**A red sign with white text

Description automatically generated**

**MACHINE LEARNING FOR ELECTRODE MOTION REDUCTION IN ELECTROCRADIOGRAM SIGNALS**

by

Ben Russell

Student Number : 40178580

Supervisor : Seán Mcloone

A Master’s Thesis

Submitted to the Department of Electronics, Electrical Engineering and Computer Science

Queens University Belfast

In Fulfilment of the Requirements

For a Master’s Degree

September 2024

**Abstract**

The aim of this thesis is to explore the application of artificial intelligence (AI) for the removal of electrode motion artifacts from ECG signals. A comprehensive reference database was developed, consisting of both clean and noisy ECG signals. The clean ECG signals were generated using three ordinary differential equations (ODE) to accurately model the heart's electrical activity. Noisy signals were synthesized by extracting noise from the MIT-BIH Noise Stress Test Database (NSTDB) and using an autoregressive model to predict new noise shapes, while maintaining the same statistical properties. Each noise instance was then scaled to various Signal-to-Noise Ratio (SNR) levels and superimposed on the clean signals to create a diverse set of corrupted ECG signals.

We developed and evaluated multiple machine and deep learning models to address the problem of noise removal. These models included Support Vector Machines (SVM), Random Forests, Principal Component Analysis (PCA) with K-Nearest Neighbours (KNN), Generative Adversarial Networks (GANs), Autoencoders, Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) units. Additionally, A custom loss function was designed to enhance the performance of these models.

The performance of each algorithm was rigorously validated using both synthetic and real ECG databases. Key performance metrics such as Root Mean Squared Error (RMSE), correlation coefficient, SNR improvement, and Wavelet Energy Based Diagnostic Distortion (WEDD) were employed to assess the efficacy of the noise removal techniques. Additionally, the results of the proposed AI-based methods were compared against traditional time-frequency based approaches to establish their relative advantages.

This study demonstrates the potential of advanced AI models in improving the quality of ECG signal processing by effectively mitigating the impact of electrode motion artifacts. The findings suggest that these AI techniques can offer significant improvements over conventional methods, paving the way for more accurate and reliable ECG analysis in clinical settings. Furthermore, since noise is a common problem across many disciplines, the models developed in this work may be utilised via transfer learning to solve problems in other industries.

**Acknowledgments**

**Table of Contents**

**List of Figures**

**List of Tables**

**Abbreviations and Symbols**

**Chapter 1: Introduction**

The Electrocardiogram (ECG) is a vital tool in modern medicine, offering a non-invasive and direct method for monitoring the electrical activity of the heart. By recording the heart’s electrical signals through electrodes placed on the skin, the ECG provides essential insights into the rhythmic patterns and conditions affecting the heart’s function. This capability makes the ECG indispensable for diagnosing various cardiac abnormalities, such as arrhythmia’s, heart disease, and myocardial infarction.

An ECG signal represents the sum of electrical potentials generated by the heart muscle during each cardiac cycle. The signal is characterized by a series of waves and complexes, most notably the p wave, QRS complex, and T wave, each corresponding to specific phases of the heart’s electrical cycle. The P wave indicates atrial depolarization, the QRS complex represents the ventricular depolarization, and the T wave is associates with ventricular repolarization. Analysing these components allows healthcare professionals to assess the timing of cardiac events, the presence of abnormal rhythms, and the health of the heart muscle.

Monitoring ECG signals is crucial for several reasons:

* **Early detection**: Continuous or periodic ECG monitoring can help detect early signs of heart disease, even before symptoms appear.
* **Diagnosis**: ECG readings are essential for diagnosing various cardiac conditions, including arrhythmia’s, ischemic heart disease, and congenital heart defects.
* **Treatment monitoring**: For patients undergoing treatment for heart conditions, ECG monitoring provides valuable feedback on the effectiveness of interventions, such as medications, pacemaker function, and recovery after cardiac procedures.
* **Prognosis**: ECG findings can inform prognosis, helping predict the likelihood of cardiac events such as sudden cardiac death or recurrence of heart attacks.

A diagram of a graph

Description automatically generated**Figure 1** shows an example of the time series electrical signal obtained from one heartbeat. It should be noted that amplitude of each waveform is in volts. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9332869/

Figure 1 iLLustration of a typical ECG signal obtained from a healthy heartbeat.

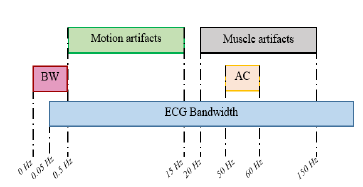
Despite its importance, the accuracy of ECG monitoring can be compromised by various sources of noise that can significantly affect the quality of the signal. There are 4 major sources of noise in ECG signals, these are 1) Baseline Wander (BW), 2) Powerline Interference (PL), 3) Muscle Artefact (MA) and 4) **Electrode Motion (EM)** [1]. BW and PL have a relatively unique frequency content and thus are easily removed by simple digital filters [2]. MA and EM noise are more challenging to remove as they can have a wide frequency content that overlaps with that of the ECG signal, specifically the PQRST complex [3]. This is shown in **figure 2** below, it should be noted that AC in this figure is referring to PL mentioned previously.

Figure 2 Illustration of noise vs signal frequency overlap.

Electrode motion noise in ECG signals refers to the interference caused by the movement of electrodes attached to the skin. When electrodes move, even slightly, they can pick up additional electrical activity not related to the heart's electrical signals. This extra activity is seen as noise in the ECG signal, which can distort the true reading. It's particularly problematic during physical activity or if the electrode doesn't adhere well to the skin. This noise appears as irregular spikes or a fuzzy baseline in the ECG trace, making it challenging to accurately interpret the heart's electrical activity. Managing electrode motion noise is crucial for ensuring reliable ECG readings, especially in scenarios requiring patient movement or long-term monitoring. **Figure 3** shows an example of an ECG signal corrupt with EM noise, it is evident that no clinical interpretation of this signal could be made by an expert or algorithm. This reinforces the need for accurate and reliable filtering.

