J. Indones. Math. Soc.
Vol. 32, No. 01 (2026), pp. 1–19.

Utilizing Trajectory Matrices and Singular Value Decomposition (SVD) for Multivariate Transformation in Time Series Analysis

Dina Prariesa[1], Udjianna Sekteria Pasaribu[2*], and Utriweni Mukhaiyar [2]

[1]Doctoral Program in Mathematics, Institut Teknologi Bandung, Indonesia
[2]Statistics Research Division, Institut Teknologi Bandung, Indonesia

Abstract. The trajectory matrix transforms univariate time series data into multivariate form using the structural properties of the Hankel Matrix (HM). Research on data matrices within Time Series Analysis (TSA) remains limited. This study examines AR models with stationary properties and applies Singular Value Decomposition (SVD) to HM in the Box-Jenkins framework. It focuses on HM properties, matrix dimension considerations in SVD, and order identification. Numerical simulations of the AR(1) and AR(2) models reveal that the PACF and SVD scree plots exhibit similar patterns. This indicates that applying SVD to HM could serve as an alternative to PACF for AR order selection. The findings highlight potential future research directions by refining, adapting, and generalizing previous studies to advance the TSA methodology.

Key words and Phrases: Autoregressive (AR) model, Hankel Matrix (HM), Singular Value Decomposition (SVD), Time Series Analysis (TSA).

1. INTRODUCTION

In TSA, hidden patterns and correlations can be revealed by transforming a univariate sequence into a multidimensional structure. A crucial step in this process is the construction of a trajectory matrix, specifically using the HM structure. In this study, a lag of 1 is chosen to construct the HM, as temporal dependencies in weakly stationary processes, particularly in AR(1) and AR(2) models, can be effectively captured. This choice is aligned with PACF interpretation and is utilized to facilitate SVD-based dimensionality reduction. By structuring the HM in this way, a balance between information preservation and computational efficiency is ensured in time series modeling.

*Corresponding author : udjianna@itb.ac.id
2020 Mathematics Subject Classification: 62M10; 62M15, 62H25
Received: 05-03-2025, accepted: 13-06-2025.

Given a time series $\{z_t, t = 1, 2, \ldots, T, T \in \mathbb{N}\}$ generated from a weakly stationary process, the objective is to transform this series into an HM of dimensions $L \times W$, constructed by defining a window length L and setting $W = T - L + 1 \in \mathbb{Z}$. Within the HM, there exists an index r such that $r \leq \min\{L, W\}$. Each data point z_t is then mapped to a lagged vector of length L, represented as:

$$\mathbf{Z}_t = (z_t, z_{t+1}, z_{t+2}, \ldots, z_{t+L-1})', \quad 1 \leq t \leq W$$

The resulting trajectory matrix has an HM structure as follows [1], [2], [3]:

$$\mathbf{H} = [\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_W] = \begin{bmatrix} z_1 & z_2 & \cdots & z_W \\ z_2 & z_3 & \cdots & z_{W+1} \\ \vdots & \vdots & \ddots & \vdots \\ z_L & z_{L+1} & \cdots & z_T \end{bmatrix} \tag{1}$$

The use of trajectory matrices with $\mathbf{H}$ aligns with the Singular Spectrum Analysis (SSA) method, introduced by Broomhead and King in 1986. SSA has been extensively studied, covering theoretical and methodological aspects [4],[5],[6], software implementation [7], [8], and practical applications in research and industry [9], [10], [11]. SSA, known for its flexibility, comprises four stages: embedding, SVD, grouping, and reconstruction. This study focuses on embedding and SVD for dimensionality reduction methods to assess their effectiveness in capturing temporal patterns in time series data.

