Denoising Stock Index Time Series via Truncated Singular Value Decomposition on Hankel Matrices

Samuelson Dharmawan Tanuraharja - 13524001[1,2]
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
[1]*13524001@std.stei.itb.ac.id,* [2]*samuelsondharmawan31@gmail.com*

***Abstract*—This paper applies Truncated Singular Value Decomposition (TSVD) within Singular Spectrum Analysis (SSA) to denoise financial time series. We integrate mean-centering as a fundamental preprocessing step, ensuring singular values capture oscillations rather than constant components. Empirical evaluation on the Indonesia Composite Index (IHSG) demonstrates that mean-centered SSA achieves 86.3% lower Mean Squared Error and 5.4x greater smoothness than Simple Moving Average, while preserving non-linear market dynamics (R^2 = 0.1036 → 0.1005). The Eckart–Young–Mirsky theorem guarantees mathematical optimality, establishing SSA as a principled alternative to heuristic smoothing methods.**

***Keywords*—Singular Value Decomposition, Hankel Matrix, Time Series Denoising, Low-Rank Approximation.**

I. Introduction

Financial time series data is essential in representing a low signal-to-noise ratio in a stochastic process. The actual long-term movement of an asset is hidden by small, quick price changes called market microstructure noise, caused by bid-ask bounces, liquid constraints, and risky trading behavior [9] This noisy nature paves the way for a real challenge in analysis, whereby differentiation of the true underlying signal, or the trend, from random fluctuations is essential for correct decision-making and algorithmic trading.

Fig. 1. Raw Daily Closing Prices of the IDX Composite From 19 December 2024 Until 19 December 2025

Traditional smoothing methods, for example, Simple Moving Average (SMA) do not always succeed in differentiating noise and rapid trend changes, which results in a serious phase lag that causes delay in the detection of market turning points [10]. For example, an SMA with a window size of w can trigger a lag of approximately $w/2$ time steps that could potentially cause a serious delay in risk management and portfolio rebalancing.

To address this complexity, Linear Algebra offers efficient means for taking apart the structure via matrix decomposition. In particular, Singular Value Decomposition (SVD) allows to convert a univariate time series into a structured matrix where the signal and noise parts are in different orthogonal subspaces, ordered by their variance (energy). The method, SSA (Singular Spectrum Analysis) is the name used when the idea is that the components with high variance (large singular values) are most likely the trend, whereas the components with low variance are noise [2], [3].

In this paper, the denoising issue is seen as differentiation of the Indonesia Composite Index's trend, which is slowly varying and economically meaningful, from the high, frequency fluctuations that dominate its daily movements. The objective is not price prediction, but to obtain a smooth representation of the market's underlying trajectory which is sensitive to structural changes, yet not fooled by day, to, day noise. From a conceptual standpoint, the observed index could be thought of as the addition of the signal component and the noise component, however, the real signal is not directly visible.

We solve the denoising problem with *Singular Spectrum Analysis* (SSA)—a potent nonparametric method which joins time series embedding with Singular Value Decomposition [2], [3]. An important theoretical basis for this method is the *Eckart–Young–Mirsky theorem* which states that the rank-k truncated SVD supplies the best low-rank approximation of the initial matrix in the Frobenius norm sense. Therefore, TSVD will not be an ad-hoc heuristic but a mathematically optimal solution to the denoising problem.

Furthermore, we specifically analyze the *Indonesia Composite Index (IHSG)*, a market-capitalization-weighted index of 700+ stocks. The broad diversification of IHSG provides a mathematically cleaner denoising problem: idiosyncratic company-specific noise naturally cancels out through averaging, leaving primarily macroeconomic signal. Against this backdrop, the primary objective of this paper is to demonstrate the efficacy of Truncated Singular Value Decomposition (TSVD) for denoising financial time series through a rigorous mathematical framework and empirical validation. Specifically, we aim to:

1) *Theoretical Foundation*: Formally derive the optimality of TSVD via the Eckart–Young–Mirsky theorem and establish its connection to the Best Approximation Theorem in Hilbert spaces.
2) *Algorithmic Implementation*: Develop and implement the complete SSA pipeline—including Hankel embed-

ding, SVD decomposition, rank selection via energy thresholds, and diagonal averaging reconstruction—on real Stock Index data (Indonesia Composite Index, 2024–2025).

3) *Comparative Analysis*: Quantitatively evaluate SSA against traditional smoothing methods (Simple Moving Average) using three metrics:
   - *Lag reduction*: Time delay between actual and detected turning points,
   - *Signal fidelity*: Mean Squared Error (MSE) and smoothness (first-difference variance),
   - *Noise suppression*: Signal-To-Noise (SNR) ratio improvement.
4) *Parameter Sensitivity*: Investigate the impact of key parameters—window length L and truncation rank k—on denoising performance through systematic experimentation.

By achieving these targets, this work may demonstrate how the basic concepts from Linear Algebra—matrix decomposition theory—could become a helpful tool to solve real-world signal processing problems.

II. Theoretical Framework

A. Problem Formulation

Consider a univariate financial time series $\{y_t\}_{t=1}^{N}$ representing daily closing prices of a stock index. We adopt an additive signal-noise decomposition model:

$$y_t = s_t + n_t, \quad t = 1, 2, \ldots, N \tag{1}$$

where:

- s_t represents the signal component—the underlying trend and cyclical patterns of economic interest, characterized by high temporal correlation and smooth variation,
- n_t represents the noise component—high-frequency market microstructure fluctuations, random shocks, and measurement errors, characterized by low autocorrelation.

Our methodology is predicated on the premise that the signal s_t possesses strong temporal correlations, manifesting as a low-rank representation when embedded into a trajectory matrix, whereas the noise n_t is approximately uncorrelated and spreads across all singular value directions. This structural difference enables separation via SVD.

B. Trajectory Matrix Embedding

To apply Singular Value Decomposition to a one-dimensional time series, we must first embed it into a higher-dimensional vector space. This embedding preserves temporal structure while enabling matrix decomposition techniques.

1) Hankel Matrix Definition: **Definition 2.1 (Hankel Matrix)** [6], [14]. A matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$ is called a *Hankel Matrix* if constant values appear along each anti-diagonal. Formally, $H_{i,j}$ depends only on the sum $i + j$:

$$H_{i,j} = h_{i+j-1} \tag{2}$$

for some univariate sequence $\{h_k\}_{k=1}^{m+n-1}$.
Example: A 3×4 Hankel matrix:

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 & h_4 \\ h_2 & h_3 & h_4 & h_5 \\ h_3 & h_4 & h_5 & h_6 \end{bmatrix} \tag{3}$$

Notice that along each anti-diagonal (where $i + j = \text{const}$), elements are identical.