A graph showing a graph

Description automatically generated with medium confidenceEM is difficult to remove from an ECG signal for various reasons. Firstly, the signal attenuated by electrode motion can be very similar to the ECG signal itself, this similarity can make it difficult to remove with digital filters which use both time and frequency domain characteristics to separate the two signals. The variability of EM noise also adds to the difficulty to remove the noise, this comes from the wide range of movements a subject can undergo such as running, walking, jumping or any other form of activity. Each activity will return a different characteristic shape of noise. Electrode placement can also vary the shape of the noise signal. Finally, individual differences such as skin type, amount of hair and other factors influencing the electrode-skin contact will affect the amount of EM noise added during movement, it is important that algorithms can deal with this problem effectively.

Figure 3 ECG SIGNAL CORRUPT WITH ELECTRODE MOTION NOISE.

The objective of this work is to develop an intelligent algorithm that can effectively reduce the amount of EM noise in an ECG signal. An up-to-date review on where the field of denoising ECG signals with both traditional filtering, and the more recent developments by using machine and deep learning approaches will be presented. This report will be structured as follows:

**Literature Review** – This section will give an overview of ECG signal processing techniques currently employed. It will then investigate Electrode Motion specifically to get an appreciation of why the noise arises and why it is problematic. Existing methods for removing EM noise will be presented, and their effectiveness will be discussed. Finally, Machine and Deep learning in biomedical signal processing will be introduced and any gaps in the literature will be identified.

**Methodology –** This section will focus on the strategy followed to build and train the algorithm. A big portion of this section will focus on the development of a large and diverse training database which was employed as a main novelty in this work to improve on previous models. This will cover the generation of clean lead II ECG signals alongside the synthetic generation of EM noise signals and their combination to form a noise corrupt ECG signal. Finally, this section will cover any technical information on all algorithm development work and propose a method for validation.

**Implementation –** This section will display all software tools used throughout the project life-cycle. Tools such as version control, data storage, training hardware and libraries/modules will be discussed. The overall workflow of the project will be presented, and any limitations will be discussed.

**Results and Discussion –** This section will present a detailed overview on the results. Visual examples of the denoised signals will be presented and compared to alternative methods. Furthermore, a quantitative results section will be presented using metrics such as Root Mean Square Error (RMSE), Signal to Noise Ratio (SNR) and Cross Correlation (CC). The meaning of these results will be discussed and the applicability to real life scenarios.

**Chapter 2: Literature Review**

ECG signal processing is a critical aspect of cardiovascular disease diagnosis and monitoring. ECG signals are often contaminated by various types of noise, which can hinder accurate interpretation. This overview covers the primary noise sources in ECG signals and the processing techniques used to mitigate these interferences. A detailed description of each noise source can be seen in table 1 below:

|  |  |  |
| --- | --- | --- |
| **Noise Source** | **Description** | **Impact** |
| Power Line Interference | PLI is caused by the 50/60 Hz power supply frequency coupling with the ECG signal. | It introduces a sinusoidal interference that can obscure the ECG signal components. |
| Baseline Wander | Low-frequency noise resulting from patient movement, respiration, and electrode impedance changes | Causes slow variations in the baseline of the ECG signal, making it difficult to identify the true isoelectric line. |
| Muscle Artefact (Electromyographic) | High-frequency noise generated by voluntary or involuntary muscle contractions. | Overlaps with the ECG signal frequency, particularly affecting the QRS complex. |
| Electrode Motion | Noise caused by the movement of electrodes relative to the skin. | Produces transient baseline shifts and spikes. Overlaps with all features in an ECG. |

Table 1 Table showing a description and the impact of each noise source to an ECG signal.

To combat the noise sources mentioned above, various techniques are employed by scientists, engineers and researchers. The most popular technique is the use of digital filters, namely low-pass and high-pass filters. Low-pass filters removed high frequency noise while high-pass filters remove low-frequency noise. Different filter designs can be employed, each with their own trade-offs, for example, Butterworth, Chebyshev and Elliptic filters are commonly used however the specific details of these is out of scope of this report. Notch filters are also used and are specifically designed for the removal of PL noise by targeting the 50/60Hz frequency.

Wavelet Transforms are another commonly used technique to remove noise, this works by decomposing the ECG signal into different frequency components, allowing for the selective removal of noise. Namely, the Discrete Wavelet Transform (DWT) is often used for its ability to handle non-stationary signals, that is, signals that statistical properties vary over time.

Empirical Mode Decomposition (EMD) decomposes ECG signals into Intrinsic Mode Functions (IMFs) for adaptive filtering. These are especially useful for dealing with non-linear and non-stationary signals.

Principle Component Analysis (PCA) and Independent Component Analysis (ICA) are used to separate the ECG signal from a noise signal based on statistical properties. PCA reduces the dimensionality of the data, while ICA separates independent sources.

Adaptive filtering uses digital filters previously mentioned, however dynamically updates the filter parameters to capture time varying noise. Adaptive filters utilise the Least Mean Squares (LMS) and Recursive Least Squares (RLS) algorithms to adjust the parameters.

Finally, Ensemble Averaging simply improves the Signal to Noise Ratio by averaging multiple ECG cycles. This reduces random noise effectively however is ineffective when continuous noise is present in the ECG signal.

1. **Digital Filtering**:
   * **Low-Pass and High-Pass Filters**:
     + **Usage**: Low-pass filters remove high-frequency noise (e.g., muscle artifacts), while high-pass filters remove low-frequency noise (e.g., baseline wander).
     + **Design**: Butterworth, Chebyshev, and Elliptic filters are commonly used.
     + **Reference**: "Digital Signal Processing: Principles, Algorithms, and Applications" by John G. Proakis and Dimitris K. Manolakis.
   * **Notch Filters**:
     + **Usage**: Specifically designed to remove PLI by targeting the 50/60 Hz frequency.
     + **Design**: Adaptive notch filters can adjust to varying interference frequencies.
     + **Reference**: "Adaptive Filter Theory" by Simon Haykin.
2. **Wavelet Transform**:
   * **Usage**: Decomposes the ECG signal into different frequency components, allowing for the selective removal of noise.
   * **Method**: Discrete Wavelet Transform (DWT) is often used for its ability to handle non-stationary signals.
   * **Reference**: "Wavelet Transforms and Their Applications" by Lokenath Debnath and Firdous Ahmad Shah.
3. **Empirical Mode Decomposition (EMD)**:
   * **Usage**: Decomposes signals into Intrinsic Mode Functions (IMFs) for adaptive filtering.
   * **Advantage**: Effective in dealing with non-linear and non-stationary data.
   * **Reference**: "Hilbert-Huang Transform and Its Applications" by Norden E. Huang and Samuel S. P. Shen.
4. **Principal Component Analysis (PCA) and Independent Component Analysis (ICA)**:
   * **Usage**: Used for separating the ECG signal from noise based on statistical properties.
   * **Method**: PCA reduces dimensionality, while ICA separates independent sources.
   * **Reference**: "Independent Component Analysis: A Tutorial Introduction" by James V. Stone.
5. **Adaptive Filtering**:
   * **Usage**: Adjusts filter parameters dynamically to track and remove time-varying noise.
   * **Method**: Least Mean Squares (LMS) and Recursive Least Squares (RLS) algorithms are common.
   * **Reference**: "Adaptive Signal Processing" by Bernard Widrow and Samuel D. Stearns.
6. **Ensemble Averaging**:
   * **Usage**: Enhances the signal-to-noise ratio by averaging multiple ECG cycles.
   * **Advantage**: Reduces random noise effectively.
   * **Reference**: "Biomedical Signal Processing: Principles and Techniques" by D. C. Reddy.

