EE69205: Signal Processing System Design
Indian Institute of Technology, Kharagpur

Denoising Speech Signal

Anirvan Krishna | Roll no. 21EE38002

1. Objective

Speech signals often get corrupted by noise, which significantly degrades the quality and intelligibility of the speech. Denoising these signals is crucial in many applications such as telecommunication, hearing aids, and speech recognition systems. The Recursive Least Squares (RLS) algorithm is an adaptive filter technique widely used for denoising because of its fast convergence and low sensitivity to noise. In this experiment, we will work with a noisy speech corpus (NOIZEUS) designed to evaluate speech enhancement algorithms. This corpus contains IEEE sentences corrupted by real-world noise at various Signal-to-Noise Ratios (SNRs). We will learn to apply the RLS adaptive filter on a noisy speech corpus and evaluate the performance of the algorithm by calculating the error (Signal-to-Distortion Ratio, SDR) for different levels of SNR (0dB, 5dB, 10dB, and 15dB).

2. Observing the Speech Signals: Clean and Noisy

For the entire course of this report, we consider the speech signal to be corrupted by station noise. We'll analyze and observe the performance of different filtering techniques for speech signal with SNR: 0 dB, 5 dB, 10 dB and 15 dB.

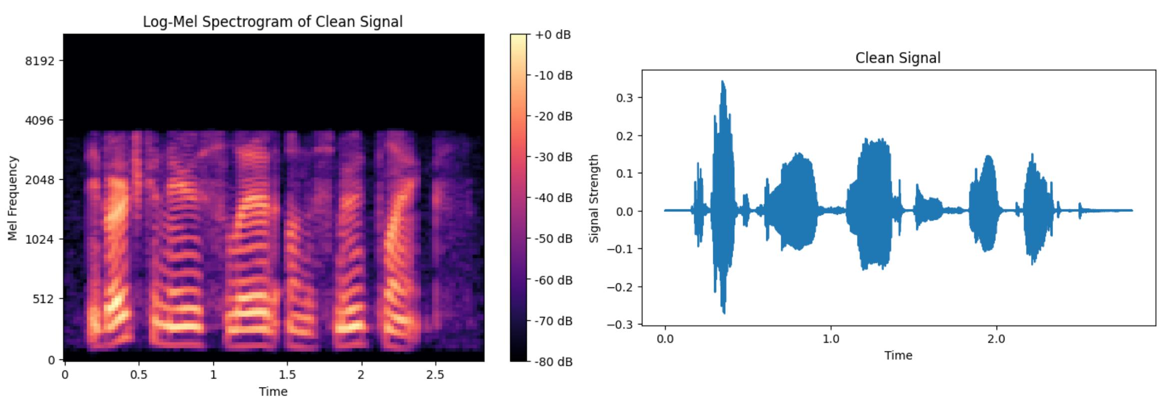

Fig. 1. Temporal and Spectral visualization of Original Speech Signal

Following this, we will look into the spectro-temporal visualizations for the noisy signals with different noise levels.

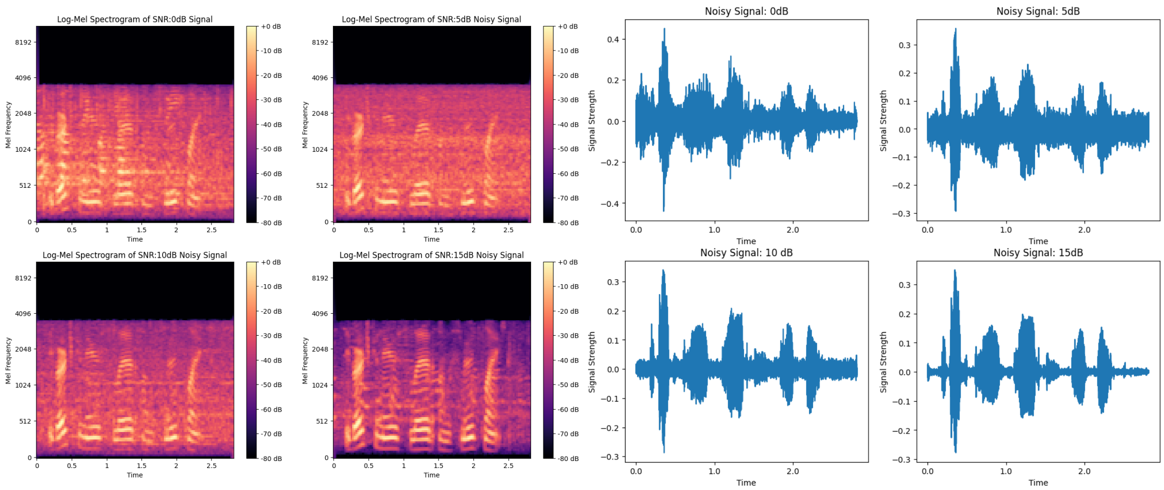

Fig. 2. Spectro-Temporal visualization of Noisy Signals for different noise levels