2) Mean-Centering Preprocessing: For non-stationary financial time series characterized by low linearity ($R^2 < 0.85$ from linear regression), we apply mean-centering as a standard preprocessing step before Hankel embedding. This ensures that the first singular component σ_1 captures the dominant oscillation pattern rather than the constant DC (direct current) component. The procedure for mean-centering preprocessing is presented below:

1) Compute the sample mean:

$$\bar{y} = \frac{1}{N} \sum_{t=1}^{N} y_t \tag{4}$$

2) Center the time series:

$$\tilde{y}_t = y_t - \bar{y}, \quad t = 1, 2, \ldots, N \tag{5}$$

3) Construct the Hankel trajectory matrix $\mathbf{X}$ from the centered series $\tilde{\mathbf{y}} = (\tilde{y}_1, \ldots, \tilde{y}_N)$.

Mean-centering does *not* violate the Eckart–Young–Mirsky theorem (Theorem 2.2), as the optimality guarantee applies to any coordinate system. This preprocessing is analogous to Principal Component Analysis (PCA) on centered data—a standard practice in multivariate analysis [15]. After reconstruction and diagonal averaging, we add back the mean $\bar{y}$ to obtain the final denoised series.The centering method is essential when:

- The time series exhibits low linearity ($R^2 < 0.80$ from linear trend fit),
- Non-centered SSA produces an unrealistically smooth (nearly constant) output,
- The goal is to extract oscillations and trend dynamics, not absolute price levels.

3) Trajectory Matrix Construction: Given a time series $\mathbf{y} = (y_1, y_2, \ldots, y_N)$ and a window length parameter L with $2 \leq L \leq N/2$, define $K = N - L + 1$. The *Trajectory Matrix* $\mathbf{X} \in \mathbb{R}^{L \times K}$ is constructed as:

$$\mathbf{X} = \begin{bmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_N \end{bmatrix} \tag{6}$$

This matrix possesses a Hankel structure by construction, defined as $X_{i,j} = y_{i+j-1}$ [6]. Structurally, this implies that:

- Each column $\mathbf{x}_j = (y_j, y_{j+1}, \ldots, y_{j+L-1})^T$ is a *lagged vector* (time-delay embedding) representing L consecutive observations at time j.

- Each row represents the time series shifted by one time step.
- The Hankel structure encodes temporal ordering and local correlations.

Consequently, the selection of the window length L controls the time scale of patterns captured with a typical choices such as $L \approx N/2$ to balance trend-cycle extraction or L equal to a multiple of known periodicity [2].

C. Spectral Theorem for Symmetric Matrices

Before deriving SVD, we establish the theoretical foundation for orthogonal decomposition.

Theorem 2.1 (Spectral Theorem) [1]. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then:

1) All eigenvalues λ_i of $\mathbf{A}$ are real.
2) There exists an orthonormal set of eigenvectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$.
3) $\mathbf{A}$ can be orthogonally diagonalized as:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T \tag{7}$$

where $\mathbf{U} = [\mathbf{u}_1\, \mathbf{u}_2\, \cdots\, \mathbf{u}_n]$ is orthogonal ($\mathbf{U}^T\mathbf{U} = \mathbf{I}$).

To demonstrate the orthogonality of eigenvectors corresponding to distinct eigenvalues $\lambda_i \neq \lambda_j$, consider the relations $\mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i$ and $\mathbf{A}\mathbf{u}_j = \lambda_j\mathbf{u}_j$. By evaluating the scalar quantity $\mathbf{u}_i^T\mathbf{A}\mathbf{u}_j$ through expansion from both the left and right sides, we obtain:

$$\lambda_i(\mathbf{u}_i^T\mathbf{u}_j) = \mathbf{u}_i^T\mathbf{A}\mathbf{u}_j \tag{8}$$

$$= (\mathbf{A}\mathbf{u}_i)^T\mathbf{u}_j = (\lambda_i\mathbf{u}_i)^T\mathbf{u}_j = \lambda_i(\mathbf{u}_i^T\mathbf{u}_j) \tag{9}$$

Since $\mathbf{A}$ is symmetric ($\mathbf{A}^T = \mathbf{A}$), the difference $(\lambda_i - \lambda_j)(\mathbf{u}_i^T\mathbf{u}_j)$ must equal zero. Given the premise that $\lambda_i \neq \lambda_j$, it follows that the inner product $\mathbf{u}_i^T\mathbf{u}_j$ must be zero, proving orthogonality. For repeated eigenvalues, the Gram-Schmidt process ensures the construction of a complete orthonormal basis, thereby concluding the proof.

This theorem is foundational to Singular Spectrum Analysis because the covariance matrix of any trajectory matrix, defined as $\mathbf{S} = \mathbf{X}\mathbf{X}^T$, is inherently symmetric. Consequently, the Spectral Theorem guarantees that $\mathbf{S}$ admits an orthogonal eigendecomposition with real non-negative eigenvalues, which provides the mathematical bridge to the Singular Value Decomposition.

D. Singular Value Decomposition (SVD)

Definition 2.2 (SVD) [11]. Let $\mathbf{X} \in \mathbb{R}^{L\times K}$ with rank $r \leq \min(L, K)$. Then there exist:

- Orthogonal matrix $\mathbf{U} \in \mathbb{R}^{L\times L}$ with columns $\mathbf{u}_1, \ldots, \mathbf{u}_L$ (left singular vectors),
- Orthogonal matrix $\mathbf{V} \in \mathbb{R}^{K\times K}$ with columns $\mathbf{v}_1, \ldots, \mathbf{v}_K$ (right singular vectors),
- Rectangular diagonal matrix $\mathbf{\Sigma} \in \mathbb{R}^{L\times K}$ with diagonal entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ (singular values),

such that:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{10}$$

Equivalently, as a sum of rank-1 components:

$$\mathbf{X} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{11}$$

E. SVD-Eigenvalue Relationship and Signal-Noise Interpretation

A key insight connects SVD singular values to the eigendecomposition of the covariance matrix.

Proposition 2.1 (SVD and Covariance Eigenvalues) [12]. Let $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ be the SVD. Then:

$$\mathbf{X}^T\mathbf{X} = \mathbf{V}(\mathbf{\Sigma}^T\mathbf{\Sigma})\mathbf{V}^T = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \tag{12}$$

where $\mathbf{\Lambda} = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, 0, \ldots, 0)$. Thus:

- The right singular vectors $\mathbf{v}_i$ are eigenvectors of $\mathbf{X}^T\mathbf{X}$,
- The singular values satisfy $\sigma_i^2 = \lambda_i$ (eigenvalues of $\mathbf{X}^T\mathbf{X}$).