The effective processing of ECG signals requires a comprehensive understanding of various noise sources and the application of appropriate filtering techniques. The choice of method depends on the specific noise characteristics and the clinical requirements of the ECG analysis.

#### References

* Proakis, J. G., & Manolakis, D. K. (1996). Digital Signal Processing: Principles, Algorithms, and Applications. Prentice-Hall.
* Haykin, S. (2002). Adaptive Filter Theory. Prentice-Hall.
* Debnath, L., & Shah, F. A. (2015). Wavelet Transforms and Their Applications. Birkhäuser.
* Huang, N. E., & Shen, S. S. P. (2005). Hilbert-Huang Transform and Its Applications. World Scientific.
* Stone, J. V. (2004). Independent Component Analysis: A Tutorial Introduction. MIT Press.
* Widrow, B., & Stearns, S. D. (1985). Adaptive Signal Processing. Prentice-Hall.
* Reddy, D. C. (2005). Biomedical Signal Processing: Principles and Techniques. McGraw-Hill.

By applying these techniques, clinicians and researchers can achieve more accurate and reliable ECG signal analysis, leading to better diagnosis and monitoring of cardiovascular conditions.

The most widely accepted explanation of EM artifacts is from Tam and Webster [4], who found that that the change in skin potential at the skin-electrolyte interface is the main source of motion artefacts. According to Edelberg’s skin model [3], the skin potential is determined by the potential across the sweat duct membrane, the total resistance in the sweat duct, the potential across the epidermis barrier membrane, and the total series resistance of the epidermis. Changes in any four variables will cause changes in the skin potential. [REQORD THIS].

The performance of many different solutions to reduce EM in ECG signals have been investigated in many studies [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23]. Among these, adaptive filters have been shown to be promising at removing EM noise, however adaptive filters require another reference signal, making them computationally complex. Although there is evidence that adaptive filters with reference signals such as accelerometer, or skin-electrode impedance signal can be useful to reduce motion artefact, there is no conclusive evidence which reference signal is performs best [24]. Furthermore, some findings contradict each other, while others are inconclusive.

Machine learning (ML) has significantly impacted biomedical signal processing by offering robust tools for the analysis and interpretation of complex biomedical data. Traditional signal processing techniques often struggle to handle the vast variability and non-stationarity of biomedical signals [Clifford, G. D., Azuaje, F., & McSharry, P. E. (2006). *Advanced methods and tools for ECG data analysis*. Artech House.] such as electrocardiograms (ECGs), electroencephalograms (EEGs), and electromyograms (EMGs). Machine learning algorithms, particularly those based on supervised and unsupervised learning, provide a way to automate feature extraction, classification, and pattern recognition in these signals. For instance, Support Vector Machines (SVMs) and Random Forests have been effectively used for detecting arrhythmias in ECG signals [Abagaro, A.M., Barki, H., Ayana, G. *et al.* Automated ECG Signals Analysis for Cardiac Abnormality Detection and Classification. *J. Electr. Eng. Technol.* **19**, 3355–3371 (2024). <https://doi.org/10.1007/s42835-024-01902-y>] and classifying different stages of sleep using EEG signals [Rashidi, S., Asl, B.M. Strength of ensemble learning in automatic sleep stages classification using single-channel EEG and ECG signals. *Med Biol Eng Comput* **62**, 997–1015 (2024). https://doi.org/10.1007/s11517-023-02980-2]. The ability of machine learning models to learn from data and improve over time makes them invaluable for predictive diagnostics and personalized medicine. Furthermore, these algorithms can integrate heterogeneous data sources, offering a comprehensive analysis that considers various physiological parameters. References such as "Biomedical Signal Processing and Machine Learning for Cardiovascular Diseases" by Ling et al. (2019) highlight the efficacy of ML methods in improving diagnostic accuracy and patient outcomes.

In the context of noise reduction in time-series signals, traditional ML methods have been used to attempt and denoise various types of signals. However, these face several issues when it comes to denoising ECG signals. Firstly, traditional ML algorithms like support vector machines (SVM), K nearest neighbours (KNN) and decision trees rely heavily on feature extraction which requires a lot of domain expertise to identify the right features characteristic of an ECG signal. Furthermore, ECG signals are typically high-dimensional data which traditional ML models can struggle with, especially when the SNR is low. ECG signals have significant temporal dependencies that need to be considered for effective denoising. Traditional ML models do not inherently model temporal relationships, making it challenging to capture the sequential nature of the data. ECG signals are non-stationary, meaning their statistical properties change over time, traditional models generally assume stationary data or require the data to be pre-processed to remove non-stationarity. In this case, the noise adds a lot of non-stationarity so it crucial that we do not remove this from the dataset. Finally, the characteristic shape of EM noise in ECG signals can vary significantly in terms of frequency and amplitude, ML models require explicit modelling which can be complex and time consuming. All of these points provide evidence as to why traditional learning algorithms are not suitable for this problem and imply that a learning technique designed for non-linear, non-stationary and highly variable data may be required.

