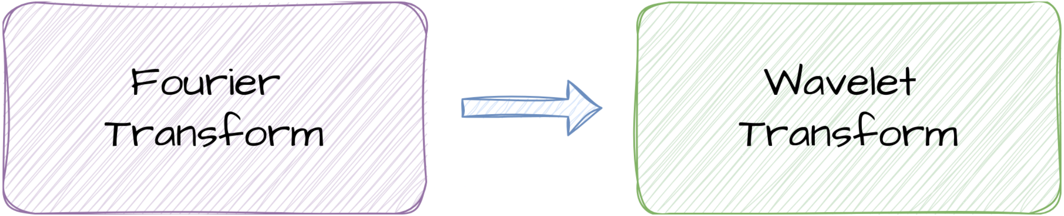
# Wavelet-Based Signal Denoising for Financial Time Series

- Siddhartha Gupte

December 5, 2024

# Problem Formulation

To evaluate wavelet-based signal denoising techniques for enhancing the analysis of financial time-series data



# Fourier Transform

## Signal in Time Domain

Continuous Signal:

$$x(t) = \text{signal in time domain}$$

Discrete Signal:

$$x[n] = \text{signal in discrete domain}$$

## Fourier Transform

Continuous Signal:

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$$

Discrete Signal:

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}$$

# Fourier Transform

## Signal in Time Domain

Continuous Signal:

$$x(t) = \text{signal in time domain}$$

Discrete Signal:

$$x[n] = \text{signal in discrete domain}$$

## Fourier Transform

Continuous Signal:

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$$

Discrete Signal:

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}$$

# Fourier Transform

## Signal in Time Domain

Continuous Signal:

$$x(t) = \text{signal in time domain}$$

Discrete Signal:

$$x[n] = \text{signal in discrete domain}$$

## Fourier Transform

Continuous Signal:

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$$

Discrete Signal:

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}$$

# Wavelet Transform

Let $x(t)$ be a signal s.t. $\int \|x(t)\|^2 dt < \infty$.

$$x(t) = \sum_{k=\infty}^{\infty} \sum_{n=\infty}^{\infty} c_{kn} \psi_{kn}$$

$c_{kn}$ - Wavelet Transform Coefficients

$\psi_{kn}$ - Wavelet Basis

$\psi_{kn}$ is derived from a single function $\psi(t)$ by the following transformations:

1. **Dilation**: $D_k(t) : \psi(t) \to \psi(f(k) \cdot t)$
2. **Translation**: $T_n : \psi(t) \to \psi(t - n)$

# Wavelet Transform

Let $x(t)$ be a signal s.t. $\int \|x(t)\|^2 dt < \infty$.

$$x(t) = \sum_{k=\infty}^{\infty} \sum_{n=\infty}^{\infty} c_{kn} \psi_{kn}$$

$c_{kn}$ - Wavelet Transform Coefficients

$\psi_{kn}$ - Wavelet Basis

$\psi_{kn}$ is derived from a single function $\psi(t)$ by the following transformations:

1. **Dilation**: $D_k(t) : \psi(t) \rightarrow \psi(f(k) \cdot t)$
2. **Translation**: $T_n : \psi(t) \rightarrow \psi(t - n)$

# Wavelet Transform

Let $x(t)$ be a signal s.t. $\int \|x(t)\|^2 dt < \infty$.

$$x(t) = \sum_{k=\infty}^{\infty} \sum_{n=\infty}^{\infty} c_{kn} \psi_{kn}$$

$c_{kn}$ - Wavelet Transform Coefficients

$\psi_{kn}$ - Wavelet Basis

$\psi_{kn}$ is derived from a single function $\psi(t)$ by the following transformations:

1. **Dilation**: $D_k(t) : \psi(t) \to \psi(f(k) \cdot t)$
2. **Translation**: $T_n : \psi(t) \to \psi(t - n)$

