Reproduction and Enhancement of the Pan-Tompkins QRS Detection Algorithm

Lara Foqaha 1220071, Taima Nasser 1222640, Veronica Wakileh 1220245

Birzeit University
Faculty of Engineering and Technology
Electrical and Computer Engineering Department

Abstract
This project aims to reproduce the classic Pan-Tompkins QRS detection algorithm using real ECG data and analyze each of its signal processing stages through digital signal processing (DSP) tools such as frequency response, pole-zero plots, and group delay. The goal is also to implement a modern adaptive thresholding strategy using least-mean-square (LMS) filtering to enhance QRS detection robustness, and evaluate and compare both the original and enhanced approaches using a real ECG dataset from the MIT-BIH Arrhythmia Database.

I. Introduction

The accurate detection of QRS complexes in electrocardiogram (ECG) signals is a fundamental task in biomedical signal processing, serving as the basis for numerous diagnostic and monitoring applications. Among the many algorithms developed for this purpose, the Pan-Tompkins algorithm remains one of the most influential due to its real-time performance, reliability, and structured signal processing pipeline. This algorithm leverages a sequence of digital processing steps (including bandpass filtering, differentiation, squaring, and moving window integration) followed by thresholding techniques to detect QRS complexes. Despite its effectiveness, challenges remain in handling noisy signals and varying QRS morphologies. Therefore, this project revisits the Pan-Tompkins algorithm with the goal of reproducing its original design and exploring enhancements through adaptive techniques. In particular, adaptive thresholding using least-mean-square (LMS) filtering is introduced to improve detection robustness under diverse signal conditions. The study also provides a detailed analysis of each processing stage and compares the performance of the original and enhanced algorithms using real ECG data from the MIT-BIH Arrhythmia Database.

II. Literature Review & Signal Flow

The Pan-Tompkins QRS detection algorithm is a widely used method in ECG signal processing, known for its reliable real-time performance in detecting QRS complexes, even in noisy conditions. It transforms raw ECG signals into a simplified form where QRS peaks are prominent, using a structured sequence of digital signal processing stages.

A. Summary of the Pan-Tompkins Algorithm Stages

1. **Bandpass Filtering**: Suppresses baseline wander and high-frequency noise, retaining the 5–15 Hz range typical of QRS complexes.
2. **Differentiation**: Highlights steep slopes, enhancing QRS features while reducing the influence of slower waves.
3. **Squaring**: Makes all values positive and emphasizes large slope changes, boosting QRS prominence.
4. **Moving Window Integration**: Smooths the signal using a short window (≈150 ms) to produce broad humps that represent QRS energy and width.
5. **Thresholding and Decision Logic**: Applies adaptive thresholds on both the integrated and bandpass-filtered signals, using signal/noise peak tracking, dual validation, and RR interval monitoring to determine QRS locations.

B. Signal Flow Diagram

The signal flow of the Pan-Tompkins algorithm:

Raw ECG → Bandpass Filter → Derivative → Squaring → Moving Window Integration → Adaptive Thresholding & Decision Logic → QRS Detection

This processing chain ensures effective enhancement and isolation of QRS complexes, enabling accurate detection in real time.

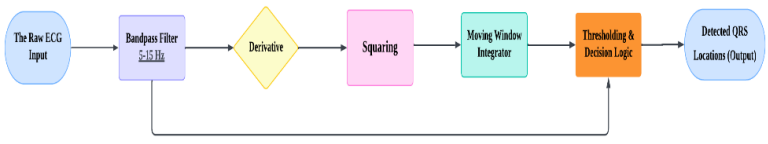

Figure 1: Signal flow diagram of the Pan-Tompkins QRS detection algorithm.

III. REPRODUCTION OF THE ALGORITHM

This section describes the reproduction of the original Pan-Tompkins algorithm using Python and ECG data from the MIT-BIH Arrhythmia Database, specifically Record 100. Each signal processing stage is implemented based on the 1985 paper, with visual inspection and signal analysis confirming the correctness of each step.

A. Loading Real ECG Data

The raw ECG signal was obtained from MIT-BIH Record 100 using the WFDB Python package. The signal is sampled at 360 Hz, and the first 5000 samples from Lead MLII were used for initial analysis.

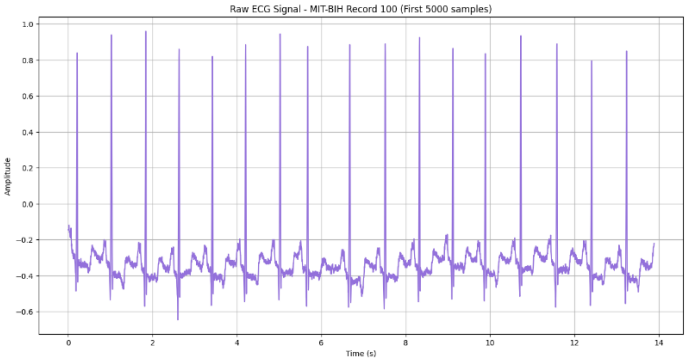

Figure 2: Figure 2: Raw ECG Signal from MIT-BIH Record 100

B. Bandpass Filtering

The bandpass filtering stage is crucial for improving the signal-to-noise ratio of the ECG by passing frequencies primarily in the 5–15 Hz range, which encompasses the spectral energy of typical QRS complexes. It is implemented by cascading a low-pass and a high-pass filter, both defined by specific difference equations from the original Pan-Tompkins paper. The low-pass filter removes high-frequency noise such as muscle artifacts and powerline interference, while the high-pass filter removes low-frequency drift and baseline wander. These filters were applied using Python's *scipy.signal.lfilter* function, resulting in a stable baseline, reduced noise, and preservation of sharp QRS peaks.