SVD involves the factorization of $\mathbf{H}$ into three elementary matrices: two orthogonal matrices ($\mathbf{U} \in \mathbb{C}^{L \times L}$ and $\mathbf{V} \in \mathbb{C}^{W \times W}$) and one diagonal matrix ($\mathbf{D} \in \mathbb{R}^{L \times W}$) containing scale factors, termed singular values (λ_i^*), arranged in descending order, from the largest to the smallest value, denoted as $\lambda_1^* \geq \lambda_2^* \geq \cdots \geq \lambda_r^* > 0$. [12], [13],[14]. The columns of $\mathbf{U}$ and $\mathbf{V}$, referred to as left and right singular vectors. This decomposition can be expressed as follows:

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^* = \mathbf{U} \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}^* \tag{2}$$

SSA is a model-free method that disregards weak stationarity and autocorrelation assumptions. Given that TSA is inherently stochastic with autocorrelation behavior, integrating SSA with classical TSA methods would be insightful. Among these, the Box-Jenkins method remains prominent, influencing diverse fields such as economics [15], social sciences [16], healthcare [17], tourism [18], industries [19], and agriculture [20]. Its broad and lasting impact is well-documented [21], highlighting its continued relevance.

The Box-Jenkins methodology consists of three main components: Autoregressive (AR), Moving Average (MA), and Integrated (I) models, which together form advanced models like ARIMA (p, d, q) [22],[23],[24],[25],[26], [27]. In TSA, AR models highlight data dependency, where the order is identified using the autocorrelation function, particularly PACF [28], [29], [30]. A PACF plot helps determine the appropriate lag order p; a sharp drop after lag p suggests an AR(p) process, while a gradual decline may indicate an ARMA process. Understanding these functions is essential for selecting an optimal forecasting model.

Inspired by this method, this study explores data matrix fields through trajectory matrices with HM structures, providing a novel analytical approach. Investigating SVD on HM is particularly intriguing, as singular values may offer crucial insights into variance within AR models. Therefore, adhering to the principle of parsimony, the study initiates with the simplest mean-centred AR(1) and AR(2) model structure for $\tilde{Z}_t = Z_t - \delta$, with a general form of

$$\begin{aligned} \mathbf{AR(1)}: &\quad \tilde{Z}_t = \phi \tilde{Z}_{t-1} + \varepsilon_t \\ \mathbf{AR(2)}: &\quad \tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \varepsilon_t \end{aligned}$$

This model provides a foundational framework for understanding dynamics and variance captured by the HM structure and SVD. Implementing SVD for dimensionality reduction presents challenges, from optimizing matrix size to analyzing singular value sequences in AR models. Addressing these challenges, this study examines the properties and conditions of $\mathbf{H}$ in SVD to enhance its implementation. Insights from this exploration are applied to simulated AR model data, enabling a comprehensive evaluation of SVD's effectiveness.

2. PROPERTIES OF $\mathbf{H}$

Since this method applies to all types of observational data, the notation X is used for generality. A dataset $\{X_n\}, n \in \mathbb{Z}$ with N observations represents one-dimensional data and is transformed into a matrix $\mathbf{H}$ with elements $x_{i,j}$. From this formulation, several properties of $\mathbf{H}$ can be deduced. Literature on matrix analysis [3],[14],[31] provides the basis for constructing lemmas and theorems in this study. These results are broadly applicable to observational data, including time series.

Definition 2.1. *Matrix* $\mathbf{H}$ *has the following properties:*

i. *A notable property of the matrix* $\mathbf{H}$ *is its anti-diagonal, which remains constant from left to right. This property is defined by the condition* $i \leq j$*, where the element* $x_{i,j} = x_{i+m,j+m}$ *for* $m = 0, \ldots, j-i$*.* $\mathbf{H}$ *exhibits a characteristic structure known as the* **Hankel Matrix**.
ii. *As a consequence of point (i), the elements of* $\mathbf{H}$ *exhibit similarity to the elements of its conjugate transpose matrix* ($\tilde{\mathbf{H}}$)*. Mathematically, this can be expressed as* $\mathbf{H} = \tilde{\mathbf{H}} = \mathbf{H}^*$*. In other words, each element* $x_{i,j} = \widetilde{x}_{j,i}$*. This property is characterized as a Hermitian matrix, which is self-adjoint.*
iii. *Due to points (i) and (ii), it follows that* $\mathbf{H}\mathbf{H}^* = \mathbf{H}^*\mathbf{H}$*. This property is characterized as a* **Normal Matrix**.