## General Form

Wavelet transform of a signal $x(t)$

$$X(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right)$$

## Special Case - Diadic DWT

$$\begin{aligned} c_{kn} &= \int_{-\infty}^{\infty} x(t)\psi_{kn}(t)dt \\ &= \int_{-\infty}^{\infty} x(t)\psi\left(2^k t - n\right)dt \end{aligned}$$

$a = 2^{-k}$

$b = 2^{-k}n$

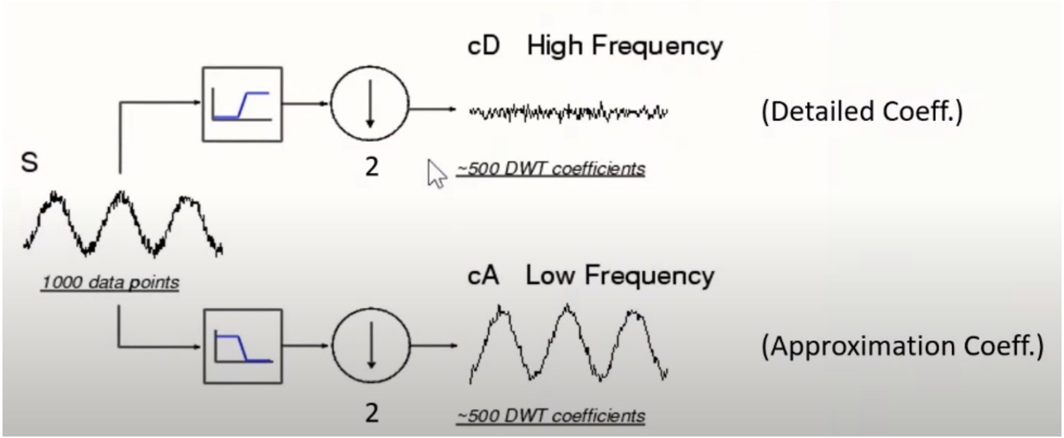



Figure: Filters

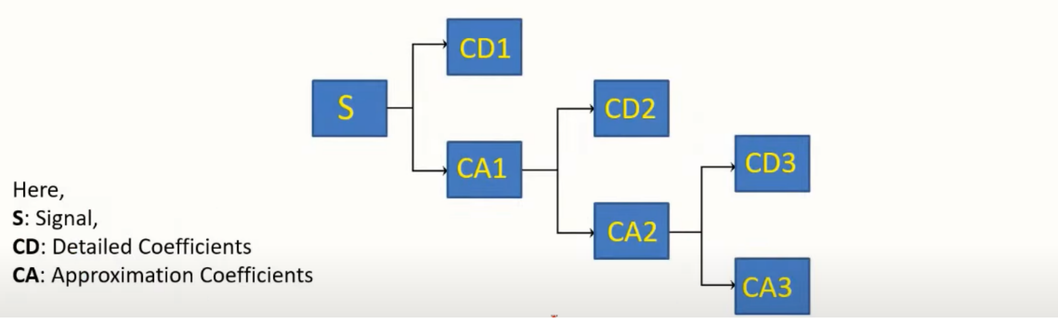



Figure: Level 3 Wavelet Decomposition Tree

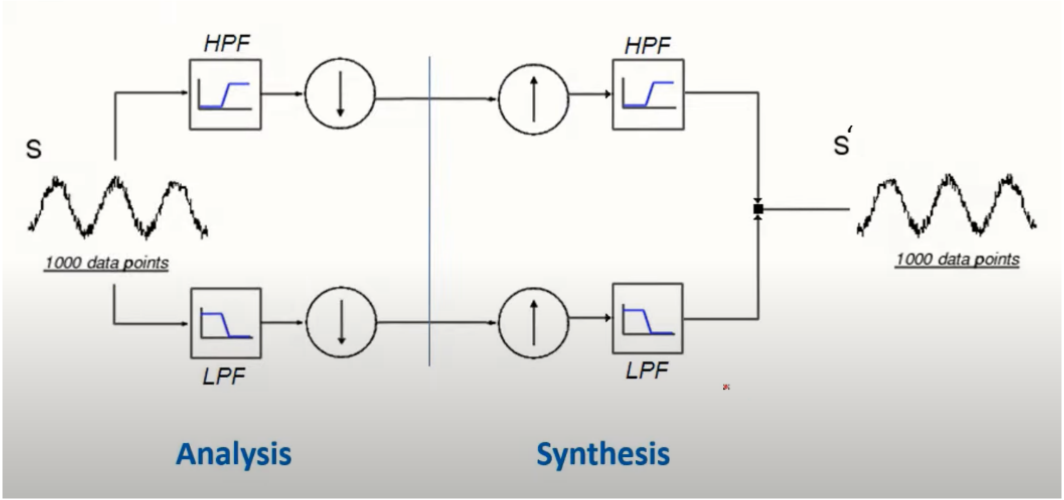



Figure: Signal Synthesis

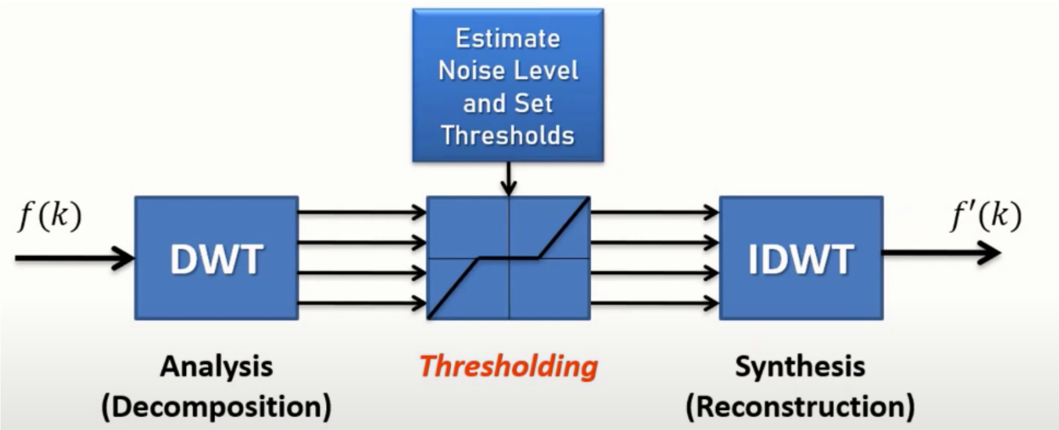



Figure: Denoising Scheme

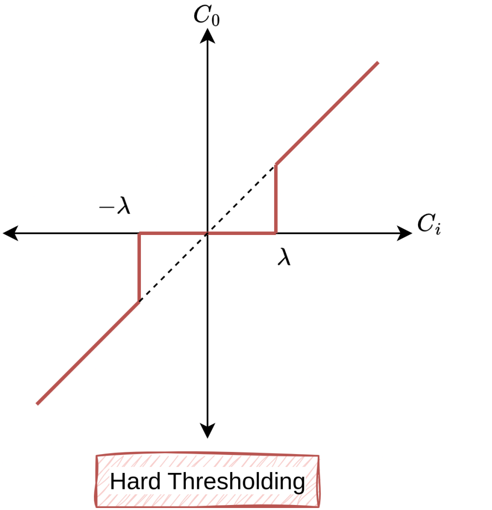
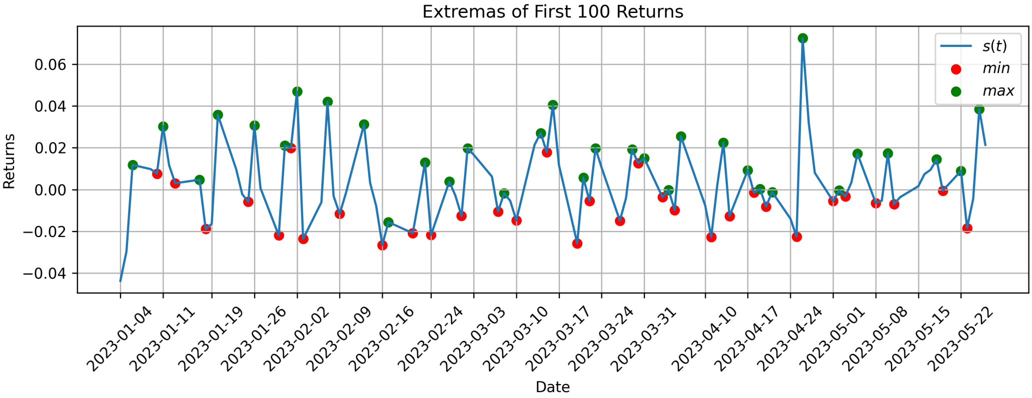