Visual inspection of the bandpass-filtered signal in Figure 3 reveals several expected effects:

- **Baseline Stability**: The slow baseline drifts present in the raw ECG have been reduced, centering the signal around zero.
- **Noise Reduction**: The output appears smoother due to attenuation of high-frequency components.
- **QRS Preservation**: QRS complexes remain sharp and prominent, confirming that the desired frequency content was preserved.

Notably, the amplitude of the signal is significantly increased compared to the raw ECG. This is expected due to the gain characteristics of the filter described in the original paper. Despite this amplitude scaling, the shape and timing of the QRS complexes are well-preserved, making the signal suitable for subsequent processing stages.

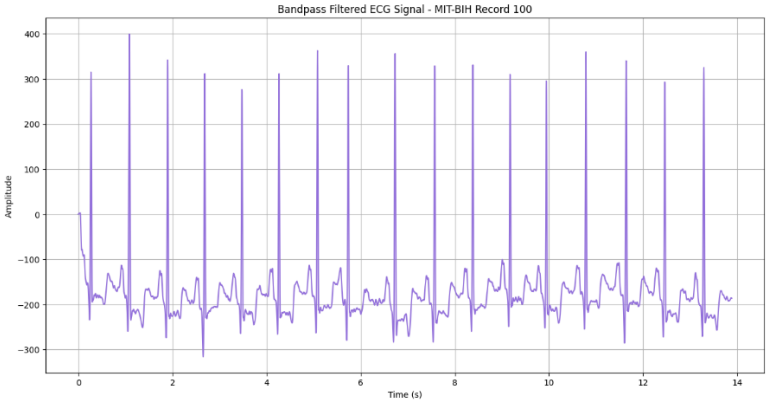

Figure 3: Bandpass Filtered ECG Signal - MIT-BIH Record 100

C. Derivative Filter

Following the bandpass stage, a five-point derivative filter is applied to extract the slope characteristics of the ECG signal. The QRS complex, known for its steep rising and falling edges, is further highlighted by this operation. The filter uses the coefficients [1, 2, 0, -2, -1] and is scaled by the factor Fs/8, where Fs is the sampling frequency (360 Hz).

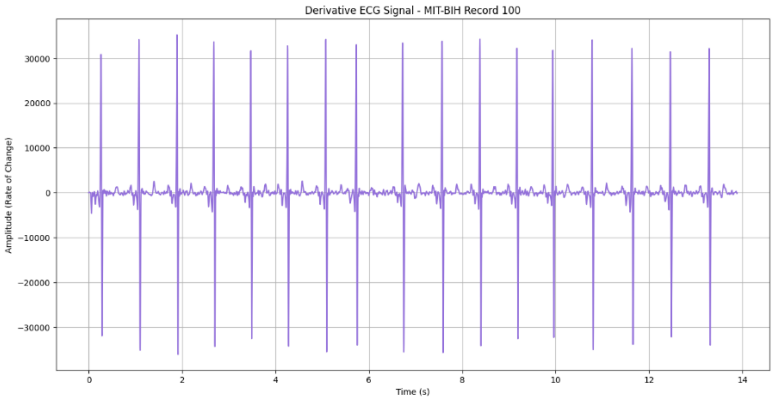

Figure 4: Derivative ECG Signal - MIT-BIH Record 100

Visual inspection of the derivative signal confirms three key effects:

- **Slope Enhancement**: Large positive and negative spikes appear where the ECG exhibits steep transitions, directly corresponding to the QRS complex.
- **Attenuation of Slower Features**: Components like P and T waves, which change more gradually, are reduced to near-zero levels.
- **Amplitude Scale**: The spikes often exceed tens of thousands in amplitude due to both differentiation and the gain from prior filtering.

These effects match the expected behavior described in the Pan-Tompkins paper. The visual output also aligns well with

Figure 1(c) from the original work. Additional DSP properties of this filter will be covered in Section IV.

D. Squaring

The output of the derivative stage is then squared point-by-point. This non-linear operation serves two purposes: it removes sign ambiguity by making all values positive, and it amplifies larger slope transitions disproportionately compared to small ones. This makes the QRS features stand out clearly while suppressing smaller noise fluctuations.

The squared signal yields sharp, high peaks corresponding to QRS complexes, which are crucial for the next stage that integrates both energy and duration information.