3. Butterworth Filtering: High Pass, Low Pass and Band Pass

In speech processing, Infinite Impulse Response (IIR) filters are widely used due to their efficiency and ability to meet specific frequency response requirements. IIR filters, unlike Finite Impulse Response (FIR) filters, have feedback components, making them recursive and typically requiring fewer coefficients for a similar frequency response. The three primary types of filters—High Pass, Low Pass, and Bandpass—serve different purposes in isolating or attenuating specific frequency components of speech signals. For these experiments, we are using a butterworth filter of order 5.

3.1. Low Pass Filtering—A Low Pass Filter (LPF) allows frequencies below a cutoff frequency ω_c to pass while attenuating higher frequencies. It is essential for removing high-frequency noise from speech signals. The transfer function of a first-order low-pass IIR filter can be expressed as:

$$|H_{LP}(j\omega)| = \frac{1}{\sqrt{1+\left(\frac{\omega}{\omega_c}\right)^{2n}}}$$

where n is filter order and ω_c is cutoff frequency.

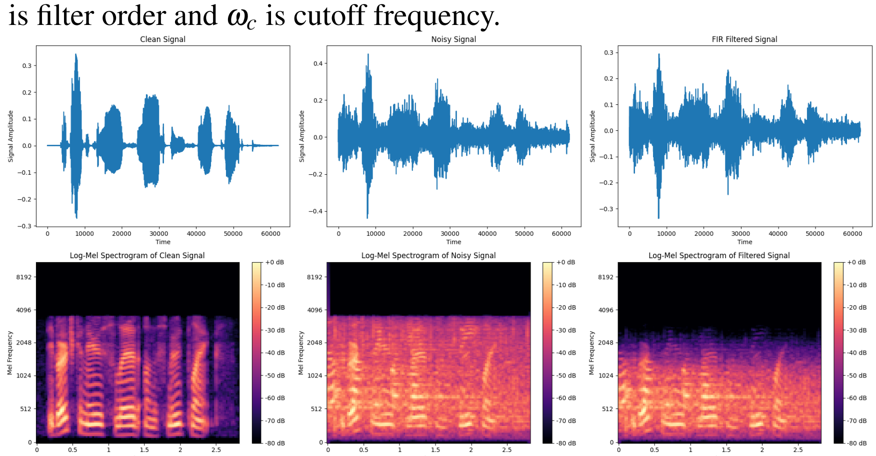

Fig. 3. Low Pass Filtering

3.2. High Pass Filtering—A High Pass Filter (HPF) allows frequencies above a cutoff frequency ω_c to pass, attenuating lower frequencies. It is used to eliminate low-frequency noise such as hum or other environmental interferences in speech. The transfer function of a first-order high-pass IIR filter can be represented as:

$$|H_{HP}(j\omega)| = \frac{1}{\sqrt{1+\left(\frac{\omega_c}{\omega}\right)^{2n}}}$$

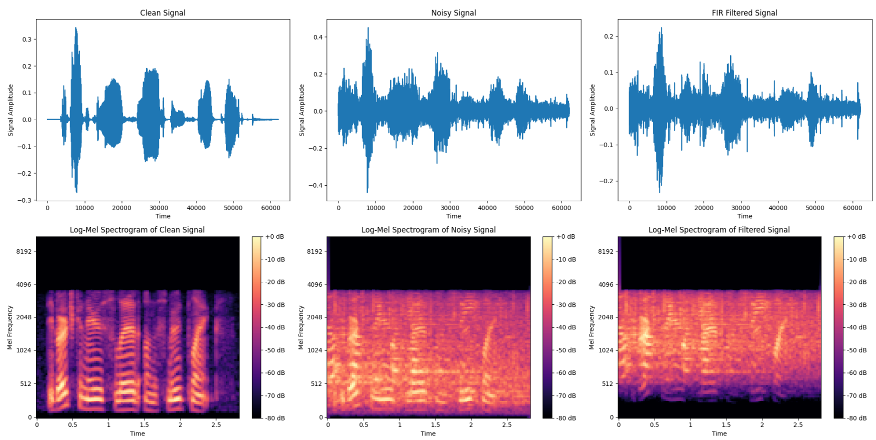

Fig. 4. High Pass Filter Implementation

3.3. Bandpass Filtering—A Bandpass Filter (BPF) allows frequencies within a certain range to pass while attenuating those outside this range. It is useful for isolating specific formant frequencies in speech, which are crucial for phonetic analysis. The transfer function of a second-order bandpass IIR filter is typically given by:

$$|H_{BP}(j\omega)| = \frac{1}{\sqrt{1+\left(\frac{\omega^2-\omega_1\omega_2}{\omega(\omega_2-\omega_1)}\right)^{2n}}}$$

where ω_1 and ω_2 are band limits of the band pass filter.