Deep learning (DL), a subset of machine learning, has revolutionized biomedical signal processing by providing more sophisticated methods for analysing complex biomedical data. Unlike traditional machine learning techniques that require manual feature extraction, deep learning models automatically learn hierarchical representations of the data, which can capture intricate patterns and temporal dependencies. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are among the most popular architectures used in this domain. CNNs have shown exceptional performance in analysing spatial data and have been used to detect abnormalities in ECG signals, such as myocardial infarctions and atrial fibrillations [Hannun et al. (2019) in "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks"]. RNNs, particularly Long Short-Term Memory (LSTM) networks, are effective in modelling sequential data, making them suitable for time-series analysis in EEG and EMG signals. These models have demonstrated superior accuracy and robustness in various biomedical applications, from disease diagnosis to brain-computer interfaces.

Despite significant advancements, several gaps remain in the application of machine learning and deep learning to biomedical signal processing. One of the primary challenges is the lack of large, high-quality annotated datasets, which are crucial for training robust and generalizable models. Many existing studies rely on small or proprietary datasets, limiting the reproducibility and comparability of results. Moreover, most current research focuses on specific types of biomedical signals, such as ECG or EEG, with less attention given to multimodal data integration, which could provide a more holistic understanding of physiological conditions. Another gap is the limited exploration of model interpretability and explainability. In clinical settings, it is vital to understand how a model arrives at a decision to ensure its reliability and to gain the trust of healthcare professionals. Additionally, there is a need for more research on the real-time implementation of these models in clinical practice, considering computational efficiency and scalability. Addressing these gaps requires a concerted effort to develop open-access datasets, improve model interpretability, and focus on translational research that bridges the gap between algorithm development and clinical application. References such as "Challenges and Opportunities in Machine Learning for Biomedical Signal Processing" by Johnson et al. (2020) discuss these issues in greater detail.

### References

* Ling, Y., et al. (2019). Biomedical Signal Processing and Machine Learning for Cardiovascular Diseases. Journal of Medical Systems, 43(6), 153.
* Hannun, A. Y., et al. (2019). Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. Nature Medicine, 25(1), 65-69.
* Johnson, A. E. W., et al. (2020). Challenges and Opportunities in Machine Learning for Biomedical Signal Processing. IEEE Transactions on Biomedical Engineering, 67(5), 1244-1261.

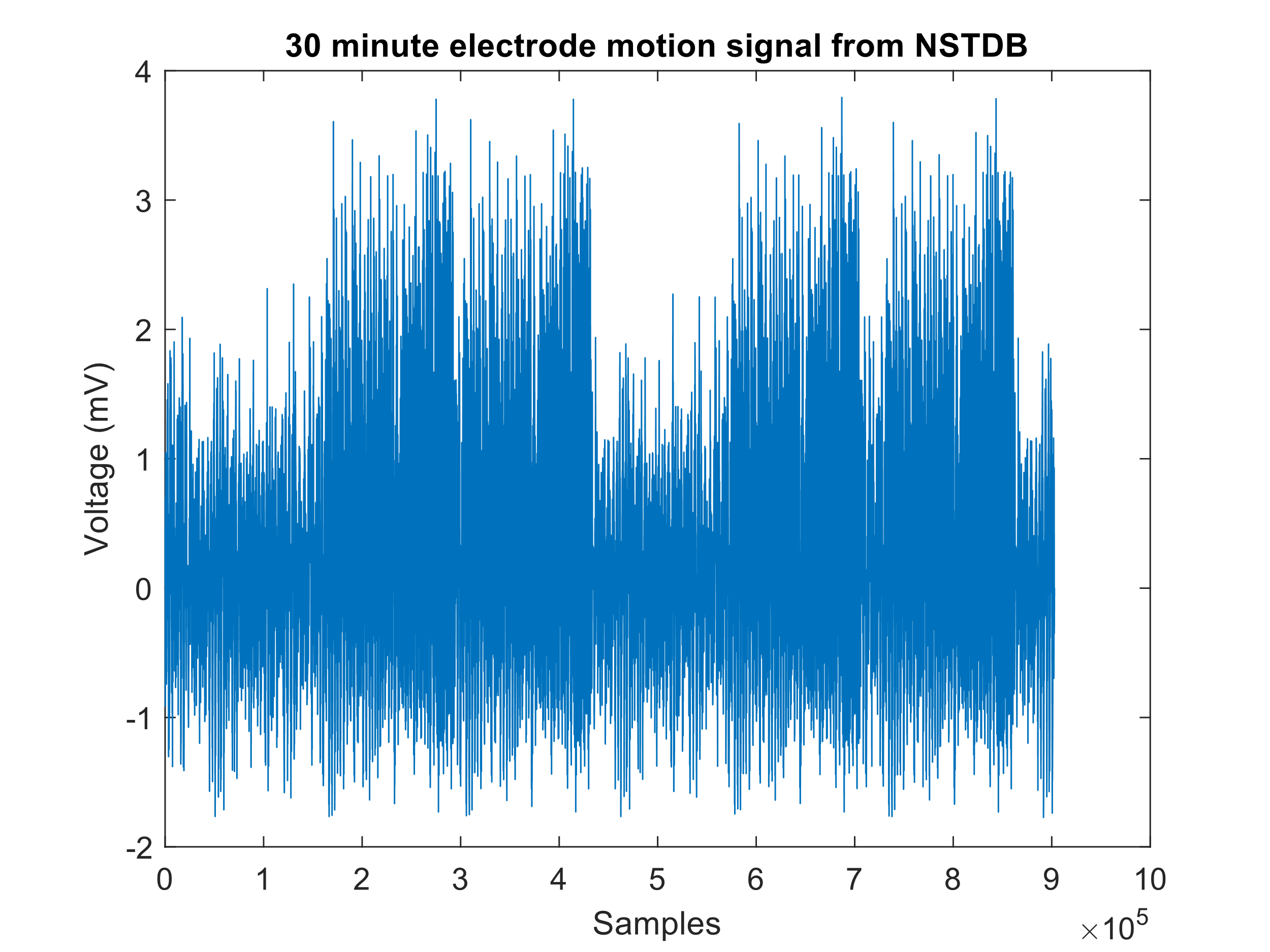
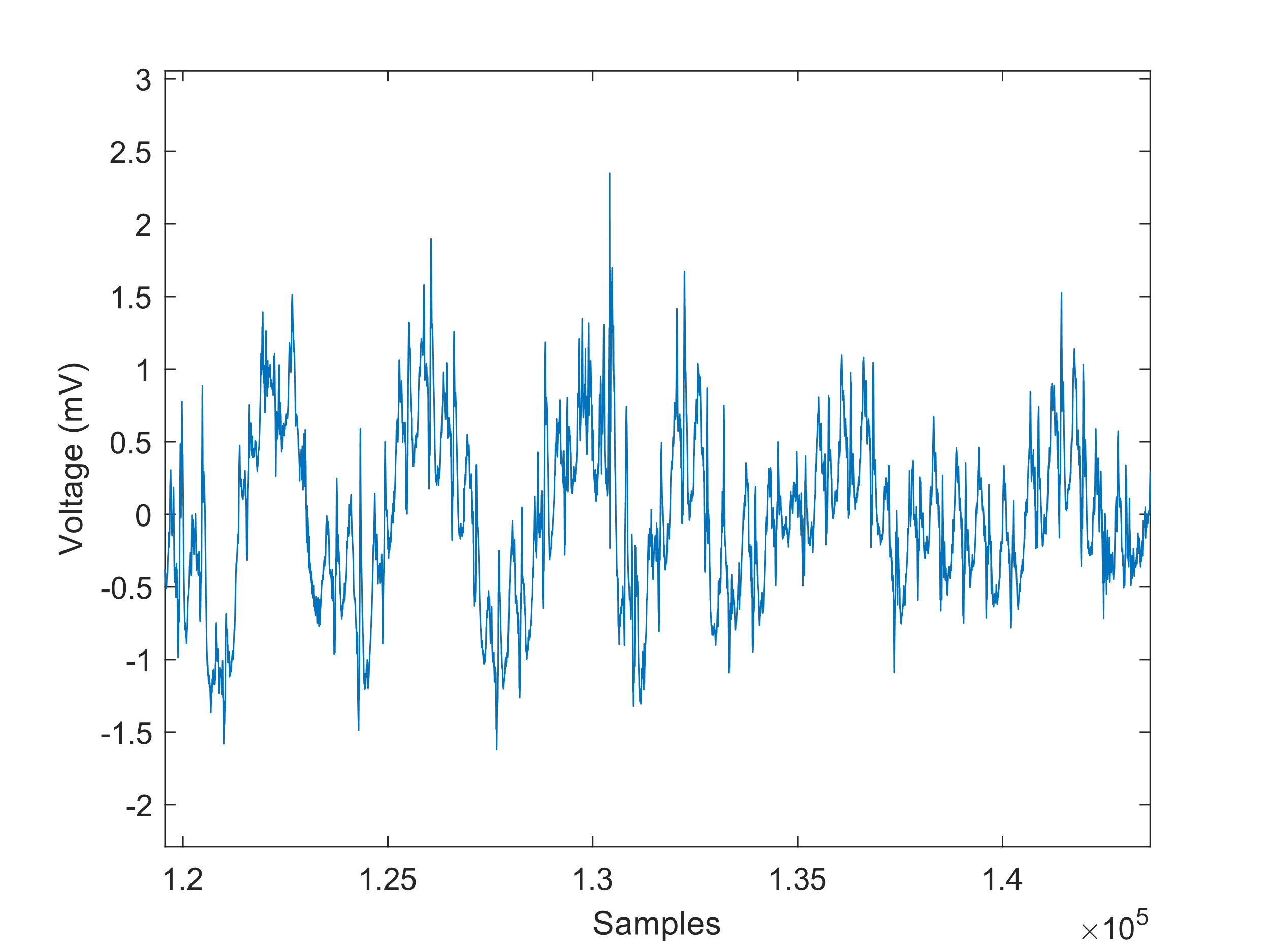
**Chapter 3: Methodology**