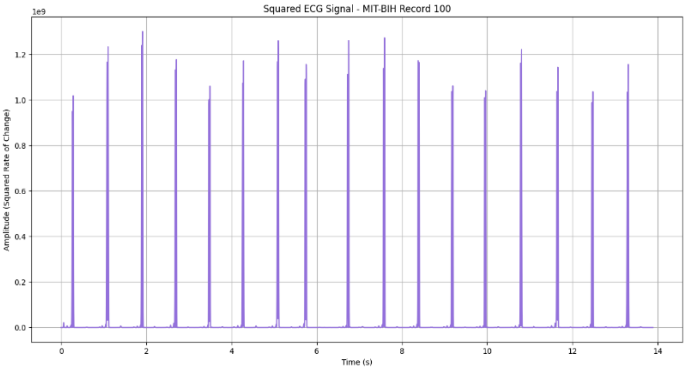

Figure 5: Squared ECG Signal- MIT-BIH Record 100

Figure 5 clearly shows that all values are non-negative. The peaks corresponding to the original QRS complexes are now large, positive, and much more separated from the surrounding signal, which has been significantly reduced in amplitude. This output emphasizes the "energy" associated with the rapid changes of the QRS.

E. Moving Window Integration

The final filtering stage in the preprocessing pipeline is the Moving Window Integrator. It computes the average of the squared signal over a short window that slides sample-by-sample. Implemented as a FIR filter with uniform coefficients (1/N, where N is the window size), this operation smooths sharp spikes into broader humps.

The integration serves two main purposes:

1. **Smoothing**: Converts sharp peaks into rounded features, making peak detection more robust.
2. **Width Information**: Encodes the energy and duration of the QRS complex, complementing the slope info from the derivative stage.

The Pan-Tompkins paper recommends a window of 150 ms. At 360 Hz, this corresponds to 54 samples. The integration filter was implemented using *scipy.signal.lfilter*, and applied to the squared signal.

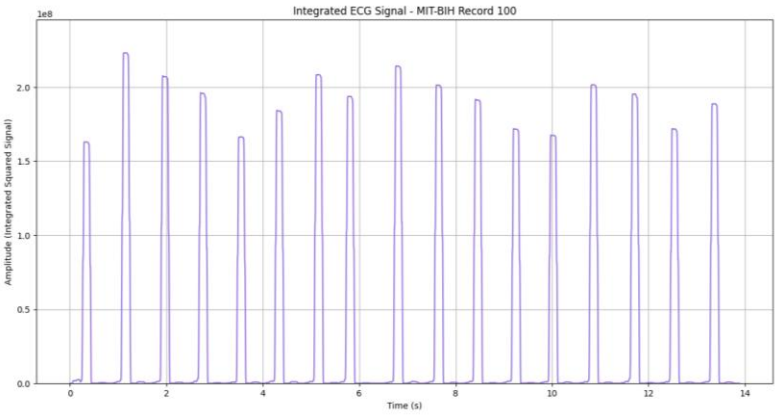

Figure 6: Integrated ECG Signal - MIT-BIH Record 100

The output signal consists of smooth, positive humps aligned with QRS complexes. This makes them easier to detect based on amplitude and duration. DSP analysis of this filter (discussed in Section IV) shows it behaves like a low-pass moving average filter with a linear phase response and consistent group delay (properties that preserve the temporal shape of the QRS while reducing high-frequency noise).

F. Thresholding and Decision Logic (Reproduction)

The final detection stage employs adaptive thresholding logic on both the integrated and filtered signals. The algorithm maintains running estimates of signal and noise peaks: SPKI/NPKI for the integrated signal and SPKF/NPKF for the bandpass-filtered signal. Thresholds are calculated as weighted averages of these peak estimates, adapting to changing ECG dynamics.

The detection procedure follows a sequence of steps:

1. **Peak Identification**: Local maxima are first identified in the integrated signal. These serve as QRS candidates.
2. **Initialization**: Signal and noise peak estimates are initialized from the first few seconds of data. The highest initial peak is assumed to be a signal (SPK), and the lowest as noise (NPK).
3. **Refractory Check**: Each peak is compared against the previous QRS detection time to enforce a 200 ms refractory period.
4. **Threshold Application**: If outside the refractory window, the peak is checked against both THRESHOLD I1 (integrated) and THRESHOLD F1 (filtered). Dual-signal validation confirms a QRS if both thresholds are exceeded. A fallback check is also done using lower thresholds (I2 and F2).
5. **State Update**: Based on classification, peak values are used to update SPK/NPK and recalculate new thresholds accordingly.

Although simplified, this logic captures the essential behavior of the original paper's adaptive mechanism. Detailed RR interval-based prediction and search-back logic are omitted in this reproduction.