The structure of HM for time series data, as defined in **Definition 2.1**, progresses either downwards or rightwards along its columns or rows as time moves forward [32]. This structural property aligns the columns of the HM as delayed versions of the time series, effectively capturing a lag-one relationship within a weak stationary process. Given the unique characteristics of HM in time series,

it is crucial to understand their general properties, as outlined in the following lemma. This lemma will lead to the properties and conditions for sufficient SVD in HM which is a benchmark for processing time series data

Lemma 2.2. *Based on* **Definition 2.1** *point (ii),* $\mathbf{H} = \widetilde{\mathbf{H}} = \mathbf{H}^*$*, hence* $\lambda^* \in \mathbb{R}$*.*

Proof. Let λ^* be an eigenvalue of $\mathbf{H}$ and y be the corresponding eigenvector. Then:

$$\begin{aligned}\lambda^*\langle y, y\rangle &= \langle \mathbf{H}y, y\rangle = \langle y, \mathbf{H}y\rangle \\ &= \langle y, \mathbf{H}^*y\rangle = \langle y, \mathbf{H}y\rangle = \langle y, \lambda^* y\rangle = \widetilde{\lambda^*}\langle y, y\rangle\end{aligned}$$

Thus, $\lambda^*(y) = \mathbf{H}(y) = \mathbf{H}^*(y) = \widetilde{\lambda^*}(y)$.
Consequently, it follows that $\lambda^* = \widetilde{\lambda^*}$, and hence $\lambda^* \in \mathbb{R}$. □

Lemma 2.3. *According to* **Definition 2.1** *point (iii),* $\mathbf{HH}^* = \mathbf{H}^*\mathbf{H}$ *are self adjoint and thus share the same* λ*, including multiplicities.*

Proof. Given $(\mathbf{HH}^*) = \mathbf{H}^*(\mathbf{H}^*)^* = \mathbf{H}^*\mathbf{H}$, it is concluded that $\mathbf{H}^*$ is self-adjoint. Similarly, $(\mathbf{HH}^*)^* = (\mathbf{H}^*)^*\mathbf{H}^* = \mathbf{H}^*\mathbf{H}$ is obtained, hence $\mathbf{HH}^*$ is also self-adjoint. Let λ be an eigenvalue of $\mathbf{HH}^*$ with eigenspace E_λ; for $v \in E_\lambda$, it is found that:

$$\mathbf{H}^*\mathbf{H}v = \lambda v$$

Multiply both sides by $\mathbf{H}$:

$$\mathbf{HH}^*\mathbf{H}v = \mathbf{H}\lambda v = \lambda \mathbf{H}v$$

Thus, $\mathbf{H}v$ is an eigenvector of $\mathbf{HH}^*$ with eigenvalue λ, because:

$$\mathbf{H}(\mathbf{H}^*v) = \lambda(\mathbf{H}v)$$

If $\lambda \neq 0$, then each eigenvector v from $\mathbf{HH}^*$ maps to an eigenvector $\mathbf{H}v$ of one-to-one. This is because if $v, w \in E_\lambda$ with $\mathbf{H}v = \mathbf{H}w$:

$$\mathbf{H}v = \mathbf{H}w \Rightarrow \mathbf{H}v - \mathbf{H}w = 0$$

which implies $\lambda v = \lambda w$, thus $v = w$. Similarly, every eigenvalue of $\mathbf{HH}^*$ maps to an eigenvalue of $\mathbf{H}^*\mathbf{H}$ with eigenvector $\mathbf{H}^*v$, mapping from the eigenspace E_λ of $\mathbf{HH}^*$, one to one, for $\lambda \neq 0$. For each $\lambda \neq 0$ of $\mathbf{HH}^*$ and $\mathbf{H}^*\mathbf{H}$, the associated eigenspaces have the same dimension. Because $\mathbf{HH}^*$ and $\mathbf{H}^*\mathbf{H}$ are self-adjoint and have the same dimension, the associated eigenvalues for $\lambda = 0$ also have the same dimension. Consequently, the multiplicities are also the same. □