Filter Type	MSE (Noisy)	MSE (Filtered)
Low Pass	0.0013	0.0025
High Pass	0.0013	0.0020
Band Pass	0.0013	0.0015

The above table discusses the mean squared error values before and after cleaning the 0 dB SNR signal. From the spectral representation of the clean and noisy signal, it is clearly visible that the both of them lie in the same frequency range. This makes it tricky to clean the signal using conventional filtering techniques. From the above table, it is clear that using strict frequency bounds for cleaning the signal leads to information loss and causes an adverse effect. This calls for the need of advanced techniques like the one discussed further.

4. Adaptive Filtering: Recursive Least Squares

The Recursive Least Squares (RLS) algorithm is a powerful adaptive filtering technique used to minimize the error between a desired signal and the output of a linear filter by adjusting the filter

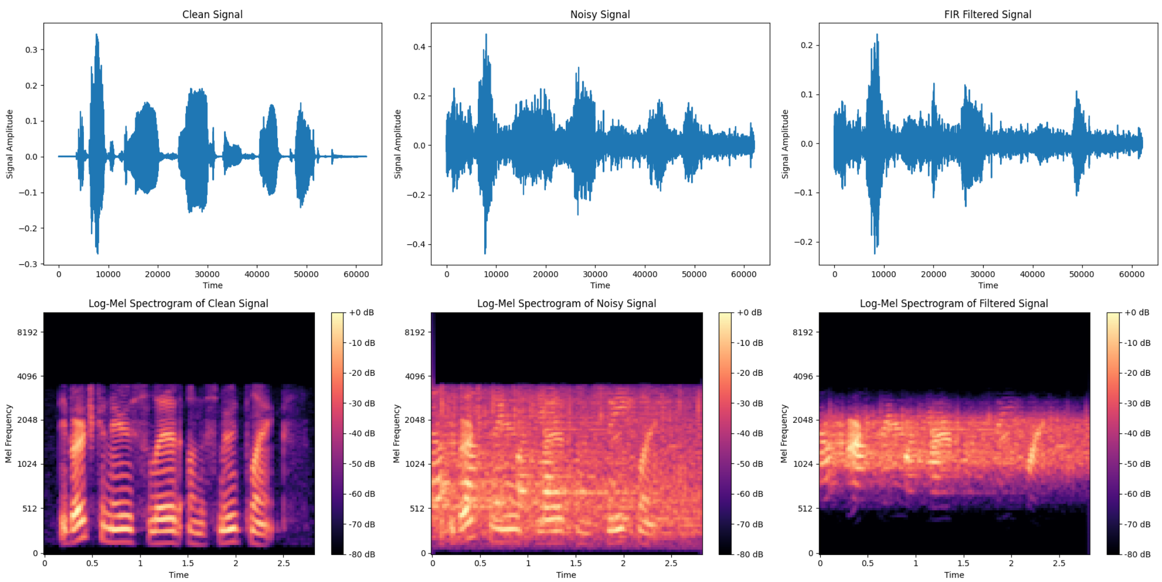

Fig. 5. Bandpass Filter Implementation

coefficients. RLS aims to minimize the weighted sum of squared errors directly, offering faster convergence at the cost of higher computational complexity. The algorithm for implementation of Recursive Least Squares has been discussed below.

