measure HRC?
- RQ3 Does increasing the HRV signal time segment length enhance HRC variance (entropy dynamics)?
- RQ4 Is increased precision associated to lack of sensitivity for demonstrating significant differences in HRC?
- RQ5 Is there a directional change in the measured entropy from corruption in which the entropy increases, decreases, or does only the variance symmetrically increase providing no actual change in complexity?
- RQ6 At what simulated SNR level do fiducial shifts render two previously distinguishable distributions indistinguishable?

TABLE II

WHITE NOISE MONTE CARLO SIMULATION SUMMARY OF PROBABILITY DISTRIBUTION OF THE FIDUCIAL SHIFT (TRIALS= $\sim 65,000$ ).

Target SNR	Distribution of Fiducial Shift by Milliseconds				
	-15.62ms	-7.81ms	0ms	7.81ms	15.62ms
2	0.01	9.74	78.9	11.21	0.10
4	0.0	8.02	82.90	9.02	0.04
6	0.0	6.52	86.29	7.16	0.01
8	0.0	5.23	89.13	5.62	0.0
10	0.0	4.20	91.35	4.44	0.0
12	0.0	3.33	93.15	3.49	0.0
14	0.0	2.65	94.57	2.79	0.0
16	0.0	2.10	95.71	2.17	0.0
18	0.0	1.66	96.60	1.73	0.0
20	0.0	1.37	97.21	1.41	0.0

TABLE III

PINK NOISE ENTROPY PERCENT CHANGE DISTRIBUTION: (TRIALS= $\sim 65,000$ ).

Time Window	Target SNR	Approximate Entropy		Permutation Entropy		Sample Entropy	
		$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
2	2	0.34	5.90	1.83	1.38	7.68	9.23
2	6	0.27	5.14	1.50	1.23	5.89	9.10
2	10	0.34	3.95	1.05	1.01	4.02	6.25
2	14	0.20	3.12	0.68	0.84	2.48	4.96
2	18	0.04	2.45	0.47	0.68	1.65	4.31
5	2	1.52	4.39	1.65	0.93	5.50	7.85
5	6	1.18	3.89	1.35	0.81	4.25	7.16
5	10	0.82	2.82	0.93	0.64	2.97	4.98
5	14	0.53	1.98	0.62	0.50	1.95	3.32
5	18	0.36	1.49	0.42	0.39	1.28	2.50
10	2	4.78	3.22	1.61	0.59	6.99	2.50
10	6	3.77	2.63	1.32	0.52	5.45	2.09
10	10	2.50	1.88	0.90	0.40	3.55	1.66
10	14	1.64	1.33	0.60	0.32	2.31	1.29
10	18	1.09	1.04	0.40	0.26	1.50	1.05

#### A. RQ1 — Fiducial R-peak Marker, Type of Noise, and SNR

We hypothesized that the fiducial marker will change as a function of the SNR and type of noise. Each row in Tables I and II present a PMF of the fiducial marker shifting for white and pink noise, respectively, as a function of the targeted SNR. The PMF was generated by part A of the simulation discussed in Figure 2 and describes the probability of the fiducial marker shifting. Each shifting in the fiducial marker is a single discretized point or index in the time series equivalent to  $F_s^{-1}$ . Since the sampling rate of the discretized signal is 128Hz, the index shifts in increments of 7.81ms. Both of these tables demonstrate that as the SNR increases, there is a greater probability of altering the fiducial marker. As a function of the type of noise applied to the signal, white noise is the most corruptive in altering the fiducial marker.

#### B. RQ2 — Type of Entropy and Precision

Due to the fiducial point of the R-peak shifting, we address how precision (i.e., standard deviation,  $\sigma$ ) can alter the proposed three types of entropy calculations that measure HRC. Tables III and IV are generated from Part B of the Simulation discussed in Figure 2 that develops distributions based on the percent change for their respective entropy measurement, SNR, Time Segment, and type of noise. To

TABLE IV

WHITE NOISE ENTROPY PERCENT CHANGE DISTRIBUTION: (TRIALS= $\sim 65,000$ ).