* Data Collection

The Physionet Database is commonly used as a publicly available database with a large amount of annotated ECG signals that can be used to both develop and validate biomedical algorithms. It is commonly used for the development of classification algorithms where the aim is to classify various cardiac abnormalities present. Certain databases have also been used to develop and validate Machine Learning Denoising algorithms, for example, XXX et all trained on the MIT-BIH database. This database was selected as the ECG records are relatively clean, and thus provide a suitable ground truth. However, this database is small and only contains 48 records, furthermore, although the signals are relatively clean, they will still contain some noise since they are real signals. These two factors were deemed enough to attempt on the development of a much larger, clean reference database that could be used to develop and validate intelligent signal conditioning algorithms.

Due to the limitations on using real ECG signals, it was deemed necessary to simulate clean ECG signals using an established model. The ECGSYN [] tool was used to simulate a large amount of clean signals with different waveform morphologies and heart rates. It was important that each signal had different morphologies too add variance into the database, the amplitude, width and slope of the P. QRS and T waves will vary between people and so this is information that should be passed into the algorithm. Furthermore, Heart Rate is another parameter that will change between subjects and so 5 different Heart Rates were generated for each clean signal.

* + Preprocessing

Noise signals were extracted from Physionets NSTDB, the EM signal was extracted alongside the sampling frequency which was 360Hz. The signal was immediately up sampled to 500Hz to match the clean signals generated later in this report. The raw noise signal after up sampling can be seen in figure X below:

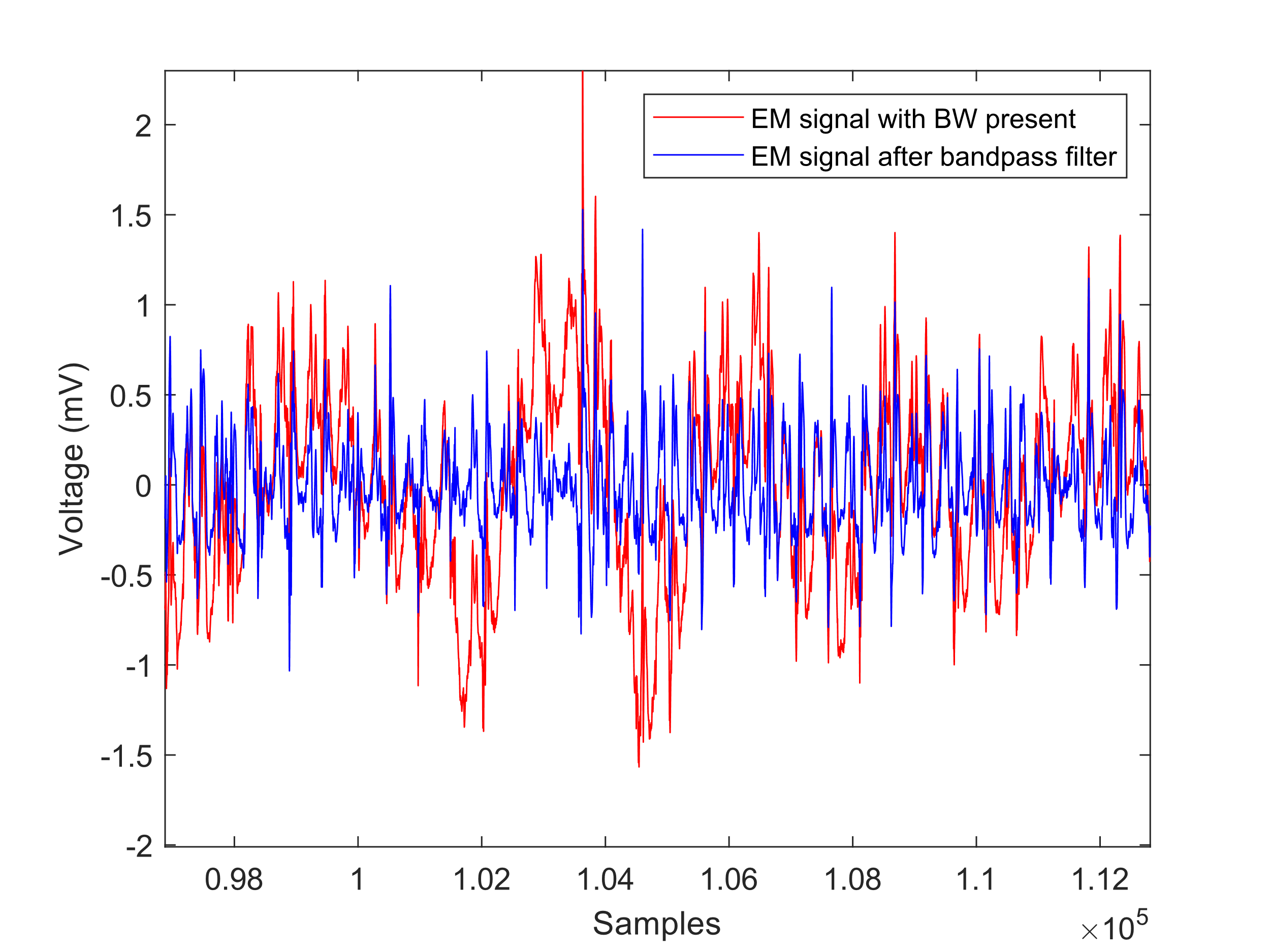
There was a small amount of BW noise identified within this signal, this noise source is out of scope for this report and so it was deemed necessary to remove this low frequency content. The signal shown above was passed through a 20th order type II Chebyshev bandpass filter with 60dB attenuation. The cut-off frequencies were set as 1Hz and 50Hz to ensure a maximum amount of EM noise is retained. The final prepared noise signal can be seen in figure X below:

Figure 4 Shows the effect of the bandpass filter.

Clean Signal Generation

To generate a dataset of clean ECG signals with varying morphologies and heart rates, Physionets ECGSYN tool was employed. An algorithm was developed with this tool embedded to ensure the following requirements were met:

1. Each clean signal would have different waveform parameters. The width, amplitude and slope angle would vary.
2. All clean signals were generated at 500Hz.
3. Each clean signal would contain morphologies representative of what would be seen in real life.
4. The clean signals would be generated at 50, 60, 70,80, 90 and 100 Beats Per Minute (BPM)

To meet these requirement (1), The algorithm utilised a Latin Hypercube Sampling (LHS) to randomly sample parameter values between defined limits. These limits refer to the minimum and maximum widths, amplitudes and slopes of each waveform. The minimum and maximum parameters set in this study can be seen in table X below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Figure 5 Parameter boundaries set for lhs.

A graph with orange dots and blue lines

Description automatically generatedLHS is utilised due to its’ ability to control sample points in the parameter. That is, we can ensure that the sampling point distribution is close to the probability density function (PDF). In this case, we assume that each parameter has an equal probability of falling in any region between the defined boundaries, thus, LHS avoids the probability that all sampling points come from the same local region. This is depicted in figure X below:

Figure 6 Depiction of LHS METHODOLOGY

Practically, and in the context of this work, this means that feature waves (P, QRS, T) will have a uniform distribution of sample values that covers the entire range.

When each parameter value is set based on the above, the ECG signal is generated by using Physionets ECGSYN algorithm. ECGSYN generates synthetic ECG signals by modeling the electrical activity of the heart as a sequence of quasi-periodic waveforms. These waveforms are constructed from a set of characteristic points corresponding to the P, Q, R, S, and T waves in a typical ECG cycle. The tool uses a combination of differential equations and Gaussian functions to produce these waveforms. The derivation of each ECG can be characterized by the following equation:

(1)

Where is the amplitude of the i-th wave (P, Q, R, S, T), is the time position of the i-th wave, and is the width (standard deviation) of the i-th wave.

A clean ECG signal with 256 beats is generated using this method. As stated above, HR and Waveform morphology are the only parameters that change in this experiment. A validation step is performed after each clean signal is generated to ensure the signal is realistic. The overall algorithm used to generate the clean signal database can be seen below:

|  |  |  |  |
| --- | --- | --- | --- |
| **Algorithm 1: Algorithm to construct clean ECG signals.** | | | |
|  | ***Input***:  HR\_TO\_GENERATE  MIN\_ANGLES\_OF\_EXTREMA  MAX\_ANGLES\_OF\_EXTREMA  MIN\_Z\_POSITION\_OF\_EXTRAMA  MAX\_Z\_POSITION\_OF\_EXTRAMA  MIN\_GAUSSIAN\_WIDTH  MAX\_GAUSSIAN\_WIDTH | | |
|  | ***Output:*** A database of clean ECG signals with different morphologies and Heart Rates. | | |
| **1** | ***Sampling 🡨*** *Generate 10,000 random sample parameters of our inputs.* | | |
| **2** | ***For (iHeartRate = 1 : numberOfHeartRates)*** *// Loop through all heart rates defined.* | | |
| **3** |  | *Define the HR to generate 10,000 signals for.* | |
| **4** |  | ***for (iSignal = 1 : numberOfSignalsPerHR)*** *// Loop through each of the 10000 signals* | |
| **5** |  |  | **Define parameter settings:** *Waveform angle, waveform amplitude, width size.* |
| **6** |  |  | ***Try:***  *Call ECGSYN function with specific parameter settings.*  ***Catch:***  ***Continue***  ***end*** |
| **7** |  |  | ***Validate :*** *Validate that the waveform amplitudes and widths are within a realistic range.* |
| **8** | ***Save :*** *Save the parameter setting matrix and signals.* | | |