Algorithm 1 Recursive Least Squares (RLS) Algorithm

Initialize filter coefficients $\mathbf{W}[0] = \mathbf{0}$
Initialize inverse correlation matrix $\mathbf{R}[0] = \mathbf{I}$ (Identity matrix)
Set forgetting factor λ
for each sample n from M to N **do**
    Extract input vector $\mathbf{X}[n] = [x[n-1], x[n-2], \ldots, x[n-M]]^T$
    Compute estimated output $\hat{d}[n] = \mathbf{W}[n-1]^T \cdot \mathbf{X}[n]$
    Compute estimation error $e[n] = d[n] - \hat{d}[n]$
    Compute Kalman gain $\mathbf{K}[n] = \frac{\mathbf{R}[n-1] \cdot \mathbf{X}[n]}{\lambda + \mathbf{X}[n]^T \cdot \mathbf{R}[n-1] \cdot \mathbf{X}[n]}$
    Update filter coefficients $\mathbf{W}[n] = \mathbf{W}[n-1] + \mathbf{K}[n] \cdot e[n]$
    Update inverse correlation matrix $\mathbf{R}[n] = \frac{1}{\lambda}\left(\mathbf{R}[n-1] - \mathbf{K}[n] \cdot \mathbf{X}[n]^T \cdot \mathbf{R}[n-1]\right)$
end for
Return denoised signal $\hat{d}[n]$

Table 1. Mean Squared Error between Original Signal and Filtered/Noisy Signal

SNR (dB)	MSE (Noisy)	MSE (Filtered)
0 dB	0.0013	0.00037
5 dB	0.00043	0.00013
10 dB	0.00013	5.2×10^{-5}
15 dB	4.3×10^{-5}	2.46×10^{-5}