Time Window	Target SNR	Approximate Entropy		Permutation Entropy		Sample Entropy	
		$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
2	2	-0.34	11.21	3.45	2.07	18.84	14.63
2	6	0.21	8.35	2.75	1.74	12.92	11.75
2	10	0.40	6.29	1.98	1.46	8.65	9.84
2	14	0.32	4.78	1.39	1.19	5.71	7.84
2	18	0.26	3.77	0.94	0.98	3.63	6.64
5	2	3.31	8.26	3.21	1.65	14.24	13.38
5	6	2.46	6.41	2.57	1.32	9.50	11.54
5	10	1.72	4.96	1.88	1.04	6.42	9.03
5	14	1.15	3.73	1.33	0.80	4.24	6.78
5	18	0.85	2.52	0.89	0.62	2.97	4.20
10	2	12.7	8.07	3.23	1.09	19.20	5.24
10	6	8.54	5.58	2.59	0.86	13.10	3.87
10	10	5.75	3.80	1.89	0.66	8.50	2.82
10	14	3.80	2.69	1.32	0.51	5.49	2.18
10	18	2.48	1.90	0.90	0.40	3.54	1.66

address RQ2 we can first visually compare the standard deviations of Tables III and IV where for each row in the table, all the variables (e.g., SNR, Time Segments) are held constant except for the type of entropy measurement (e.g., sample entropy). We demonstrate that the standard deviation of the permutation entropy is respectively lower than either approximate or sample entropy across all SNR levels, time segments, and type of noise. Thus from the Monte Carlo simulation, we demonstrated that permutation entropy provides a more precise HRC measurement than either approximate or sample entropy. This finding is likely due to the manner in which permutation entropy is calculated versus sample and approximate entropy. Permutation entropy analyzes HRV by rank order of the R-R interval timing rather than their distance criteria like sample and approximate entropy. Outside of false positives and missed detections of the QRS wave, corruption of the fiducial shift and the timing of QRS wave indices has a stronger effect on distance than on the rank order.

For both white and pink noise, the one-way ANOVA indicated that there were significant differences in the mean log standard deviation of the percent error of the three entropy measures ( $F(2, 31575) = 12440$ ,  $p < 0.05$  and  $F(2, 30957) = 11156$ ,  $p < 0.05$  for white and pink noise respectively). Post-hoc comparisons using the Tukey-Kramer procedure revealed that all entropy methods produced significantly different mean log standard deviations of the percent error ( $p < 0.05$ ) for both white and pink noise with permutation entropy having the lowest mean log standard deviation of the percent error followed by approximate entropy and then sample entropy.

#### C. RQ3 — Noisy Time Segments and HRC Variance

Increased time segments provide more data which should aid in reducing the aleatory uncertainty for entropy measurements. Therefore we hypothesize that increased time segment lengths at a fixed SNR value would result in more stable and precise HRC measurements. Through the proposed Monte

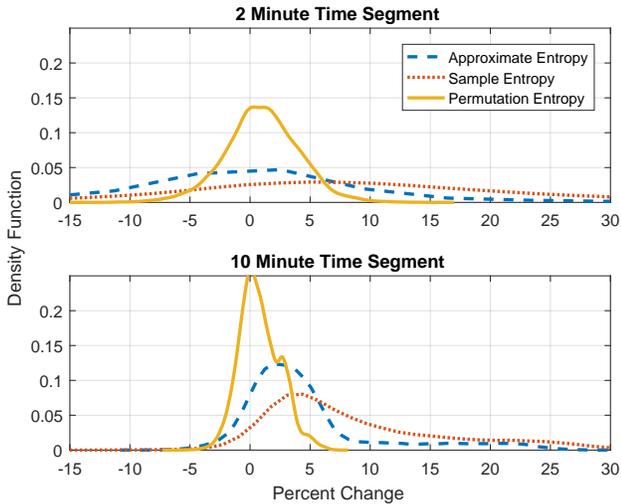


Fig. 3. Percent Change Entropy Distribution (SNR=2db, White Noise)

Carlo Simulation (Part B) in Figure 2, the SNR values were kept fixed but the time windows varied by 2, 5 and 10 minutes for pink and white noise. By examining Tables III and IV at their respective SNR levels, we note that the standard deviation and expectation decreases as the window size increases for both white and pink noise.