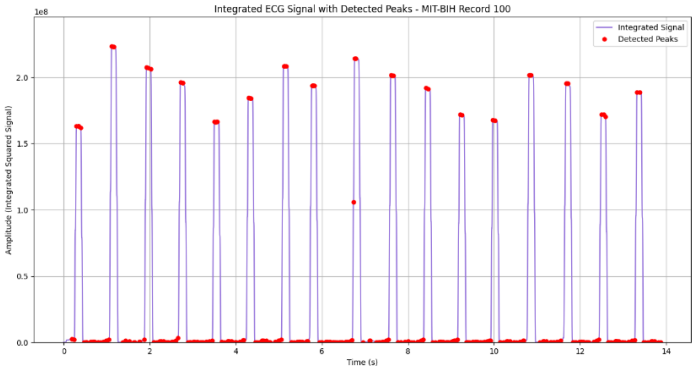

Figure 7: Integrated Signal with Detected Peaks

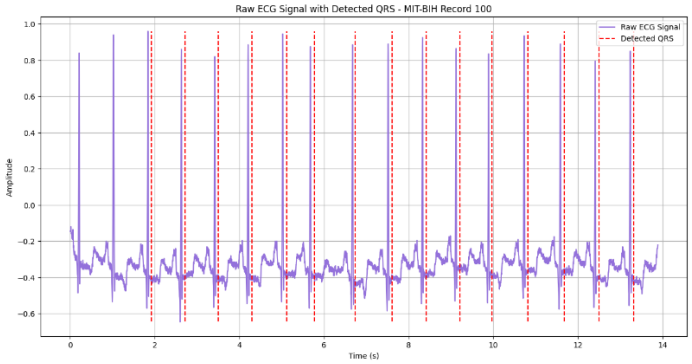

Figure 8: Raw ECG Signal with Detected QRS MIT-BIH Record 100

Figure 7 shows peak detections on the integrated signal and Figure 8 presents the final QRS detection result overlaid on the raw ECG waveform. The results confirm that QRS complexes were reliably identified at the correct positions, validating the performance of the reproduced detection logic.

IV. DSP ANALYSIS OF FILTERS

To validate the behavior of the filtering stages used in the Pan-Tompkins algorithm, this section presents a digital signal processing (DSP) analysis for each of the linear filters: Bandpass, Derivative, and Moving Window Integrator. For each filter, the magnitude and phase response, group delay, and pole-zero characteristics are computed and interpreted. These analyses provide insight into how each filter shapes the ECG signal to emphasize QRS complexes while minimizing noise and distortion.

A. Bandpass Filter Analysis

The bandpass filter is implemented by cascading a second-order low-pass and a high-pass filter. Its goal is to preserve the energy in the QRS frequency band (5–15 Hz) while attenuating baseline wander and high-frequency noise.

- **Magnitude Response**: The filter shows gain within the desired passband, with deep notches at 60 Hz and 120 Hz, effectively rejecting powerline noise.

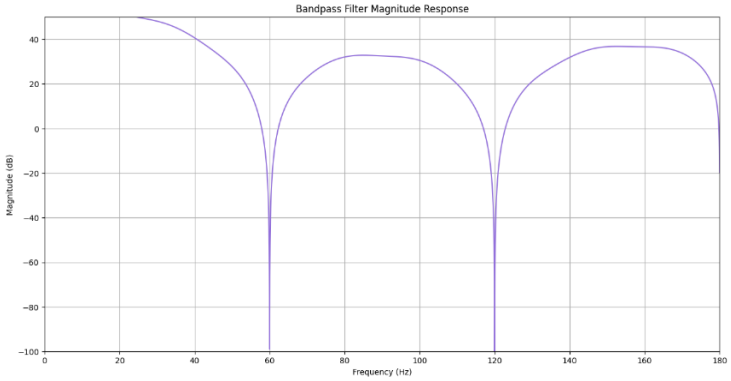

Figure 9: Magnitude Response of Bandpass Filter

The plot exhibits a typical bandpass shape with regions of high gain and deep attenuation. It shows considerable gain in the passband, with prominent peaks around +35 dB and +38 dB, corresponding to enhanced QRS energy. These peaks align with the amplitude boost observed in the filtered signal. The magnitude response also displays sharp notches at 60 Hz and 120 Hz, effectively suppressing powerline interference and its harmonic. Notably, due to the actual 360 Hz sampling rate (higher than the original design assumption of 200 Hz), the passband appears broader and more complex than initially specified, which is acceptable as it still targets the QRS spectrum while filtering out undesired components.

- **Phase Response**: The phase response ideally should be linear to ensure consistent delay across all frequencies and preserve waveform shape. While the response of this IIR filter is not perfectly linear, it remains relatively flat in the QRS-relevant frequency bands (roughly 0–50 Hz and above 70 Hz). This indicates limited phase distortion, which helps maintain the integrity of the signal's shape.

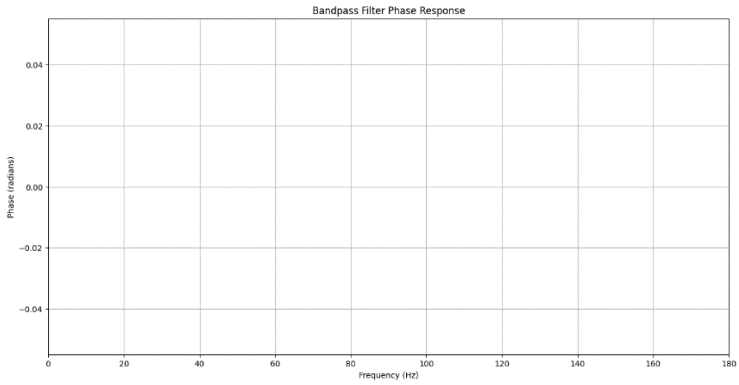

Figure 10: Phase Response of Bandpass Filter

The plot shows the phase is generally flat in the main frequency range of interest, confirming that although not strictly linear, the filter's distortion is minimal and acceptable for QRS detection purposes.