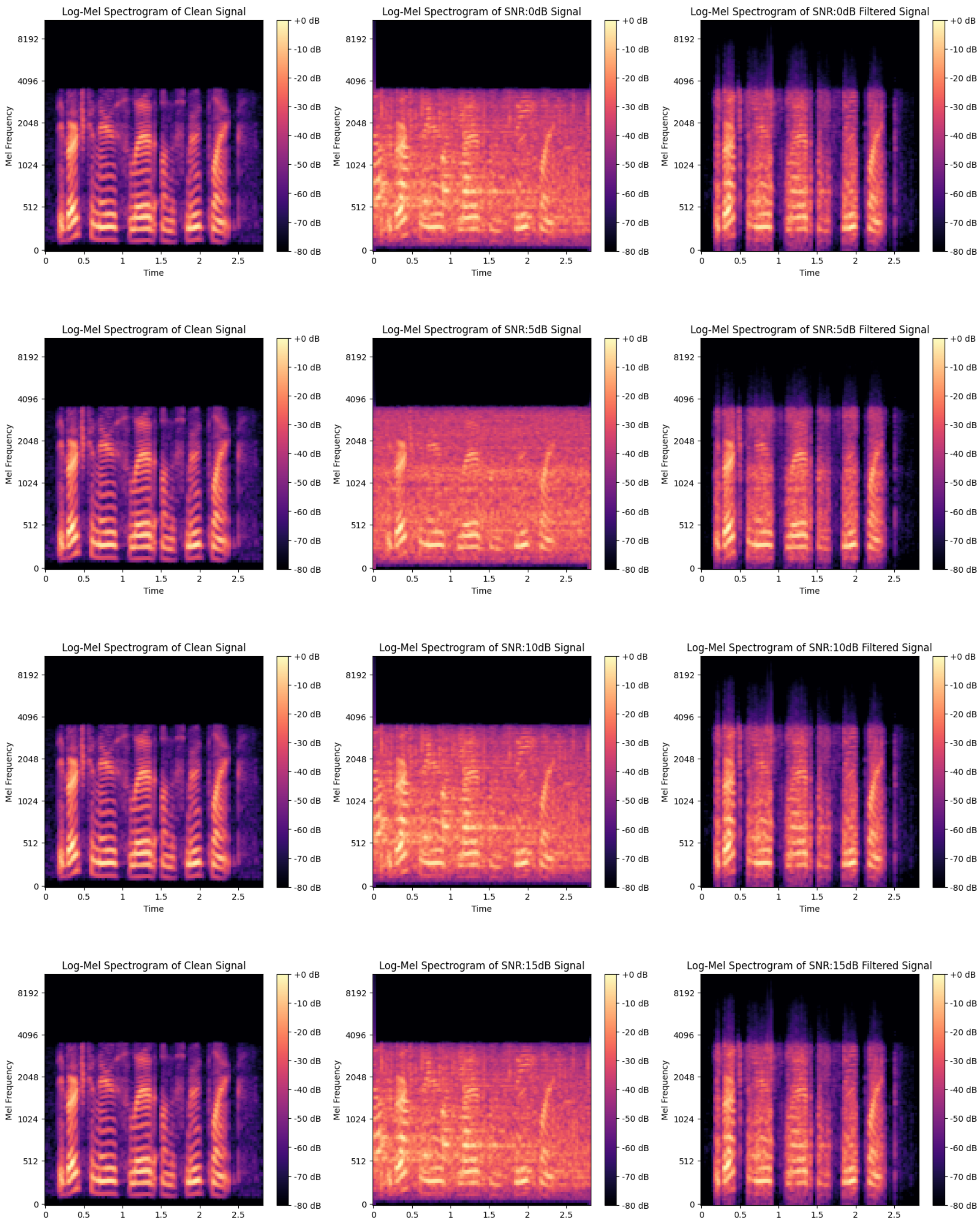

Fig. 6. Spectrographic Comparision of Original, Noisy and RLS Filtered Signals

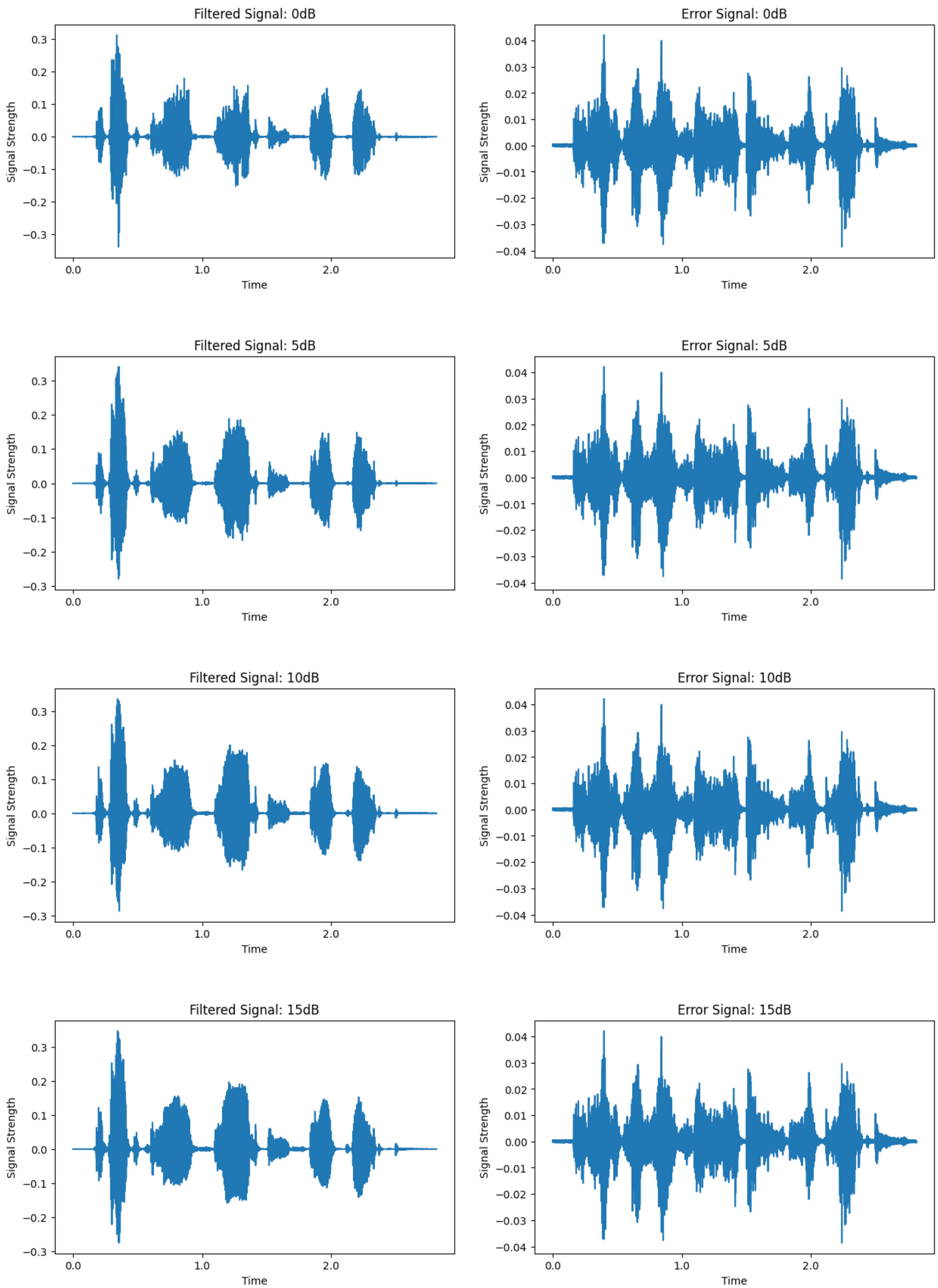

Fig. 7. Separating the Error Signal from the Filtered Speech