Based on the Monte Carlo simulated samples, a pictorial representation of the distribution was created using kernel density estimation for white noise SNR levels of 2 and 20 at 2 and 10 minute segments shown in Figures 3 and 4. Figure 3 visually demonstrated Sample Entropy extending beyond a 30% change from its original entropy calculation. However, as the time segment increases we can note the mass of the distribution becomes more centralized, thus increasing the precision of the entropy measurement. This effect of increasing the windowing segment length is consistent for all three entropy calculation methods. In regards to real-time health monitoring, we are met with a trade-off between time resolution and the precision of the entropy measurement. For example, as we aim to characterize a subject's physiology within a smaller time frame, the precision of the measurement worsens. Thus, when designing these systems, one should consider the level of imprecision which can exist before unacceptable Type 1 or Type 2 errors are encountered in the system.

#### D. RQ4 — Insight HRC Sensitivity

Although some HRC measurements may not offer ideal precision or stability (e.g., Sample and Approximate Entropy), we hypothesize that this lack of precision allows for higher sensitivity in distinguishing features between two cohorts. That is, the increased variation in the measurement that results in a decrease in precision provides information that can aid in demonstrating independence between two classes or distributions.

Table V provides the p-value for the Wilcoxon Rank Sum Test for the three proposed entropy measurements at 2 and 5 minutes, which evaluates the hypoxic and non-hypoxic

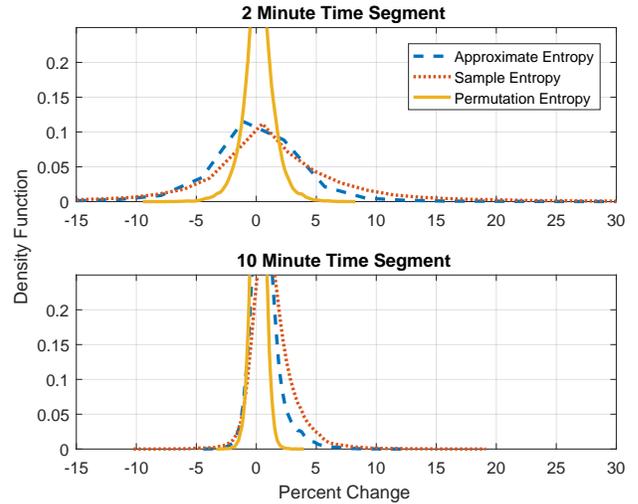


Fig. 4. Percent Change Entropy Distribution (SNR=20db, White Noise)

cohorts of the presented NASA experiment. In Table V, we note that permutation entropy had a lower P-Value for both sets and for all time segments than approximate and sample entropy, demonstrating independence between two classes or distributions. Based on the results and discussion of RQ3, permutation entropy was also shown to be the most precise measurement from the other proposed entropy measurements. Thus, the results of this data do not fully support our hypothesis that these imprecise measurements are more sensitive to noise but are better at distinguishing independence. Hence, there is no trade-off between increased precision and the ability to distinguish the two distributions of hypoxia and non-hypoxia, a known cause of autonomic nerve system activation which alters HRC dynamics.