- **Group Delay**: Group delay represents the average delay experienced by each frequency component as it passes through the filter. A constant group delay across the passband ensures signal features are uniformly delayed, preserving waveform shape.

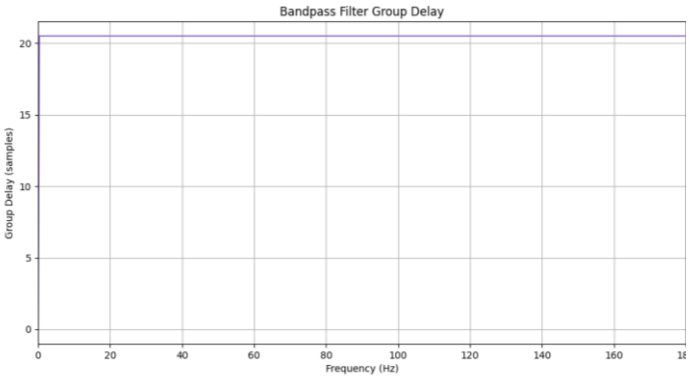

Figure 11: Group Delay of Bandpass Filter

The group delay plot shows that the delay remains fairly constant across the main frequency range of interest (outside the notches), especially where the filter gain is highest. The average delay is approximately 20.5 samples, which aligns with the processing latency observed in the algorithm. This consistent delay explains why the detected QRS markers appear uniformly shifted in time, and helps maintain the internal timing structure of the QRS complexes.

- **Pole-Zero Plot**: The pole-zero plot provides insight into the filter's frequency behavior and confirms that the implemented difference equations match the intended design.

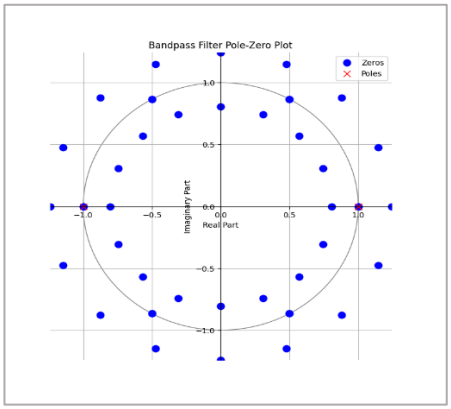

Figure 12: Pole-Zero Plot of Bandpass Filter

The plot shows poles at z = 1 and z = –1, corresponding to DC and the Nyquist frequency, respectively. These are typical of second-order IIR filter structures. The zeros are distributed precisely on the unit circle and are positioned to cancel out frequencies at 60 Hz and 120 Hz (explaining the strong attenuation observed in the magnitude response). The density and placement of the zeros create complex notch behavior around those frequencies and validate that the filtering characteristics match what was described in the original Pan-Tompkins equations. This confirms both stability and effectiveness in rejecting powerline interference.

B. Derivative Filter Analysis

The Derivative filter is the second linear processing stage. Its purpose is to calculate the instantaneous slope of the signal, emphasizing the rapid changes characteristic of the QRS complex. This implementation is based on equation (9) from the Pan-Tompkins paper, implemented causally with a 2 sample delay, and scaled by a factor of Fs/8.

The filter's coefficients b_deriv = [1, 2, 0, -2, -1] and a_deriv = [1] correspond to the difference equation:
$y[n] = (Fs/8) * (x[n] + 2x[n-1] - 2x[n-3] - x[n-4])$.

The derivative filter is a 5-tap FIR filter used to highlight rapid slope changes in the ECG.

- **Magnitude Response**:

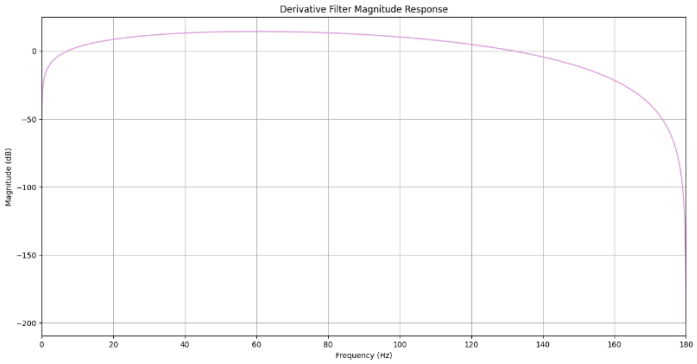

Figure 13: Magnitude Response of Derivative Filter

The magnitude response of an ideal derivative increases linearly with frequency. This filter approximates that behavior, starting from 0 Hz and rising through mid-band before rolling off at higher frequencies. This shape enhances QRS slopes while suppressing baseline drift and high-frequency noise, aiding in robust detection of rapid signal transitions.