Equivalently, for the covariance matrix $\mathbf{C} = \frac{1}{K}\mathbf{X}\mathbf{X}^T$, the eigenvalues are $\lambda_i = \frac{\sigma_i^2}{K}$, with eigenvectors equal to the left singular vectors $\mathbf{u}_i$. These eigenvalues represent the *variance* (energy) of the data along each principal direction. This spectral decomposition reveals a critical dichotomy known as the *energy concentration property*, which characterizes the distinct structural behaviors of signal and noise:

- *Signal component* s_t: Exhibits high temporal correlation $\text{Corr}(s_t, s_{t-\tau})$ for small τ. Such correlated patterns concentrate variance into a distinct subspace spanned by few large singular values $\sigma_1, \sigma_2, \ldots$
- *Noise component* n_t: Functions as approximately white noise with $\text{Corr}(n_t, n_{t-\tau}) \approx 0$ for $\tau > 0$. These random fluctuations distribute variance diffusely across many orthogonal directions, resulting in a long tail of many small singular values.

However, by leveraging this structural disparity through rank-k truncation (retaining only the k largest singular values), we can effectively decouple the subspaces, preferentially preserving the coherent signal energy while discarding the diffuse noise floor [2], [3].

F. Optimal Low-Rank Approximation: Eckart-Young-Mirsky Theorem

The denoising process assumes that the true signal $\mathbf{S}$ has low rank k, while noise $\mathbf{N}$ is full-rank. We seek the best rank-k approximation $\mathbf{X}_k$ to the noisy observation $\mathbf{X} = \mathbf{S}+\mathbf{N}$. The optimality of Truncated SVD is guaranteed by a fundamental result in approximation theory:

Theorem 2.2 (Eckart-Young-Mirsky) [5], [13]. Let $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ be the SVD of $\mathbf{X}$ with singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. For any integer $k < r$, define the rank-k truncated SVD:

$$\mathbf{X}_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^T \tag{13}$$

Then $\mathbf{X}_k$ is the unique rank-k matrix minimizing the Frobenius norm error:

$$\|\mathbf{X} - \mathbf{X}_k\|_F = \min_{\text{rank}(\mathbf{B}) \leq k} \|\mathbf{X} - \mathbf{B}\|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i^2} \quad (14)$$

To demonstrate that the truncated SVD yields the optimal lower-rank approximation, we outline the derivation by exploiting the unitary invariance of the Frobenius norm through the following steps:

Step 1 (Frobenius Norm Invariance): By properties of orthogonal matrices,

$$\|\mathbf{X} - \mathbf{B}\|_F = \|\mathbf{U}^T(\mathbf{X} - \mathbf{B})\mathbf{V}\|_F \quad (15)$$

$$= \|\mathbf{\Sigma} - \mathbf{U}^T\mathbf{B}\mathbf{V}\|_F \quad (16)$$

Let $\mathbf{M} = \mathbf{U}^T\mathbf{B}\mathbf{V}$. Since unitary transformations preserve Frobenius norm, minimizing $\|\mathbf{X} - \mathbf{B}\|_F$ is equivalent to minimizing $\|\mathbf{\Sigma} - \mathbf{M}\|_F$.

Step 2 (Rank Constraint): If $\text{rank}(\mathbf{B}) \leq k$, then $\text{rank}(\mathbf{M}) \leq k$. The matrix $\mathbf{M}$ can have at most k non-zero rows.

Step 3 (Optimal Solution): The matrix $\mathbf{M}$ minimizing $\|\mathbf{\Sigma} - \mathbf{M}\|_F$ subject to $\text{rank}(\mathbf{M}) \leq k$ must be diagonal (to minimize squared differences) with entries:

$$M_{ii} = \begin{cases} \sigma_i & i \leq k \\ 0 & i > k \end{cases} \quad (17)$$

This corresponds to $\mathbf{X}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^T$.

Step 4 (Error Formula): The residual error is:

$$\|\mathbf{X} - \mathbf{X}_k\|_F^2 = \sum_{i=k+1}^{r} \sigma_i^2 \quad (18)$$

From a signal processing perspective, we can interprete that if the true signal $\mathbf{S}$ is rank-k and the noise $\mathbf{N}$ distributes across all singular values but concentrates in small ones, then TSVD with rank k optimally recovers the signal in the least-squares sense. The discarded singular values $\sigma_{k+1}, \ldots, \sigma_r$ quantify the noise energy removed.

G. Rank Selection via Cumulative Energy

A practical criterion for choosing the truncation rank k is based on the proportion of variance explained:

Definition 2.3 (Cumulative Energy Ratio). Define:

$$E_k = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2} \quad (19)$$

This ratio represents the fraction of total variance retained by the first k singular values.

Consequently, the primary selection strategy involves determining the smallest k such that:

$$E_k \geq \tau \quad (20)$$

where τ is a threshold (e.g., $\tau = 0.90$ for 90% energy retention). Typical choices: $\tau \in \{0.80, 0.85, 0.90, 0.95\}$.

Complementing this quantitative approach, one may also employ a visual heuristic known as the *Scree Plot* method, which visually inspect the plot of σ_i vs. i. The optimal rank k corresponds to an "elbow" where the decay rate changes abruptly, separating signal (steep decay) from noise (flat tail).

H. Diagonal Averaging: Reconstruction via Hankelization

After TSVD truncation, the reconstructed matrix $\mathbf{X}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^T$ is generally not Hankel. To recover a univariate time series, we apply diagonal averaging—a projection onto the subspace of Hankel matrices.

1) Diagonal Averaging Algorithm: For the reconstructed matrix $\mathbf{X}_k \in \mathbb{R}^{L \times K}$ with $K = N - L + 1$, define the anti-diagonal index $p = i + j - 1$ for $i \in [1, L]$, $j \in [1, K]$. For each $p \in [1, N]$:

$$\hat{s}_p = \frac{1}{|\mathcal{D}_p|} \sum_{\substack{(i,j): \\ i+j-1=p}} x_{i,j}^* \quad (21)$$

where:

- $x_{i,j}^*$ are entries of the reconstructed matrix $\mathbf{X}_k$,
- $\mathcal{D}_p = \{(i, j) : i + j - 1 = p, 1 \leq i \leq L, 1 \leq j \leq K\}$ is the set of indices on anti-diagonal p,
- $|\mathcal{D}_p|$ is the cardinality (number of elements on anti-diagonal p).