TABLE V  
SIGNIFICANCE LEVEL (P-VALUE) OF WILCOXON RANK SUM TEST  
BETWEEN HYPOXIC AND NON-HYPOXIC COHORTS

Time	<i>PermEnt<sub>0</sub></i>	<i>ApproxEnt</i>	<i>SampEnt</i>
2 mins	<b>p = 0.022</b>	p = 0.55	p = 0.099
5 mins	<b>p = 0.032</b>	p = 0.067	<b>p = 0.034</b>

#### E. RQ5 — What is the Directional Change in Entropy

As noise alters the fiducial point, we aim to understand how this affects the direction of the entropy measurements. For example as noise increases and alters the fiducial point, does the measured entropy increase, decrease, or is there no clear trend apparent (symmetric)? We hypothesized that increased perturbations of the fiducial point from increased noise would make the HRV signal less predictive, thus increasing the entropy measured in the signal. There is some support of this hypothesis from the results provided in RQ2 and RQ3, where you can note a general positive shift in the reported expectation ( $\mu$ ) of the calculated percent change. This is also demonstrated by Figures 3 and 4, where the percent change of the distributions are skewed further to the right as SNR increases.

TABLE VI  
WHITE NOISE ENTROPY MOVEMENT (TRIALS= $\sim$  250).

Time Window	Target SNR	Approximate Entropy		Permutation Entropy		Sample Entropy	
		$H_\mu$	$NH_\mu$	$H_\mu$	$NH_\mu$	$H_\mu$	$NH_\mu$
2	2	0.679	0.691	0.888	0.909	1.345	1.483
2	6	0.672	0.688	0.883	0.906	1.332	1.470
2	10	0.669	0.687	0.880	0.903	1.319	1.465
2	14	0.667	0.686	0.878	0.902	1.310	1.462
2	18	0.665	0.685	0.876	0.901	1.307	1.459
5	2	0.911	0.975	0.897	0.914	1.227	1.407
5	6	0.903	0.970	0.892	0.911	1.206	1.393
5	10	0.895	0.967	0.888	0.909	1.191	1.384
5	14	0.892	0.964	0.886	0.908	1.182	1.384
5	18	0.888	0.962	0.884	0.907	1.177	1.377

However, in order to properly support our hypothesis RQ5, provides the raw calculated values which were averaged over subjects and trials. More specifically, Table VI provides the mean value of these distributions specifically associated to either the Hypoxic ( $H_\mu$ ) or Non-Hypoxic ( $NH_\mu$ ) case across SNR levels and time windows. We can note a consistent linearly increasing mean entropy as noise increases from the initial uncorrupted dataset for all time segments and entropy measurements. This finding is critical since healthier individuals (an individual with a lower sympathetic response) have higher entropy measurement producing more complex heart patterns and higher variability. Hence this informs us that if corruption occurs and the fiducial point is altered, subjects will appear healthier than they actually are.

#### F. RQ6 — Indistinguishable Independent Distributions

We demonstrated in RQ5 that there is a clear directional change in the measured entropy when noise is introduced into the signal. Because the noise does not alter the measured entropy distribution symmetrically (e.g., Figure 3) and thus shifts the entire distribution in one direction, we must then inquire about the implications of how this statistically affects our readings and to what degree it affects statistical independence. For example, if a subject has a lower entropy measurement, we perceive them to have a higher amount of physiological stress (i.e., increase sympathetic response). However if noise is introduced into the signal, causing an increase in entropy, what is the likelihood that the subject would statistically be viewed as someone not under stress?

Since hypoxia is known to cause an increase in sympathetic response, we address this research question through an approach similar to that of RQ4 using the NASA hypoxia dataset. We hypothesized that as SNR increases, it becomes increasingly difficult to discriminate ( $\alpha = 0.05$ ) between the two cohorts of hypoxia and non-hypoxia HRC measurements. The simulation is iterated 250 times for each window time segment, cohort, type of entropy measurement, and SNR for their respective cohorts, hypoxia and non-hypoxia. In Table VII, ( $Sig_\mu$ ) denotes the simulations where noise was only added to the hypoxia cohort data. However, ( $dSig_\mu$ ) denotes the simulations where noise was added to both the hypoxia and non-hypoxia data. The percentage of statistical tests that were significant from the 250 simulations are shown

TABLE VII  
EVALUATING THE UNCERTAINTY IN INDISTINGUISHABLE INDEPENDENT: (TRIALS= $\sim$  250).