- **Phase Response**:

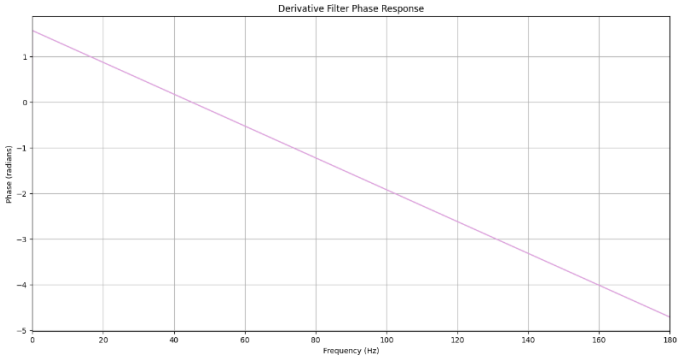

Figure 14: Phase Response of Derivative Filter

The phase response shows a nearly linear decrease with frequency, indicating the filter applies a consistent delay across the passband. This helps preserve waveform shape and the timing of the QRS features.

- **Group Delay**:

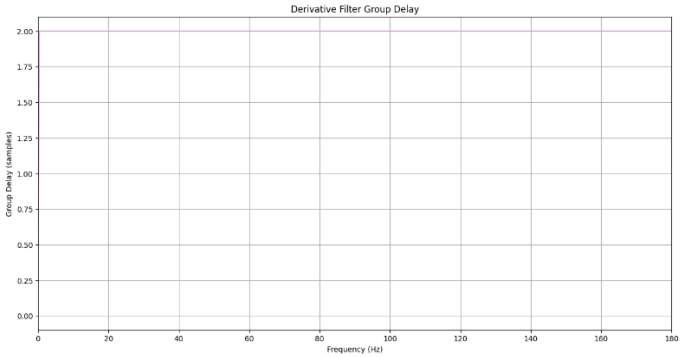

Figure 15: Group Delay of Derivative Filter

The group delay remains constant across all frequencies at exactly 2 samples, consistent with a 5-tap linear-phase FIR filter. This matches the theoretical delay of (N−1)/2 = 2 for N = 5 taps. Such uniform delay avoids phase distortion and helps preserve the shape of the ECG waveform.

- **Pole-Zero Plot**:

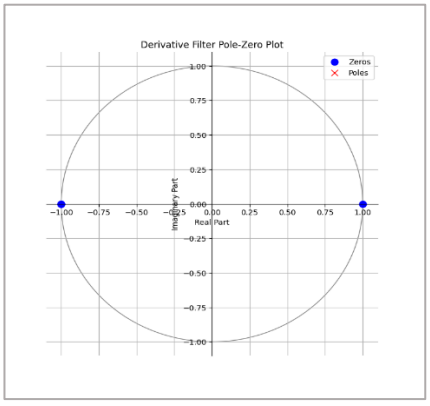

Figure 16: Pole-Zero Plot of Derivative Filter

The derivative filter, being an FIR filter, has no poles. Its zeros are positioned on the unit circle, with specific locations at z = 1 (DC) and z = –1 (Nyquist), resulting in zero gain at 0 Hz and 180 Hz. These attenuations are reflected in the magnitude response. The zero placement shapes the filter's frequency response and enhances detection of high-slope QRS transitions while suppressing baseline and high-frequency content.

C. Moving Window Integration Filter Analysis

The Moving Window Integration (MWI) filter is a linear-stage FIR filter that follows the squaring operation. It smooths sharp peaks and encodes QRS energy and width to facilitate detection. The filter is a rectangular moving average defined by the difference equation:

$y[n] = (1/N) * \Sigma_{k=0}^{n-1} x[n-k]$

For a window duration of 150 ms at 360 Hz, the window size N is set to 54 samples. The coefficients are b_integ = [1/N, 1/N, ..., 1/N] and a_integ = [1].

- **Magnitude Response**:

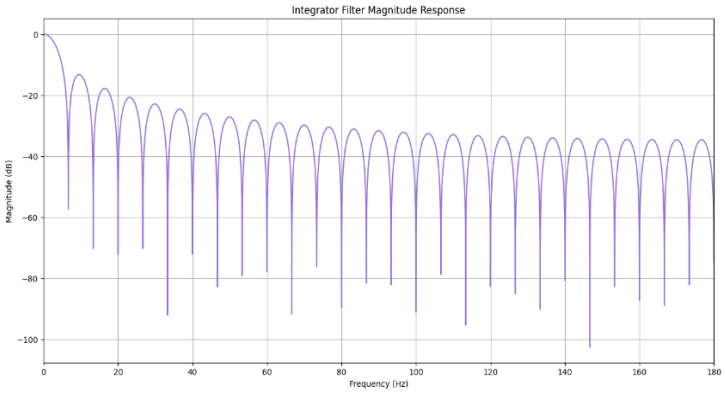

Figure 17: Magnitude Response of Integrator Filter

The response shows maximum gain at 0 Hz (DC) and decreases with frequency, displaying regular notches at intervals of Fs/N. For Fs = 360 Hz and N = 54, notches occur around 6.67 Hz, 13.33 Hz, 20 Hz, etc. These notches effectively attenuate periodic noise with durations equal to or submultiples of the window length, reinforcing the smoothing role of this moving average filter.