Figure: Thresholding Methods

# Empirical Mode Decomposition

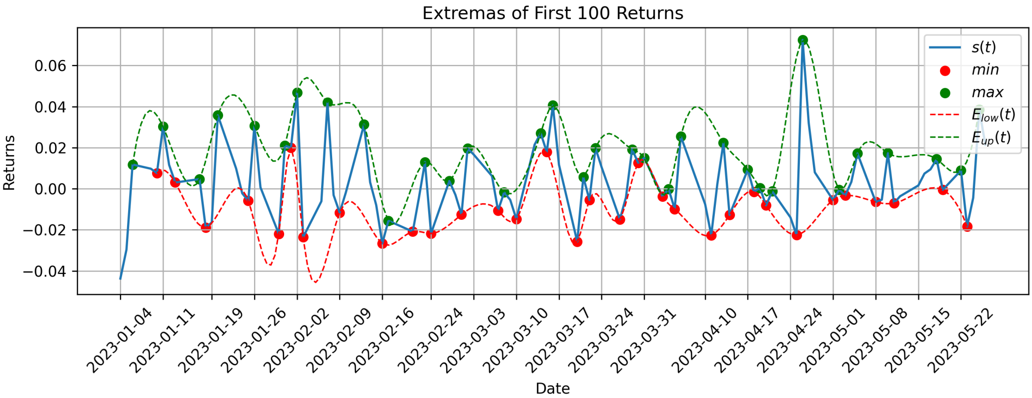
## Step 1: Identify Extremas

Identify the extremas of the signal $s(t)$



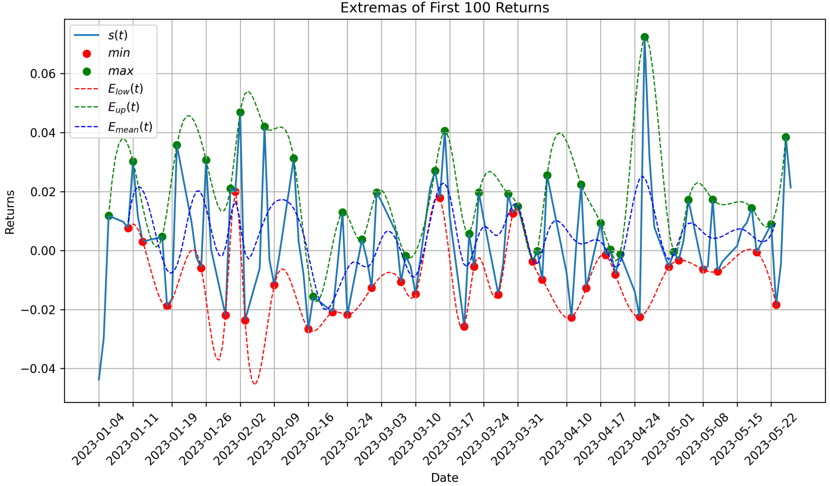
# Step 2: Interpolate

## Interpolate between maximas and minimas



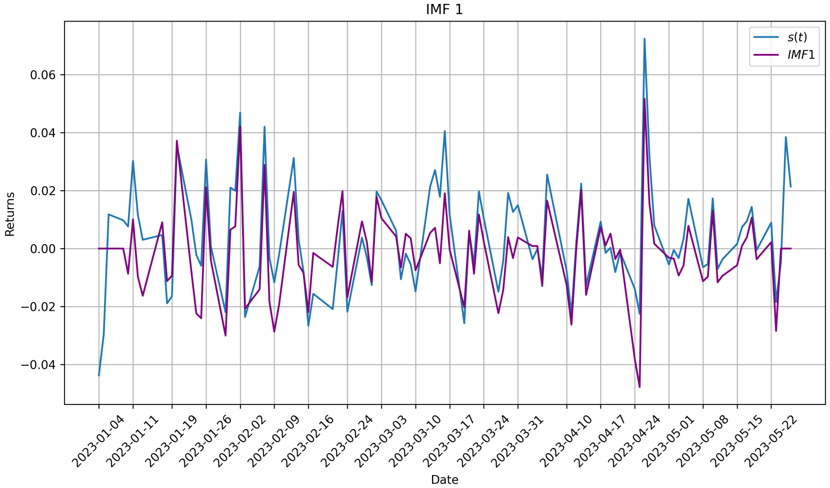
# Step 3: Mean Envelope

## Construct mean of upper and lower envelope



# Step 4: Extract the Details

## Keep extracting the details until zero-mean residual



# ICEEMDAN + Wavelets

## Algorithm

1. Use ICEEMDAN to decompose the time series $x(t)$ into different IMF components and a residual component $R$.

$$x(t) = \sum_{k=1}^{K} \mathrm{IMF}_k + R$$

2. Divide the set of IMF components into two sets. If $i$ is the separation point, then the two components are $\sum_{k=1}^{i} \mathrm{IMF}_k$ and $\sum_{k=i+1}^{K} \mathrm{IMF}_k$:

$$x(t) = \sum_{k=1}^{i} \mathrm{IMF}_k + \sum_{k=i+1}^{K} \mathrm{IMF}_k + R$$

# ICEEMDAN + Wavelets

## Algorithm

1. Use ICEEMDAN to decompose the time series $x(t)$ into different IMF components and a residual component $R$.

$$x(t) = \sum_{k=1}^{K} \mathrm{IMF}_k + R$$