Time Window	Target SNR	Approximate Entropy		Permutation Entropy		Sample Entropy	
		$Sig_\mu$	$dSig_\mu$	$Sig_\mu$	$dSig_\mu$	$Sig_\mu$	$dSig_\mu$
2	2	0%	0%	0%	70.0%	0%	0.4%
2	6	0%	0%	0.4%	83.2%	0%	0.8%
2	10	0%	0%	5.2%	88.8%	0%	0.2%
2	14	0%	0%	33.6%	92.4%	0%	1.2%
2	18	0%	0%	72.0%	96.4%	0%	0.4%
5	2	0%	0%	0%	24.4%	0%	18.0%
5	6	0%	0%	0%	47.2%	0%	51.6%
5	10	0%	0%	0%	69.6%	1.2%	88.4%
5	14	0%	0.8%	1.2%	78.4%	26.8%	96.4%
5	18	0%	0%	20%	85.2%	79.2%	99.6%

in each column of Table VII. This percentage is intended to provide contextual meaning of the likelihood that the two cohorts are statistically independent as a function of SNR, windowed time segment and type of entropy measurement. From Table VII can observe that there is a much smaller percentage of statistical independence when noise is only introduced to the hypoxia cohort. We already know from Table V, when no noise is introduced, that Permutation Entropy for a 2 minute time segment is significant, ( $p = 0.022$ ). However for the  $Sig_\mu$  case of corrupting only the hypoxic cohort with an SNR of 10, we expect the two classes to be independent only 5.2%. This means that based on the 250 simulations that we ran, only 13 of the 250 generated a p-value less than 0.05. Similarly for the  $dSig_\mu$  case when both the hypoxic and non-hypoxic cohorts are corrupted, 222 out of the 250 simulations generated a p-value less than 0.05. Thus, we expect to achieve significance 88.8% of the time when both signals are corrupted. This discrepancy between the  $Sig_\mu$  and  $dSig_\mu$  cases is due to the impact on how the noise shifts the means of the distributions. Hypoxia readings have a lower entropy measurement since the subject is under stress. However if noise is present strictly for the hypoxia data, it biases the distribution into making the collected data look as if the subject is not under stress and thus non-hypoxic. Therefore, it becomes more difficult to statistically distinguish between the two cohorts. Machine learning approaches specifically designed to detect anomalies in the data, such as single class Support Vector Machines (SVMs) [44], [45], are designed to learn the decision boundaries strictly on a single class. Thus as incoming data is altered by noise and the distribution is being shifted outside classifier boundary, the algorithm will classify it as an anomaly causing a false positive. This finding is paramount if we know that hypoxia HRV signals are being corrupted, and can be combated by introducing noise into the control cohort (i.e., non-hypoxia) to appropriately adjust the results of our statistical inference (re-train our ML algorithm for more generalized classifier boundaries to incorporate various noise levels) or simply bias the distribution based on your incoming SNR.

#### IV. CONCLUSION

Current literature about handling HRC calculations focuses on reducing the false positive and false negative rates of detecting the QRS waveform in the ECG signal. The location of the R-peak is typically extrapolated from the detected QRS waveform and little attention is given to how errors regarding the location of the R-peak can alter HRC readings. In this paper we present a Monte Carlo simulation framework that evaluates the effects of ECG signal corruption on the fiducial point of the R-peak and how it effects HRC measurements when using sample, approximate, and permutation entropy.