- **Phase Response**:

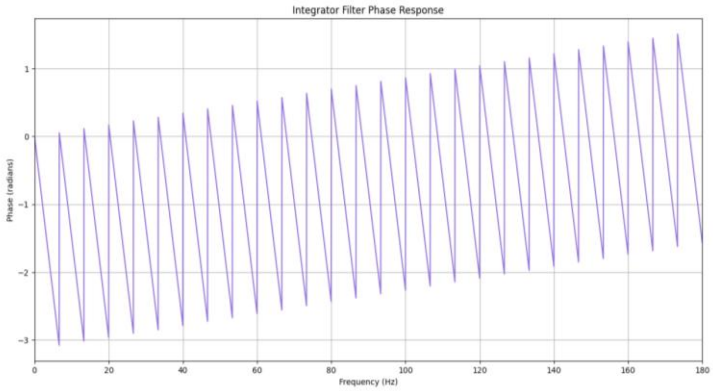

Figure 18: Phase Response of Integrator Filter

The phase response displays a linear pattern with sharp π-radian jumps, typical of linear-phase FIR filters when unwrapped. This reflects consistent delay across frequency segments and confirms predictable timing behavior.

- **Group Delay**:

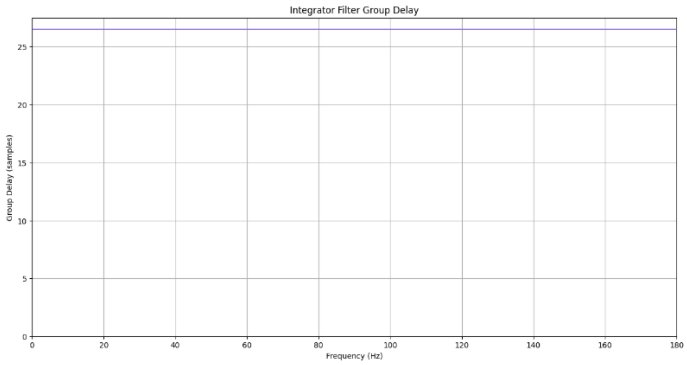

Figure 19: Group Delay of Integrator Filter

The group delay is constant across the frequency range and equals (N–1)/2 = 26.5 samples for N = 54. This consistent delay avoids phase distortion and preserves the shape and alignment of the integrated signal peaks.

- **Pole-Zero Plot**:

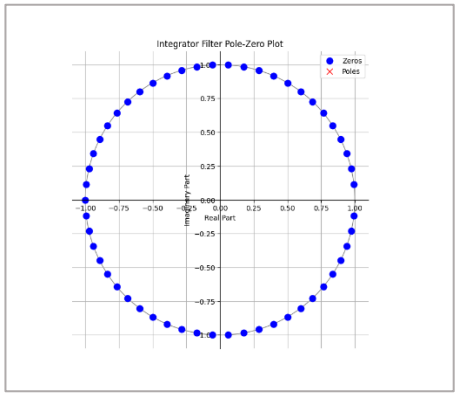

Figure 20: Pole-Zero Plot of Integrator Filter

As an FIR filter, all poles are at the origin (z = 0). The 53 zeros are equally spaced on the unit circle and correspond exactly to the notch frequencies seen in the magnitude response (multiples of Fs/N). This configuration ensures selective attenuation at predictable intervals, reinforcing the filter's smoothing behavior and signal shaping role for reliable QRS detection.

V. ADAPTIVE THRESHOLDING USING LMS

A. Objective

Improving the robustness and dependability of the QRS detection thresholding, particularly in noisy environments, is the main goal of this task. An adaptive thresholding strategy based on the concepts of Least Mean Squares LMS filtering was used to accomplish this improvement.

B. Methodology

The LMS update mechanism served as the foundation for the development of an adaptive threshold known as T_{LMS}. This method uses the amplitude of each integrated signal peak from the Pan-Tompkins preprocessing stages to dynamically update the detection threshold.

A modified LMS update equation is used by the adaptation rule:

$T_{LMS}(i+1) = T_{LMS}(i) + \mu . (P_i - T_{LMS}(i))$

Here, μ is a learning rate parameter that has been empirically set to 0.01 for stable adaptation, and P_i is the amplitude of the current peak in the integrated signal.

The threshold is continuously raised for higher peaks and lowered for lower peaks as the rule moves closer to the current peak amplitude.

C. The adaptive thresholding logic for QRS complex detection modality

If a peak's amplitude exceeds the adaptive threshold T_{LMS}, it is first identified as a QRS candidate. Each candidate peak must also pass a secondary validation check using the absolute amplitude of the corresponding bandpass-filtered signal in order to guarantee robustness comparable to the original Pan-Tompkins logic. In particular, this validation step makes use of the previously defined static threshold from the Pan-Tompkins initialization phase. To avoid multiple detections of the same QRS event, a 200 ms refractory period is enforced after detection. The QRS complex indices found using this adaptive LMS technique are saved and then assessed.

D. Basis and Nature of Adaptation

Although the implemented method adapts the detection threshold directly using the instantaneous difference between the peak amplitude and the current threshold, it conceptually borrows from the traditional LMS adaptive filtering, where weights are updated based on minimizing an error signal. This adaptive approach uses LMS-like concepts especially designed for threshold adaptation, as opposed to conventional LMS filtering, which usually modifies filter coefficients. With this method, the threshold can change dynamically in response to changes in the properties of the ECG peaks that are encountered in real time. Because it determines the threshold's adaptive response's stability and speed, the parameter μ is essential.