Furthermore, figure X shows examples of 4 clean signals generated for each HR:

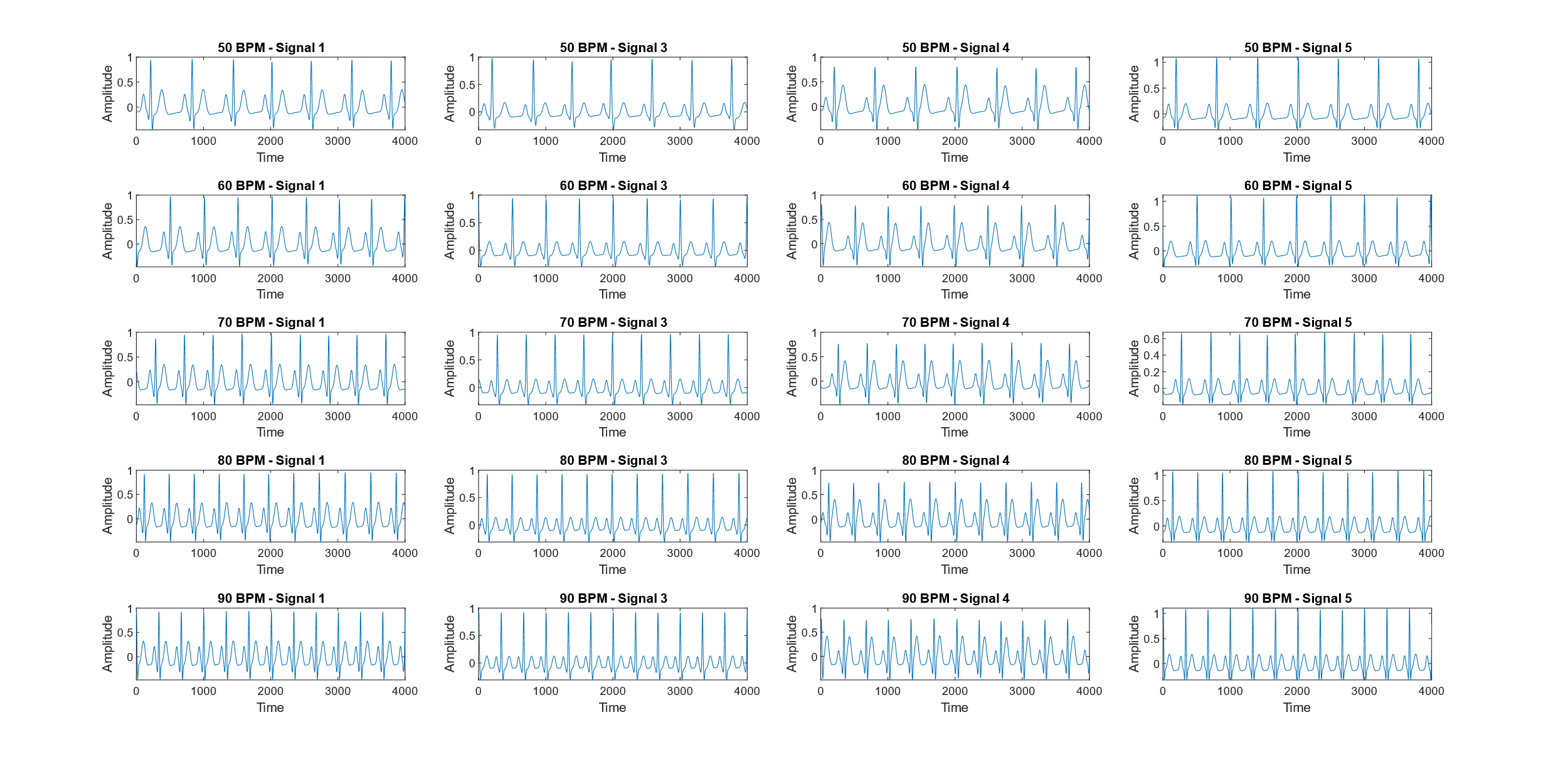


Figure 7 Plots showing 4 different clean signals for all heart rates.

It should be noted that for this experiment, 200 clean signals were generated for each HR, providing a large database containing the necessary variability.

The final step of the database generation is to corrupt each clean ECG signal with various amounts and types of noise. This stage is complex, and the overall algorithm can be seen below:

|  |  |  |  |
| --- | --- | --- | --- |
| **Algorithm 2 : Algorithm to construct Noisy ECG signals.** | | | |
|  | ***Input***: Path to noise signal  Path to clean ECG signals  Sampling frequency  ECG Length  Number of noise sections  SNR Levels  Number of generated noise signals. | | |
|  | ***Output:*** A database of noisy ECG signals with different morphologies and Heart Rates. | | |
| **1** | ***Load Noise Data 🡨*** *Read in 30 minute noise signal* | | |
| **2** | *Segment the noise signal into 30 second strips.* | | |
| **3** |  | ***for (iSegment = 1 : Number of noise sections)***  *Predict a new 30 second noise strip using auto-regressive modelling.* | |
| **4** |  |  | ***Save :*** *Save each noise signal to a table on MATLAB.* |
| **5** | **for (iCleanEcgSignal = 1 : numberOfEcgSignal)** | | |
| **6** | **Load ECG Data 🡨** Read in 30 second record. | | |
| **7** | **Load QRS Locations 🡨** Load in the QRS locations. | | |
| **8** | **Calculate** the peak to peak amplitude of QRS peaks. | | |
| **9** | **Convert** peak to peak amplitude to power | | |
| **10** | **Scale** the noise signal to each required SNR level. | | |
| **11** | **Add** the noise signal to the clean ECG record. | | |
| **12** | **Save** : Save the table which contains all SNR levels and noise corrupt  Signals for one clean ECG record. | | |
| **13** | **end** | | |
|  |  | | |