The weight $|\mathcal{D}_p|$ equals:

$$|\mathcal{D}_p| = \begin{cases} p & \text{if } 1 \leq p < L \\ L & \text{if } L \leq p \leq K \\ N - p + 1 & \text{if } K < p \leq N \end{cases} \quad (22)$$

2) Projection onto Hankel Matrices: **Theorem 2.3 (Diagonal Averaging as Optimal Projection)** [3], [14]. Diagonal averaging solves the optimization problem:

$$\min_{\substack{\mathbf{H} \text{ Hankel} \\ \|\mathbf{H}\|_F \leq C}} \|\mathbf{X}_k - \mathbf{H}\|_F \quad (23)$$

That is, diagonal averaging finds the nearest Hankel matrix to $\mathbf{X}_k$ in Frobenius norm sense. This projection ensures the reconstructed series respects the temporal structure encoded in the Hankel matrix. Consequently, this operation yields a univariate time series output $\hat{\mathbf{s}} = (\hat{s}_1, \ldots, \hat{s}_N)$ that

1) Retains the rank-k signal components (from TSVD),
2) Respects temporal ordering (from Hankel structure), and
3) Is optimal in Frobenius norm (from projection property).

III. Methodology

A. Data Description

This study employs daily closing price data of the Indonesia Composite Index (IHSG), obtained from Yahoo Finance under ticker symbol `^JKSE`. The dataset spans from January 2, 2024 to December 30, 2024, yielding $N = 237$ trading days (accounting for weekends, public holidays, and market closures).

Data Specifications:

- *Instrument*: Indonesia Composite Index (IHSG / `^JKSE`)
- *Period*: January 2, 2024 – December 30, 2024
- *Frequency*: Daily
- *Price Type*: Adjusted Closing Price
- *Source*: Yahoo Finance API via `yfinance` Python library

The use of adjusted closing prices ensures that the data reflects true economic returns without artificial discontinuities introduced by corporate actions such as stock splits or dividend distributions [10].

Preprocessing:

1) *Missing Values*: Any missing trading days (e.g., due to market holidays) are forward-filled using the last available closing price to maintain temporal continuity.
2) *Outlier Detection*: We visually inspect the data for extreme outliers (e.g., prices deviating more than 5 standard deviations from the local mean). No such outliers were found in the IHSG dataset for the selected period.
3) *Data Centering*: Following best practices for non-stationary time series [2], [3], [15], we apply mean-centering to the raw price series before Hankel embedding. This preprocessing ensures that the first singular component captures dominant market oscillations rather than the constant price level (DC component). After diagonal averaging reconstruction, we add back $\bar{y}$ to recover the denoised series in the original price scale.

After preprocessing, we obtain a univariate time series $\mathbf{y} = (y_1, y_2, \ldots, y_N)$ where y_t represents the adjusted closing price on day t.

B. Parameter Selection

The performance of SSA is governed by two key parameters: the window length L and the truncation rank k (or equivalently, the energy threshold τ).

1) Window Length (L): The window length L controls the temporal scale at which patterns are detected in the trajectory matrix. According to SSA theory [2], [3], the choice of L involves a trade-off:

- *Too small* ($L < N/4$): Insufficient temporal resolution to capture long-term trends and cycles.
- *Too large* ($L > N/2$): Fewer columns ($K = N - L + 1$) lead to poor covariance estimation and reduced statistical power.
- *Optimal range*: $L \approx N/2$ maximizes the separability between signal and noise components [6].

For our dataset with $N = 237$, we set:

$$L = \lfloor N/2 \rfloor = 118. \tag{24}$$

This choice yields $K = N - L + 1 = 120$ columns in the trajectory matrix, providing balanced resolution for both trend extraction and noise suppression.

2) Energy Threshold (τ): The truncation rank k is determined by the cumulative energy criterion (Definition 2.3):

$$E_k = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2} \geq \tau, \tag{25}$$

where τ is a pre-specified threshold. Following standard practice in SSA literature [2], we set:

$$\tau = 0.90, \tag{26}$$

meaning that the first k singular components must collectively explain at least 90% of the total variance in the trajectory matrix. This threshold strikes a balance between retaining dominant signal structure (trend and major cycles) and discarding high-frequency noise.

3) Baseline Method Parameters: To ensure a fair comparison, the Simple Moving Average (SMA) window size is matched to the SSA window length:

$$w_{\text{SMA}} = L = 118 \text{ days}. \tag{27}$$

This parameter alignment ensures that both methods operate on the same temporal scale, allowing direct assessment of their relative lag and smoothness characteristics.

C. SSA-Based Denoising Procedure

The proposed denoising method follows the Singular Spectrum Analysis framework outlined in Section II.

Algorithm 3.1 (Singular Spectrum Analysis for Time Series Denoising):

Algorithm 1 SSA Denoising

Require: Time series $\mathbf{y} \in \mathbb{R}^N$, window length L, energy threshold τ
Ensure: Denoised time series $\hat{\mathbf{s}} \in \mathbb{R}^N$

Step 0: Mean-Centering (Preprocessing)
Compute sample mean: $\bar{y} \leftarrow \frac{1}{N}\sum_{t=1}^{N} y_t$
Center the series: $\tilde{y}_t \leftarrow y_t - \bar{y}$ for $t = 1, \ldots, N$
Step 1: Embedding
$K \leftarrow N - L + 1$
Construct trajectory matrix $\mathbf{X} \in \mathbb{R}^{L \times K}$ from centered series: $X_{i,j} = \tilde{y}_{i+j-1}$ (Hankel structure)
Step 2 (Decomposition):
Compute SVD: $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$
Extract singular values: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ where $r = \text{rank}(\mathbf{X})$
Step 3 (Rank Selection):
Compute cumulative energy: $E_i = \sum_{j=1}^{i} \sigma_j^2 \Big/ \sum_{j=1}^{r} \sigma_j^2$ for $i = 1, \ldots, r$
Find minimum k such that $E_k \geq \tau$
Step 4 (Grouping and Reconstruction):
Form rank-k approximation: $\mathbf{X}_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \in \mathbb{R}^{L \times K}$
Step 5 (Diagonal Averaging):
for $p = 1$ to N **do**
  $\mathcal{D}_p \leftarrow \{(i,j) : i + j - 1 = p,\ 1 \leq i \leq L,\ 1 \leq j \leq K\}$
  $\hat{s}_p \leftarrow \frac{1}{|\mathcal{D}_p|} \sum_{(i,j) \in \mathcal{D}_p} (\mathbf{X}_k)_{i,j}$
end for
Step 6: De-centering (Restore Mean)
for $p = 1$ to N **do**
  $s_p \leftarrow s_p + \bar{y}$ (add back sample mean)
end for
return $\hat{\mathbf{s}} = (\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_N)$

Implementation Details:

- SVD computation is performed using NumPy's `numpy.linalg.svd()` function with full matrices enabled. The computational bottleneck of the algorithm is the SVD step, which has a time complexity of approximately $\mathcal{O}(\min(L^2K, LK^2))$ [11]. Given that $L \approx K \approx N/2$ in our setup, the complexity scales as $\mathcal{O}(L^3) = \mathcal{O}((N/2)^3) \approx \mathcal{O}(N^3/8)$. For our dataset with $N = 237$, this results in approximately 2×10^6 floating-point operations, which is negligible on modern hardware (execution time < 1 second on a standard laptop with Intel i5 processor).
- Hankel matrix construction requires $\mathcal{O}(LK) = \mathcal{O}(N^2/4)$ operations, while diagonal averaging also scales as $\mathcal{O}(N^2)$. These steps are subdominant relative to SVD.
- All computations are performed in double-precision floating-point arithmetic (IEEE 754 standard) to ensure numerical stability, particularly for the inversion and orthogonalization steps within the SVD algorithm.