Through the use of Monte Carlo simulations, we are able to characterize PMF distributions and how the fiducial point shifts based on signal quality of the ECG. White noise was shown to cause higher perturbations in the fiducial point of the R-peak when compared to pink noise. This characterization allowed us to run additional Monte Carlo trials in order to evaluate changes in the precision of the proposed entropy measurement, in which permutation entropy is demonstrated to be most precise during corruption. From these findings, we utilized a secondary data set that addresses the sensitivity of demonstrating a statistical difference between hypoxia vs non-hypoxia caused by altered heart rate dynamics from the autonomic nervous system. This analysis showed that permutation entropy not only had better precision under noisy environments but was also sensitive statistically for 2 and 5 minute time segments between the two cohorts. Where as, Approximate entropy was not significant for either 5 or 10 minutes and sample entropy was only significant for 5 minute time segments. This work then demonstrated that as perturbations of the R-peak increased, as a function decreasing SNR, the entropy of the signal increased. We demonstrated that corrupted ECG signals entropy calculations have the potential to have biased means. However, permutation entropy showed to have a stronger precision and sensitivity which was able to still out perform sample and approximate entropy. Thus, sample and approximate entropy have a greater likelihood of showing inaccurate statistical changes in heart dynamics during hypoxia for 2 minute segments.

HRC calculations are critical to the implementation of statistical modeling techniques in numerous bio-informatics domains for physiological insight. Thus we were able to address critical fundamental design questions, allowing researchers to evaluate what type of entropy measurements are best, suggestions on how to handle ECG signal corruption of the fiducial point, and ideal time window segments base on the type of environment imposed on their telemetry system.

#### REFERENCES

- [1] N. T. Liu, L. C. Cancio, J. Salinas, and A. I. Batchinsky, "Reliable real-time calculation of heart-rate complexity in critically ill patients using multiple noisy waveform sources," *J Clin Monit Comput*, vol. 28, no. 2, pp. 123–131, Apr 2014.
- [2] J. Harrison, K. Izzetoglu, H. Ayaz, B. Willems, S. Hah, H. Woo, P. A. Shewokis, S. C. Bunce, and B. Onaral, "Human Performance Assessment Study in Aviation Using Functional Near Infrared Spectroscopy BT - Foundations of Augmented Cognition: 7th International