* Noise Modelling
  + Types of Motion Artifacts
  + Simulation of Noisy Signals
* Machine Learning Techniques
  + Supervised Learning
  + Unsupervised Learning
  + Deep Learning Models
* Model Training and Validation
  + Training Procedures
  + Evaluation Metrics

**Chapter 4: Implementation**

* Software and Tools Used
* Implementation Details
* Workflow
* Challenges and Solutions

**Chapter 5: Results and Discussion**

* Performance of Machine Learning Models
* Comparison with Existing Methods
* Analysis of Results
* Implications and Applications

**Chapter 6: Conclusion**

* Summary of Findings
* Contributions to the Field
* Limitations of the Study
* Future Work

**References**

**[1]** Ben Jabeur, T., Bashier, E., Sandhu, Q., Bwalya, K. J., & Joshua, A. (2024). Noise and Artifacts Elimination in ECG Signals Using Wavelet, Variational Mode Decomposition and Nonlocal Means Algorithm. arXiv. The paper discusses the elimination of various noise types in ECG signals, including Baseline Wander (BW), Powerline Interference (PL), Muscle Artifact (MA), and Electrode Motion (EM) artifacts.

[2] Sörnmo, L., & Laguna, P. (2005). **"Bioelectrical Signal Processing in Cardiac and Neurological Applications"**

**[3]** Dai, X., & Bai, Y. (2021). Denoising ECG by Adaptive Filter with Empirical Mode Decomposition. *IEEE Access*, 9, 71659-71668. doi: 10.1109/ACCESS.2021.3080795.

[4] Tam H., Webster J.G. Minimizing Electrode Motion Artifact by Skin Abrasion. *IEEE Trans. Biomed. Eng.*1977;**BME-24**:134–139. doi: 10.1109/TBME.1977.326117

[5] An X., Stylios G.K. Comparison of Motion Artefact Reduction Methods and the Implementation of Adaptive Motion Artefact Reduction in Wearable Electrocardiogram Monitoring. *Sensors.*2020;**20**:1468. doi: 10.3390/s20051468

[6] Łęski J.M., Henzel N. ECG baseline wander and powerline interference reduction using nonlinear filter bank. *Signal Process.*2005;**85**:781–793. doi: 10.1016/j.sigpro.2004.12.001

[7] Chouhan V.S., Mehta S.S. Total removal of baseline drift from ECG signal; Proceedings of the International Conference on Computing: Theory and Applications (ICCTA’07); Kolkata, India. 5–7 March 2007; pp. 512–515.

[8] El-Dahshan E.-S.A. Genetic algorithm and wavelet hybrid scheme for ECG signal denoising. *Telecommun. Syst.*2011;**46**:209–215. doi: 10.1007/s11235-010-9286-2.

[9] Karthikeyan P., Murugappan M., Yaacob S. ECG signal denoising using wavelet thresholding techniques in human stress assessment. *Int. J. Electr. Eng. Inform.*2012;**4**:306. doi: 10.15676/ijeei.2012.4.2.9.

[10] Singh B.N., Tiwari A.K. Optimal selection of wavelet basis function applied to ECG signal denoising. *Digit. Signal Process.*2006;**16**:275–287. doi: 10.1016/j.dsp.2005.12.003.

[11] Strasser F., Muma M., Zoubir A.M. Motion artifact removal in ECG signals using multi-resolution thresholding; Proceedings of the 20th European Signal Processing Conference (EUSIPCO); Bucharest, Romania. 27–31 August 2012; pp. 899–903.

[12] Zhang D. Wavelet approach for ECG baseline wander correction and noise reduction; Proceedings of the 27th Annual Conference on Engineering in Medicine and Biology; Shanghai, China. 17–18 January 2006; pp. 1212–1215.

[13] Zhang D. Wavelet approach for ECG baseline wander correction and noise reduction; Proceedings of the 27th Annual Conference on Engineering in Medicine and Biology; Shanghai, China. 17–18 January 2006; pp. 1212–1215.

[14] Blanco-Velasco M., Weng B., Barner K.E. ECG signal denoising and baseline wander correction based on the empirical mode decomposition. *Comput. Biol. Med.*2008;**38**:1–13. doi: 10.1016/j.compbiomed.2007.06.003.

[15] Pan N., Mang V., Un M.P., Hang P.S. Accurate removal of baseline wander in ECG using empirical mode decomposition; Proceedings of the 2007 Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging; Hangzhou, China. 12–14 October 2007; pp. 177–180.

[16] Zhao Z.-D., Chen Y.-Q. A new method for removal of baseline wander and power line interference in ECG signals; Proceedings of the Machine Learning and Cybernetics; Dalian, China. 13–16 August 2006; pp. 4342–4347.

[17] Kabir M.A., Shahnaz C. Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains. *Biomed. Signal Process. Control.*2012;**7**:481–489. doi: 10.1016/j.bspc.2011.11.003.

[18] Thakor N.V., Zhu Y.S. Applications of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection. *IEEE Trans. Biomed. Eng.*1991;**38**:785–794. doi: 10.1109/10.83591.

[19]  Zhang Z., Silva I., Wu D., Zheng J., Wu H., Wang W. Adaptive motion artefact reduction in respiration and ECG signals for wearable healthcare monitoring systems. *Med. Biol. Eng. Comput.*2014;**52**:1019–1030. doi: 10.1007/s11517-014-1201-7.

[20] Iyer V., Ploysongsang Y., Ramamoorthy P. Adaptive filtering in biological signal processing. *Crit. Rev. Biomed. Eng.*1990;**17**:531–584.

[21] Correa A.G., Laciar E., Patino H.D., Valentinuzzi M.E. Artifact removal from EEG signals using adaptive filters in cascade. *J. Phys. Conf. Ser.*2007;**90**:012081. doi: 10.1088/1742-6596/90/1/012081.

[22]  Romero I., Geng D., Berset T. Adaptive filtering in ECG denoising: A comparative study; Proceedings of the 2012 Computing in Cardiology; Krakow, Poland. 9–12 September 2012; pp. 45–48.

[23]  Liu S.-H. Motion artifact reduction in electrocardiogram using adaptive filter. *J. Med. Biol. Eng.*2011;**31**:67–72. doi: 10.5405/jmbe.676.

A**ppendices**

* Additional Data
* Code Listings
* Supplementary Material