2. Divide the set of IMF components into two sets. If $i$ is the separation point, then the two components are $\sum_{k=1}^{i} \mathrm{IMF}_k$ and $\sum_{k=i+1}^{K} \mathrm{IMF}_k$:

$$x(t) = \sum_{k=1}^{i} \mathrm{IMF}_k + \sum_{k=i+1}^{K} \mathrm{IMF}_k + R$$

3. Apply wavelet transform on the noise component $x(t)_{\text{noise}}$, pass it through suitable threshold and then invert it using inverse wavelet transform to get the denoised signal.

$$\begin{aligned} x(t) &= x(t)_{\text{noise}} + x(t)_{\text{non-noise}} \\ &= \epsilon(t) + \underbrace{\vec{x}(t) + x(t)_{\text{non-noise}}}_{\tilde{x}(t)} \\ &= \tilde{x}(t) + \epsilon(t) \end{aligned}$$

3. $\tilde{x}(t)$ is the denoised signal and $\epsilon(t)$ is the final noise component separated from the original signal.

3. Apply wavelet transform on the noise component $x(t)_{\text{noise}}$, pass it through suitable threshold and then invert it using inverse wavelet transform to get the denoised signal.

$$\begin{aligned} x(t) &= x(t)_{\text{noise}} + x(t)_{\text{non-noise}} \\ &= \epsilon(t) + \underbrace{\vec{x}(t) + x(t)_{\text{non-noise}}}_{\tilde{x}(t)} \\ &= \tilde{x}(t) + \epsilon(t) \end{aligned}$$

3. $\tilde{x}(t)$ is the denoised signal and $\epsilon(t)$ is the final noise component separated from the original signal.