- Conference, AC 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013. Procees," D. D. Schmorrow and C. M. Fidopiastis, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 433–442.
- [3] J. E. Mercado, L. Reinerman-Jones, D. Barber, and R. Leis, "Investigating Workload Measures in the Nuclear Domain," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, 2014, pp. 205–209.
- [4] S. G. Charlton, "Measurement of cognitive states in test and evaluation," Mahwah, NJ, pp. 97–126, 2001.
- [5] G. F. Wilson, "Psychophysiological Test Methods and Procedures," in *Handbook of Human Factors Testing and Evaluation*, 2nd ed., S. G. Charlton and T. G. O'Brien, Eds. Mahwah, NJ: LAWRENCE ERLBAUM ASSOCIATES, 2001, ch. 7, pp. 127–156.
- [6] B. Cain, "A Review of the Mental Workload Literature," Defence Research and Development Canada Toronto, Toronto, Ontario, Tech. Rep., 2007.
- [7] L. Salahuddin, J. Cho, M. G. Jeong, and D. Kim, "Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2007, pp. 4656–4659.
- [8] D. J. Plews, P. B. Laursen, J. Stanley, A. E. Kilding, and M. Buchheit, "Training adaptation and heart rate variability in elite endurance athletes: Opening the door to effective monitoring," *Sports Medicine*, vol. 43, no. 9, pp. 773–781, 2013.
- [9] K. Itao, T. Umeda, G. Lopez, and M. Kinjo, "Human recorder system development for sensing the autonomic nervous system," in *2008 IEEE Sensors*, Oct 2008, pp. 423–426.
- [10] M. M. Teixeira, Renata Roland and, T. V. d. S. Santos, J. T. M. Bernardes, L. G. Peixoto, O. L. Bocanegra, M. B. Neto, and F. S. Espindola, "Chronic stress induces a hyporeactivity of the autonomic nervous system in response to acute mental stressor and impairs cognitive performance in business executives," *PLoS ONE*, vol. 10, no. 3, 2015.
- [11] B. Cinaz, R. La Marca, B. Arnrich, and G. Tröster, "Monitoring of mental workload levels," *Proceedings of the IADIS International Conference e-Health 2010, EH, Part of the IADIS Multi Conference on Computer Science and Information Systems 2010, MCCSIS 2010*, pp. 189–193, 2010.
- [12] R. W. Backs, "Going beyond heart rate: Autonomic space and cardiovascular assessment of mental workload," *The International Journal of Aviation Psychology*, vol. 5, no. 1, pp. 25–48, 1995.
- [13] J. Sztajzel, "Heart rate variability: A noninvasive electrocardiographic method to measure the autonomic nervous system," *Swiss Medical Weekly*, vol. 134, no. 35-36, pp. 514–522, 2004.
- [14] Y. Zhong, K.-M. Jan, K. H. Ju, and K. H. Chon, "Quantifying cardiac sympathetic and parasympathetic nervous activities using principal dynamic modes analysis of heart rate variability," *American journal of physiology. Heart and circulatory physiology*, vol. 291, no. 3, pp. H1475–83, 2006.
- [15] G. Berntson and J. Stowell, "Ecg artifacts and heart period variability: don't miss a beat!" *Psychophysiology*, vol. 35, no. 1, pp. 127–32, 1998.
- [16] S. Laborde, E. Mosley, and J. F. Thayer, "Heart rate variability and cardiac vagal tone in psychophysiological research recommendations for experiment planning, data analysis, and data reporting," *Frontiers in Psychology*, vol. 8, p. 213, 2017.
- [17] C. Schubert, M. Lambertz, R. Nelesen, W. Bardwell, J.-B. Choi, and J. Dimsdale, "Effects of stress on heart rate complexity a comparison between short-term and chronic stress," vol. 80, pp. 325–32, 12 2008.
- [18] N. Napoli and L. Barnes, "A Dempster-Shafer Approach for Corrupted Electrocardiograms Signals," *Twenty-Ninth International Florida Artificial Intelligence Research Society Conference*, 2016.
- [19] A. H. Khandoker, H. F. Jelinek, and M. Palaniswami, "Identifying diabetic patients with cardiac autonomic neuropathy by heart rate complexity analysis," *Biomed Eng Online*, vol. 8, p. 3, 2009.
- [20] S. J.-J. Leistedt, P. Linkowski, J. Lanquart, J. E. Mietus, R. Davis, A. L. Goldberger, and M. Costa, "Decreased neuroautonomic complexity in men during an acute major depressive episode: analysis of heart rate dynamics," *Translational Psychiatry*, vol. 8, pp. 1–7, July 2011.
- [21] Q. Zheng, C. Chen, Z. Li, A. Huang, B. Jiao, X. Duan, and L. Xie, "A novel multi-resolution svm (mr-svm) algorithm to detect ecg signal anomaly in we-care project," in *2013 ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC)*, Feb 2013, pp. 1–6.