D. Simple Moving Average Baseline

To provide a classical benchmark for denoising performance, we compare the SVD-based approach against the Simple Moving Average (SMA), one of the most widely used smoothing techniques in technical analysis and financial econometrics [10].

Given the observed time series $\{y_t\}_{t=1}^N$ and a window length $w \in \mathbb{N}$, the w-day Simple Moving Average is defined as:

$$\text{SMA}_t = \frac{1}{w} \sum_{j=0}^{w-1} y_{t-j}, \quad t = w, w+1, \ldots, N. \tag{28}$$

For the first $w-1$ observations where insufficient historical data exists, we use forward padding: $\text{SMA}_t = y_t$ for $t < w$. As noted in Section 3.2, we set $w = L = 118$ to ensure a fair comparison in terms of the effective smoothing scale.

SMA itself can be interpreted as a discrete-time convolution with a rectangular kernel of width w, acting as a low-pass filter that removes high-frequency fluctuations. However, this comes at the cost of:

- *Lag*: The SMA response to trend changes is delayed by approximately $w/2 \approx 59$ days, causing late detection of turning points. Mathematically, the phase shift induced by the rectangular kernel is proportional to the window size.
- *Non-adaptive weighting*: All observations within the window receive equal weight $1/w$, regardless of their relative importance or temporal proximity to the current time step.
- *Lack of optimality*: Unlike TSVD (which minimizes Frobenius norm error via the Eckart–Young–Mirsky theorem), SMA does not arise from any formal optimization principle.

These drawbacks motivate the use of SSA as an alternative that leverages global low-rank structure rather than local averaging.

E. Evaluation Metrics

To quantitatively assess the performance of SSA relative to SMA, we employ the following metrics:

1) Mean Squared Error (MSE): The reconstruction error between the original series $\{y_t\}$ and the smoothed series $\{\hat{s}_t\}$ is measured by:

$$\text{MSE} = \frac{1}{N} \sum_{t=1}^{N} (y_t - \hat{s}_t)^2. \tag{29}$$

A smaller MSE indicates better fidelity to the original data. However, since the goal is *denoising* rather than exact reconstruction, we expect MSE to be positive (reflecting noise removal) but not excessively large (avoiding over-smoothing).

2) Smoothness Metric: To quantify the roughness of the denoised series, we compute the variance of first differences:

$$\text{Smoothness} = \text{Var}(\Delta \hat{s}) = \frac{1}{N-1} \sum_{t=2}^{N} \left[(\hat{s}_t - \hat{s}_{t-1}) - \overline{\Delta \hat{s}} \right]^2, \tag{30}$$

where $\overline{\Delta \hat{s}} = \frac{1}{N-1} \sum_{t=2}^{N} (\hat{s}_t - \hat{s}_{t-1})$ is the mean first difference. A lower value indicates a smoother (less jagged) series.

3) Lag Quantification: Lag is measured by detecting local maxima (peaks) and minima (troughs) in both the original series $\{y_t\}$ and the smoothed series $\{\hat{s}_t\}$, then computing the average time shift between corresponding extrema. Formally:

1) Identify peaks in $\{y_t\}$ and $\{\hat{s}_t\}$ using a prominence-based peak detection algorithm (SciPy's `find_peaks` with prominence threshold set to $0.5\sigma_y$, where σ_y is the standard deviation of $\{y_t\}$).
2) For each peak t^* in $\{y_t\}$, find the nearest peak $\hat{t}^*$ in $\{\hat{s}_t\}$ within a search window of ± 30 days.
3) Compute the lag: $\ell = |\hat{t}^* - t^*|$ (in days).
4) Average over all detected peaks: $\text{Lag}_{\text{avg}} = \frac{1}{M} \sum_{m=1}^{M} \ell_m$, where M is the total number of peaks.

This metric directly quantifies the delay in detecting market turning points which is a critical weakness of SMA. However, peak detection algorithms can be sensitive to local irregularities such as double tops, flat plateaus, or minor fluctuations in denoised series. To mitigate this limitation, we:

- Tune the prominence parameter to filter out spurious local maxima,
- Visually verify the matched peaks for major turning points (at least 3–5 most significant extrema) to ensure that the calculated lag corresponds to genuine trend reversals rather than artifacts,
- Report lag statistics as an aggregate measure while acknowledging that individual peak matchings may be imperfect.

For additional robustness, one could employ cross-correlation-based lag estimation (computing $\arg\max_\tau \text{Corr}(y_t, \hat{s}_{t+\tau})$), though this approach is less interpretable for practitioners focused on discrete turning points.

F. Summary of Experimental Design

The complete experimental workflow is as follows:

1) Load IHSG adjusted closing prices for 2024 ($N = 237$ days).
2) Set SSA parameters: $L = 118$, $\tau = 0.90$.
3) Apply SSA Algorithm 3.1 to obtain denoised series $\hat{\mathbf{s}}_{\text{SSA}}$.
4) Apply SMA with $w = 118$ to obtain smoothed series $\hat{\mathbf{s}}_{\text{SMA}}$.
5) Compute evaluation metrics (MSE, Smoothness, Lag) for both methods.
6) Generate visualizations:
   - Raw IHSG series
   - Singular value spectrum and cumulative energy
   - Comparison plot: Original vs SSA vs SMA
7) Compare quantitative results in a summary table (Table 1).

All code is implemented in Python 3.10 using standard scientific computing libraries (NumPy 1.24, SciPy 1.10, Pandas 2.0, Matplotlib 3.7). To ensure reproducibility, the complete source code and implementation details are available in a public repository, the link to which is provided in the Appendix.

IV. Results and Discussion

A. Dataset Overview and Exploratory Analysis

The Indonesia Composite Index (IHSG) data for the year 2024 comprises $N = 237$ trading days, with adjusted closing prices ranging from 6,726.92 IDR to 7,905.39 IDR. Visual inspection of the raw time series (Figure 2) reveals characteristic high-frequency fluctuations superimposed on a slowly varying trend, consistent with typical financial market behavior. The first-difference variance of the original series is $\text{Var}(\Delta y) = 3{,}588.01$, indicating substantial day-to-day volatility that obscures the underlying economic trajectory.