# Dataset

1. *S&P500* returns were used.
2. Time-Period: Jan 1, 2001 to April 20, 2024

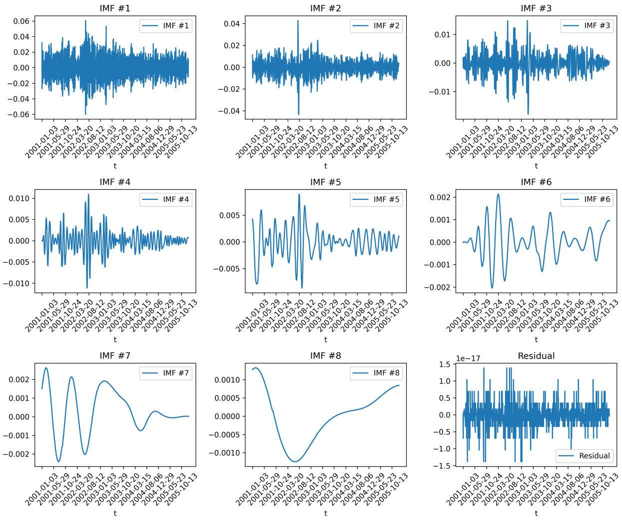



Figure: Decomposition into IMFs

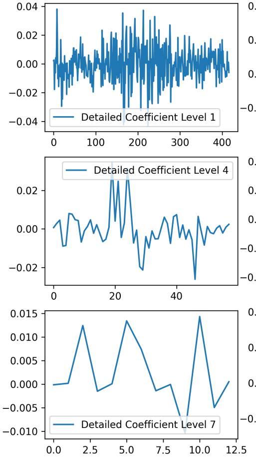



Figure: Detailed and Approximate Coefficients

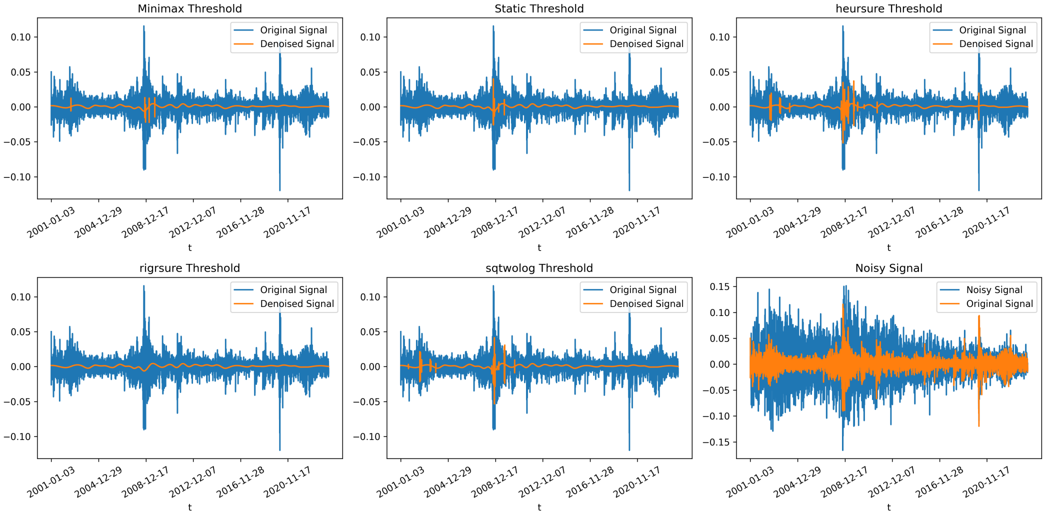



Figure: Result for Different Thresholds

# Result

$$\mathrm{SNR} = \frac{||origSig||_2^2}{||origSig - denSig||_2^2} \tag{1}$$

- SNR (Denoised Signal): 0.996
- SNR (Noisy Signal): 0.425