- [22] A. Ukil, S. Bandyopadhyay, C. Puri, and A. Pal, "Heart-trend: An affordable heart condition monitoring system exploiting morphological pattern," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 6260–6264.
- [23] G. Clifford, W. Long, G. Moody, and P. Szolovits, "Robust parameter extraction for decision support using multimodal intensive care data," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1887, pp. 411–429, 2009.
- [24] N. T. Liu, A. I. Batchinsky, L. C. Cancio, and J. Salinas, "The impact of noise on the reliability of heart-rate variability and complexity analysis in trauma patient," *Comp. in Biology and Medicine*, vol. 43, no. 11, pp. 1955 – 1964, 2013.
- [25] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, p. e215, Jun. 2000.
- [26] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. Jour. of Physio. - Heart and Circ. Physio.*, vol. 278, no. 6, pp. 2039–2049, 2000.
- [27] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Phys. Rev. Lett.*, vol. 89, Jul 2002.
- [28] S. M. Pincus, "Approximate entropy as a measure of system complexity," *PNAS*, vol. 88, no. 6, pp. 2297–2301, 1991.
- [29] M. Riedl, A. Müller, and N. Wessel, "Practical considerations of permutation entropy," *The Euro. Phys. Jour. Special Topics*, vol. 222, no. 2, pp. 249–262, 2013.
- [30] K. Kim, H. Baek, Y. Lim, and K. Park, "Effect of missing rr-interval data on nonlinear heart rate variability analysis," *Comp. Methods and Programs in Biomed.*, vol. 106, no. 3, pp. 210 – 218, 2012.
- [31] R. J. Ellis, B. Zhu, J. Koenig, J. F. Thayer, and Y. Wang, "A careful look at ecg sampling frequency and r-peak interpolation on short-term measures of heart rate variability," *Physiological Measurement*, vol. 36, no. 9, p. 1827, 2015.
- [32] K. K. Kim, Y. G. Lim, J. S. Kim, and K. S. Park, "Effect of missing rr-interval data on heart rate variability analysis in the time domain," *Physiological Measurement*, vol. 28, no. 12, p. 1485, 2007.
- [33] K. Kim, J. Kim, Y. Lim, and K. Park, "The effect of missing rr-interval data on heart rate variability analysis in the frequency domain," *Physiological Measurement*, vol. 30, no. 10, pp. 210 – 218, 2009.
- [34] F. Gustafsson, "Determining the initial states in forward-backward filtering," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 988–992, Apr 1996.
- [35] C. Chew and E. Zahedi, "Limb cardiovascular system identification using adaptive filtering," *International Federation for Medical and Biological Proceedings*, pp. 406–409, 2006.
- [36] H. Limaye and V. Deshmukh, "Ecg noise sources and various noise removal techniques: A survey," *International Journal of Application or Innovation in Engineering and Management*, vol. 5, pp. 406–409, 2016.
- [37] G. D. Clifford, F. Azuaje, and P. McSharry, "Ecg statistics, noise, artifacts, and missing data," in *Advanced Methods And Tools for ECG Data Analysis*. Norwood, MA, USA: Artech House, Inc., 2006, ch. 3, pp. 70–71.
- [38] N. J. Kasdin, "Discrete simulation of colored noise and stochastic processes and 1/f alpha; power law noise generation," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 802–827, May 1995.
- [39] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomed. Engin.*, vol. BME-32, no. 3, pp. 230–236, 1985.
- [40] C. Bandt and B. Pompe, "Permutation Entropy: A Natural Complexity Measure for Time Series," *Phys. Review Letters*, vol. 88, no. 17, Apr 2002.
- [41] D. Gradwell, "Hypoxia and hyperventilation," in *Ernstings's Aviation Medicine*, 4th ed., D. Gradwell and D. J. Rainford, Eds. CRC Press, 2006, ch. 3, p. 4156.
- [42] C. Stephens, K. Kennedy, B. Crook, R. Williams, M. Last, and P. Schutte, "Mild normobaric hypoxia exposure for human-autonomy system testing," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 61, 2017.
- [43] C. Stephens, K. Kennedy, N. Napoli, M. Demas, B. Crook, R. Williams, M. Last, and P. Schutte, "Effects on task performance and psychophysiological measures of performance during normobaric hypoxia exposure," *Proceedings of the International Symposium on Aviation Psychology*, vol. Dayton, OH, 2017.
- [44] J. Su, Y. Long, X. Qiu, S. Li, and D. Liu, *Anomaly Detection of Single Sensors Using OCSVM\_KNN*. Cham: Springer International Publishing, 2015, pp. 217–230.
- [45] Y. Chen, J. Qian, and V. Saligrama, "A new one-class svm for anomaly detection," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 3567–3571.