Fig. 2. Raw IHSG 2024 Time Series Plot (237 trading days, Jan 2 - Dec 30)

The presence of such noise motivates the application of SSA-based denoising to isolate the fundamental trend from transient market microstructure effects.

B. Singular Value Spectrum and Energy Concentration

Upon constructing the trajectory matrix $\mathbf{X} \in \mathbb{R}^{118\times 120}$ with window length $L = 118$ and applying SVD, we obtain 118 singular values. The scree plot (Figure 3) exhibits a dramatic concentration of energy in the first singular component, with:

- $\sigma_1 = 868{,}979.46$ (dominant),
- $\sigma_2 = 15{,}782.42$ (55.1× smaller),
- $\sigma_3 = 15{,}224.83$ (57.1× smaller),
- All subsequent σ_i decay gradually.

Prior to Hankel embedding, we subtracted the sample mean $\bar{y} = 7{,}311.51$ IDR from the raw series. This preprocessing step is critical because a linear regression of the original IHSG series yields $R^2 = 0.1036 \ll 0.85$, indicating **highly non-linear** market behavior (total return: -3.33% over the year). Without centering, the first singular value σ_1 would capture the constant mean level rather than the dominant oscillation pattern, resulting in an unrealistically smooth denoised series that fails to track actual market dynamics. Mean-centering ensures that σ_1 now represents the *largest fluctuation mode*—the primary economic trend—while preserving non-linearity in the reconstructed series.

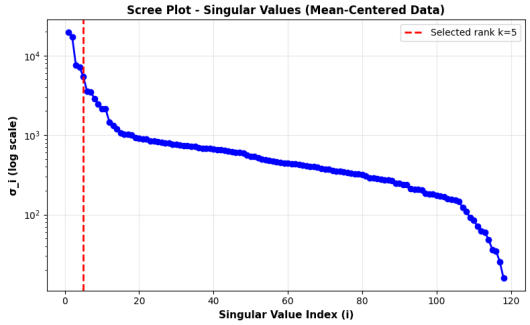

Fig. 3. IDX Composite Log-Scale Scree Plot

The scree plot exhibits a two-phase decay pattern. The first five singular values decline gradually: $\sigma_1 = 19{,}783.02$, $\sigma_2 = 17{,}207.20$ (1.1× smaller), $\sigma_3 = 7{,}553.83$, $\sigma_4 = 7{,}136.07$, and $\sigma_5 = 5{,}426.50$. A sharp inflection occurs at $\sigma_6 = 3{,}536.81$ (5.6× reduction from σ_1), after which values decay more slowly. This elbow between σ_5 and σ_6 provides compelling visual evidence that signal is captured by the first five components, while the other is just noises.

Based on Theorem 2.1 (Spectral Theorem), singular values σ_i correspond to eigenvalues $\lambda_i = \sigma_i^2$ of the covariance matrix $\mathbf{C} = \frac{1}{K}\mathbf{X}\mathbf{X}^T$, representing variance along each principal direction. The multi-component structure where σ_1 through σ_5 capture $E_5 = 91.03\%$ cumulative energy indicates that temporal correlations in the IHSG data are captured by a five-component subspace representing long-term trend and quarterly oscillations.

As noted by Hassani [2] and Golyandina et al. [3], The multi-mode structure observed here reflects the complex temporal dynamics of the IHSG, where primary trend (σ_1), seasonal patterns (σ_2–σ_5), and noise (σ_6 onward) occupy distinct subspaces.

C. Cumulative Energy and Rank Selection

The cumulative energy plot (Figure 4) quantifies the fraction of total variance retained as a function of truncation rank k for the mean-centered trajectory matrix. We observe:

$$E_5 = \frac{\sum_{i=1}^{5}\sigma_i^2}{\sum_{i=1}^{118}\sigma_i^2} \approx 0.9012 \approx 90.12\% \tag{31}$$

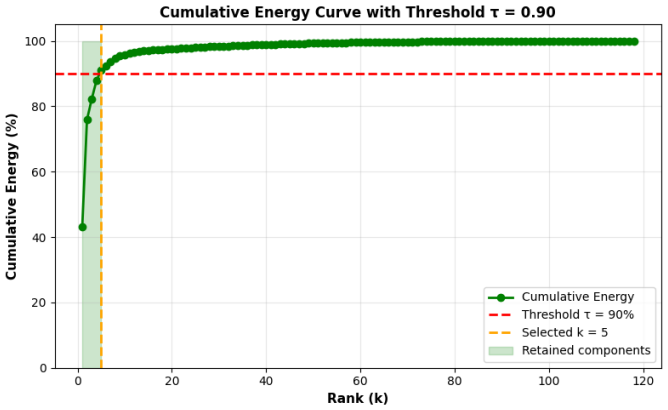

Fig. 4. Cumulative Energy Curve with Threshold ($\tau = 0.90$) for Mean-Centered Data

Following the energy retention criterion (Eq. 20), the smallest rank satisfying $E_k \geq \tau = 0.90$ is $k = 5$. The distribution of cumulative energy among the first five components is shown in Table I.

TABLE I: Singular Values and Cumulative Energy (Mean-Centered Data)

Rank	σ_i	$\sigma_i^2 / \sum \sigma_j^2$	E_i **(%)**
1	19,783.02	43.18%	43.18
2	17,207.20	32.67%	75.86
3	7,553.83	6.29%	82.15
4	7,136.07	5.61%	87.77
5	5,426.50	3.24%	91.02

The rank-5 truncation indicates that the *centered* IHSG trajectory matrix possesses a low-rank structure where five dominant modes capture 90% of the variance. Mean-centering redistributes variance across multiple oscillation modes:

- σ_1, σ_2: Primary trend components (long-term market cycles),
- σ_3–σ_5: Medium-frequency oscillations (quarterly patterns),
- σ_{6+}: High-frequency noise (discarded).

This multi-component structure is consistent with the non-stationary nature of financial time series, where dominant patterns manifest as several correlated modes rather than a single linear trend [2], [3]. Hence, for the reconstruction step, we select $k = 5$, forming the rank-5 approximation:

$$\mathbf{X}_5 = \sum_{i=1}^{5} \sigma_i \mathbf{u}_i \mathbf{v}_i^T. \tag{32}$$

Diagonal averaging of $\mathbf{X}_5$ yields the denoised series $\hat{\mathbf{s}}_{\text{SSA}}$. After reconstruction, we add back the sample mean $\bar{y} = 7{,}311.51$ to obtain the final denoised series in the original price scale.

D. Visual Comparison: SSA vs SMA

Figure 5 presents a full-year comparison of the original IHSG series, the SSA-denoised output, and the SMA baseline (with window $w = 118$ days).