VI. EVALUATION AND COMPARSION

Three implemented QRS detection methods—the standard Adaptive Pan-Tompkins algorithm, a Static Threshold method, and an LMS-based Adaptive Threshold method—are quantitatively evaluated and their performances compared in this evaluation task. Samples representing clean conditions (Records 100, 101) and noisy conditions (Records 108, 203, 228) were chosen from among the ECG records from the MIT-BIH Arrhythmia Database. Sensitivity Sens, Positive Predictive Value PPV, and the F1 Score were used as key performance metrics to compare each method's output to standard annotations (i.e. ground truth).

A. Results Discussion

Both adaptive methods (Pan-Tompkins and LMS) demonstrated high accuracy and low false-positive rates on clean ECG signals (records 100 and 101), as evidenced by their high PPV values. The Static Threshold method's sensitivity to even slight changes in the signal was demonstrated by the large number of false positives it produced.

All methods' performance declined in noisy environments (records 108, 203, and 228). The Static Threshold method's low sensitivity and high false positives demonstrated its serious flaws and made it inappropriate for noisy environments.

```
[Running] python -u "c:\Users\Reem\Documents\dsp proj\code\qrs_evaluation.py"
--- Starting Evaluation Across Records ---

Method                  TP    FP    FN    Sens    PPV     F1
-----------------------------------------------------------
100                    101     0     3   0.971  1.000  0.985
                       103   177     1   0.990  0.368  0.536
                        18     0    86   0.173  1.000  0.295
101                     92     0     6   0.939  1.000  0.968
                        97   159     1   0.990  0.379  0.548
                        96     0     2   0.980  1.000  0.990
108                     74    82    11   0.871  0.474  0.614
                         2     9    83   0.024  0.182  0.042
                        24     1    61   0.282  0.960  0.436
203                    128    17    23   0.848  0.883  0.865
                        67    68    84   0.444  0.496  0.469
                       116     5    35   0.768  0.959  0.853
228                     93     8    21   0.816  0.921  0.865
                        13    13   101   0.114  0.500  0.186
                        48     4    66   0.421  0.923  0.578

--- Evaluation complete for all records ---
Script finished.
```

Figure 21: Evaluation Across clean and Noisy Records Result

Superior F1 scores were a consistent result of the Adaptive Pan-Tompkins method's balanced performance, which included controlled false positives and reasonable sensitivity.

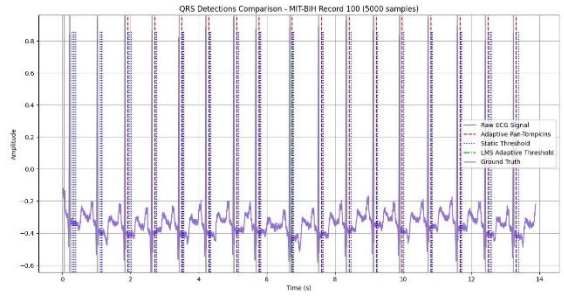

Figure 22: Clean ECG Signals - MIT-BIH Record 100

Strong specificity (high PPV) and frequently lower sensitivity were demonstrated by the LMS Adaptive Threshold method, suggesting possible problems with threshold adaptation in highly variable or noisy environments. This performance pattern points to the necessity of improved adaptation logic or parameter optimization.

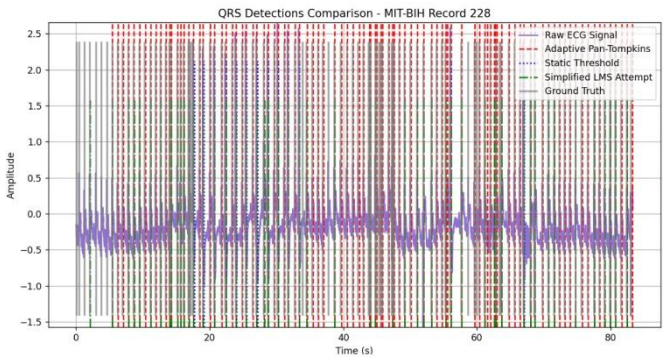

Figure 23: Final Comparison Plot for a Representative Noisy - MIT-BIH Record 228

The assessment highlights the superiority of adaptive thresholding techniques over static thresholding, with the Pan-Tompkins approach outperforming the LMS-based method in most cases, especially in terms of sensitivity across a range of ECG signal qualities. To optimize LMS performance consistency, more methodological improvements and parameter tuning are advised.

CONCLUSION

For QRS detection, this study contrasted the LMS-based Adaptive Threshold, Static Threshold, and Adaptive Pan-Tompkins approaches. Although it lacked some sophisticated features, the Adaptive Pan-Tompkins approach demonstrated great robustness and adaptability, which occasionally resulted in detecting errors. Despite its simplicity, the Static Threshold approach's sensitivity caused it to perform poorly in noisy environments. Although the LMS-based approach was simple to use and offered excellent precision, it was unstable in a variety of noise scenarios. Future developments should concentrate on incorporating search back mechanisms, T-wave discrimination for Pan-Tompkins, and comprehensive RR interval logic. They should also refine the LMS approach by improving its theoretical foundation and enhancing the learning rate parameter.