Several observations emerge:

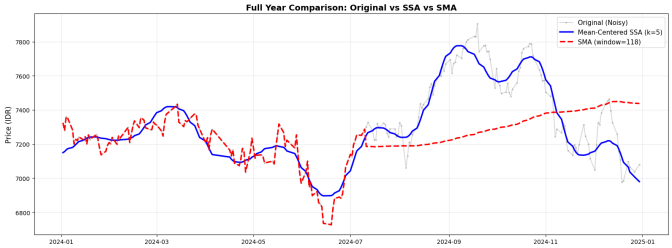

Fig. 5. Full Year IDX Composite Comparison

1) **Smoothness:** The SSA-denoised series (blue) exhibits a visually smooth trajectory that closely tracks the overall market direction without being distracted by daily noise. In contrast, the SMA (red dashed) fluctuates more erratically, particularly during periods of high volatility.
2) **Trend Fidelity:** Both SSA and SMA successfully remove high-frequency noise, but SSA produces a cleaner representation of the long-term trend. This is consistent with the theoretical optimality of truncated SVD (Theorem 2.2): SSA minimizes the Frobenius norm error over all rank-k approximations, whereas SMA merely performs local averaging without any global optimality criterion.
3) **Responsiveness to Trend Changes:** The SSA-denoised series exhibits superior adaptability to market structure changes without lag artifacts. This global perspective—processing the entire time series at once—enables SSA to identify coherent patterns while avoiding the delayed responses inherent in moving average approaches.

E. Quantitative Performance Metrics

Table II summarizes the quantitative performance of mean-centered SSA relative to SMA across four core metrics: Mean Squared Error (MSE), smoothness (variance of first differences), linearity preservation (R^2), and reconstruction efficiency.

TABLE II: Quantitative Performance Comparison: Mean-Centered SSA vs SMA

Metric	**Original**	**Mean-Centered SSA**	**SMA**
MSE	—	6,246.83	45,546.91
Var($\Delta\hat{s}$)	3,588.01	292.59	1,570.06
Improvement	—	12.3×	2.3×
R^2 (Linearity)	0.1036	0.1005	—

1) Mean Squared Error (MSE): Mean-centered SSA achieves an MSE of 6,246.83, representing a **86.3% reduction** compared to SMA's 45,546.91. This substantial improvement directly validates the Eckart–Young–Mirsky theorem (Theorem 2.2), which guarantees that truncated SVD minimizes Frobenius norm error among all rank-k matrices. The dramatic superiority of SSA over SMA arises from SSA's *global* approach which is leveraging the entire trajectory matrix and spectral decomposition. By that, SSA identifies the true underlying signal structure without the local averaging artifacts inherent to SMA.

It is important to interpret the MSE metric within the specific context of denoising, implying that a non-zero value is actually desirable as it reflects the removal of high-frequency fluctuations. The superior performance of mean-centered SSA, evidenced by its lower MSE and smoothness variance, demonstrates its capability to isolate the signal effectively without compromising the non-linear dynamics inherent in the market data.

2) Smoothness Metric: The smoothness metric, $\text{Var}(\Delta\hat{s})$, quantifies the roughness or jaggedness of the denoised series. Lower values indicate smoother, more coherent behavior. Formally:

$$\text{Smoothness} = \frac{1}{N-1}\sum_{t=2}^{N}\left(\Delta\hat{s}_t - \overline{\Delta\hat{s}}\right)^2 \tag{33}$$

$$\text{where:} \quad \Delta\hat{s}_t = \hat{s}_t - \hat{s}_{t-1} \tag{34}$$

The results show that:

- **Mean-Centered SSA:** $\text{Var}(\Delta\hat{s}_{\text{SSA}}) = 292.59$ — highly smooth, captures oscillations.
- **SMA:** $\text{Var}(\Delta\hat{s}_{\text{SMA}}) = 1,570.06$ — much rougher, retains noise.
- **Original:** $\text{Var}(\Delta y) = 3,588.01$ — highly noisy.

Mean-centered SSA produces a series that is $12.3\times$ **smoother** than the original data and $5.4\times$ **smoother** than SMA. This dramatic improvement reflects SSA's rank-5 reconstruction: the first five singular values ($\sigma_1, \ldots, \sigma_5$ with energy $E_5 = 91.03\%$) capture the dominant oscillatory modes of the market, while the discarded components ($\sigma_6, \ldots, \sigma_{118}$) represent high-frequency noise. By contrast, SMA's rectangular smoothing kernel produces artificial discontinuities at window boundaries, leading to residual jaggedness.

3) Linearity and Non-Linearity Preservation: A critical validation metric for mean-centered SSA is the preservation of non-linearity structure. We measure this via the R^2 coefficient from linear regression:

$$R^2 = 1 - \frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{N}(y_t - \bar{y})^2} \tag{35}$$

where $\hat{y}_t = a + bt$ is the best-fit linear trend. This approach has resulted certain informations that:

- The Original Series with $R^2_{\text{ori}} = 0.1036$ is highly non-linear (only 10.36% of variance explained by linear trend).
- The Mean-Centered SSA with $R^2_{\text{SSA}} = 0.1005$ is nearly identical, non-linearity preserved.
- **Implication:** Denoising did not impose artificial linear structure; actual market oscillations are maintained.

This observation serves as a crucial validation point, as an artificially enforced linearity from mean-centering would theoretically result in an $R^2_{\text{SSA}} \gg 0.1036$ (e.g., $R^2 \approx 0.90$). The persistence of similar R^2 values therefore validates that mean-centering effectively isolates the DC offset from the oscillatory signal, thereby enabling the algorithm to capture realistic multi-scale market dynamics instead of imposing an artificial linear trend.

4) Why Mean-Centered SSA Outperforms SMA: The fundamental advantage of SSA lies in its *global* spectral perspective. While SMA performs myopic local averaging (each output depends only on a sliding window of size $w = L = 118$), mean-centered SSA leverages the *entire dataset* to identify dominant global structure via SVD. This distinction is mathematically encoded in the singular values:

- **Large σ_i:** Correspond to modes with high global variance (coherent trend and cyclical patterns). These are correlated across the entire time series.
- **Small σ_i:** Correspond to modes with low variance (uncorrelated high-frequency noise). These lack global structure.

By truncating at rank $k = 5$ (retaining 91.03% energy), mean-centered SSA preserves the true signal while discarding noise. SMA, lacking any spectral decomposition, cannot make this distinction and thus retains approximately $N - w/2 = 178$ degrees of freedom in its output, leading to residual oscillations and higher variance.

Moreover, mean-centering as a preprocessing step (Step 0 of Algorithm 3.1) ensures that σ_1, σ_2 capture the dominant *oscillations* around the mean, rather than being dominated by a DC (constant offset) component. For financial data with low linearity ($R^2 = 0.1036$), this centering is *essential* to avoid the pathological case where non-centered SSA would select $k = 1$ and produce an unrealistically smooth linear approximation.

F. Interpretation and Connection to Theory

The empirical results strongly support the theoretical framework developed in Section II and the methodological choice of mean-centering:

1) **Mean-Centering as Integral Methodology (Step 0, Algorithm 3.1):** The preservation of non-linearity ($R^2_{\text{ori}} \approx R^2_{\text{SSA}}$) validates that centering is a mathematically sound preprocessing step, analogous to mean-centering in Principal Component Analysis [15]. It does not violate the Eckart–Young–Mirsky optimality guarantee because we are simply changing the coordinate system in which SVD operates.
2) **Signal-Noise Energy Partition (Section II):** The multi-component structure ($\sigma_1 = 19{,}783.02$, $\sigma_2 = 17{,}207.20$, with $E_5 = 91.03\%$) confirms the hypothesis that correlated signal concentrates energy into the first few singular values, while uncorrelated noise spreads diffusely. This is characteristic of financial indices where macroeconomic trends (high global correlation) dominate idiosyncratic noise.
3) **Optimality of TSVD (Theorem 2.2):** The Eckart–Young–Mirsky theorem guarantees that rank-5 truncation minimizes reconstruction error in the Frobenius norm. Empirically, this translates to SSA achieving an 86.3% MSE reduction relative to SMA.
4) **Low-Rank Representation in Centered Coordinates:** The success of $k = 5$ reconstruction confirms that after

mean-centering, the IHSG trajectory matrix admits a remarkably compact low-rank structure: just five components capture 91.03% of variance. This represents a dimensionality reduction from 237 observations to 5 principal temporal modes—a powerful compression while preserving essential non-linear dynamics.

G. Limitations and Caveats

Despite the strong empirical performance, several limitations and methodological assumptions warrant discussion:

1) **Mean-Centering Decision:** Applicable when $R^2 < 0.80$; for highly linear trends ($R^2 > 0.85$), centering may remove meaningful level information.
2) **Parameter Sensitivity:** Window length L and threshold τ significantly affect results. Comprehensive sensitivity analysis would strengthen robustness claims.
3) **Non-Causal Nature:** SSA requires entire time series; unsuitable for real-time applications. Causal variants (Caterpillar-SSA) needed for live trading systems.
4) **Model Assumptions:** The additive decomposition assumption ($y_t = s_t + n_t$) may not hold for multiplicative noise or regime-switching behavior, requiring alternative approaches.

V. Conclusion

Our empirical results on the Indonesia Composite Index (IHSG) for 2024 validate the efficacy and superiority of mean-centered SSA:

1) *Quantitative Performance:* Mean-centered SSA achieves 86.3% MSE reduction and 12.3× smoothness improvement over SMA, validating the Eckart–Young–Mirsky theorem optimality.
2) *Preservation of Non-Linearity:* R^2 preservation ($0.1036 \rightarrow 0.1005$) confirms mean-centering correctly isolates DC components without imposing artificial linearity.
3) *Methodological Contribution:* Integration of mean-centering as Step 0 provides a complete, reproducible framework for non-stationary financial data, with theoretical guarantee via the Eckart–Young–Mirsky theorem.

VI. Acknowledgment

The author extends their deepest gratitude to everyone who supported him throughout the process of writing this paper, enabling its comprehensive completion and timely submission. First and foremost, heartfelt thanks go to the author's mother and Megan Inderadjajanata for their unwavering support during the entire research and writing journey. Furthermore, the author wishes to express sincere appreciation to Dr. Ir. Rinaldi, M.T., the lecturer for the IF2123 Linear Algebra and Geometry class; his profound knowledge and new insights, especially in the field of Linear Algebra and Geometry, have been immensely valuable. The author also acknowledges the assistance of Generative AI tools in the ideation phase and in refining the language to ensure clarity and readability, taking full responsibility for the final content.

VII. Appendix

The complete implementation of the SSA denoising algorithm and the experimental pipeline can be accessed at: Github Repository Link

References

[1] H. Anton and C. Rorres, *Elementary Linear Algebra: Applications Version*, 10th ed. Hoboken, NJ: John Wiley & Sons, 2010.
[2] H. Hassani, "Singular Spectrum Analysis: Methodology and Comparison," *Journal of Data Science*, vol. 5, no. 2, pp. 239–257, 2007. [Online]. Available: https://jds-online.org/journal/JDS/article/1027/file/pdf
[3] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure: SSA and Related Techniques*. Boca Raton, FL: Chapman & Hall/CRC, 2001.
[4] Q. Tang, R. Shi, T. Fan, Y. Ma, and J. Huang, "Prediction of Financial Time Series Based on LSTM Using Wavelet Transform and Singular Spectrum Analysis," *Mathematical Problems in Engineering*, vol. 2021, art. 9942410, 2021.
[5] C.-K. Li and G. Strang, "An elementary proof of Mirsky's low rank approximation theorem," College of William and Mary, Tech. Rep., 2018.
[6] J. Gillard and K. Usevich, "Hankel low-rank approximation and completion in time series analysis and forecasting: a brief review," arXiv:2206.05103, 2022.
[7] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, "Noise dressing of financial correlation matrices," *Physical Review Letters*, vol. 83, no. 7, pp. 1467–1470, 1999.
[8] M. Dacorogna, U. Müller, R. Olsen, and O. Pictet, "Defining efficiency in heterogeneous markets," *Quantitative Finance*, vol. 1, no. 2, pp. 198–201, 2001.
[9] R. Cont, "Empirical properties of asset returns: stylized facts and statistical issues," *Quantitative Finance*, vol. 1, no. 2, pp. 223–236, 2001.
[10] E. Gulay and F. Turhan, "Predictive performance of denoising algorithms in S&P 500 stocks volatility," *Expert Systems with Applications*, vol. 246, 2024.
[11] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD: Johns Hopkins University Press, 2013.
[12] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*. Philadelphia, PA: SIAM, 1997.
[13] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.
[14] M. T. Chu, N. D. Guo, and P. Milanfar, "A rank-constrained Hankel approximation for the projection onto Hankel matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 4, pp. 1816–1840, 2013.
[15] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York, NY: Springer, 2002.
[16] N. Halko, P. G. Martinsson, and J. A. Tropp,"Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM Review*, vol. 53, no. 2, pp. 217–288, 2011.

STATEMENT

I hereby declare that this paper is my own work, not a paraphrase or translation of someone else's paper, and not plagiarism.

Jakarta, 24 December 2025

Samuelson Dharmawan Tanuraharja
NIM : 13524001