The matrix $\mathbf{H}$ in Eq. (1), which consists of $\{\bar{z}_t\}$ reflects its properties through in the multiplication with its transpose, namely $\mathbf{HH}^*$ and $\mathbf{H}^*\mathbf{H}$, resulting in a matrix that resembles the autocovariance matrix ($\mathbf{C}$). The $\{c_{ij}\}$ of ($\mathbf{C}$) represents the covariance between $\{\bar{z}_t\}$ at different lags defined as:

$$C_{ij} = \gamma(|i-j|) = \mathrm{Cov}(\bar{z}_t, \bar{z}_{t+|i-j|}) = \mathbb{E}[\bar{z}_t \bar{z}_{t+|i-j|}] \tag{3}$$

Since the process is weakly stationary, the $\{c_{ij}\}$ only depend on the lag $|i-j|$. The multiplication of the $\mathbf{HH}^*$ results in an $L \times L$ matrix:

$$(\mathbf{HH}^*)_{ij} = \sum_{t=1}^{W} \bar{z}_{i+t-1}\bar{z}_{j+t-1}$$

Meanwhile, the multiplication $\mathbf{H}^*\mathbf{H}$ produces a $W \times W$ matrix:

$$(\mathbf{H}^*\mathbf{H})_{ij} = \sum_{t=1}^{L} \bar{z}_{t+i-1}\bar{z}_{t+j-1}$$

When normalized by the length L or W, these forms approximate C_{ij} as follows:

$$\frac{1}{W}(\mathbf{HH}^*)_{ij} \approx C_{ij} = \gamma(|i-j|) \approx \frac{1}{L}(\mathbf{H}^*\mathbf{H})_{ij}$$

Although $\mathbf{HH}^*$ and $\mathbf{H}^*\mathbf{H}$ have different dimensions, they exhibit similar spectra of eigenvalues and share the same eigenvalues (λ_r) as stated in **Lemma 2.3**. Since they are symmetric and positive semi-definite, λ_r are guaranteed to be non-negative real numbers, as described in **Lemma 2.2**. Moreover, the singular values of $\mathbf{H}$ (λ_r^*) are directly related to these eigenvalues, as

$$\lambda_r^* = \sqrt{\lambda_r}, \quad \text{with} \quad r = \min(L, W).$$

This implies that the choice of L or W does not affect the fundamental spectral properties of the matrices. Both matrices capture essential statistical properties of TSA. For instance, in autocorrelation behavior, the autocorrelation matrix ($\mathbf{R}$) is obtained by normalizing equation (3) based on its variance:

$$R_{ij} = \rho(i-j) = \frac{\gamma(|i-j|)}{\gamma(0)}$$

with $\rho(k)$ is the autocorrelation function at lag k and $\gamma(0)$ is the time series variance. The relationship between $\mathbf{C}$ and $\mathbf{R}$ is summarized as:

$$R_{ij} = \frac{C_{ij}}{C_{ii}}$$

Since $C_{ii} = \gamma(0)$ for all i, each element R_{ij} is obtained by normalizing C_{ij} with the variance $\gamma(0)$, yielding a correlation value within [-1,1], i.e., $-1 < \rho(i-j) < 1$. To analyze the relationship between singular value sequences and autocorrelation behavior, it is crucial to consider implications from **Definition 2.1** , point *(iii)*, and **Lemma 2.3**. Given that $\mathbf{H}$ is a normal matrix, the following theorem applies in the SVD process.

Theorem 2.4.

$$\mathbf{H} = \mathbf{UDV}^* \Leftrightarrow \mathbf{H}^*\mathbf{H} = \mathbf{HH}^*.$$

Proof. ($\rightarrow$)

$$\mathbf{H}^*\mathbf{H} = (\mathbf{UDV}^*)^*\mathbf{UDV}^* = \mathbf{VDU}^*\mathbf{UDV}^* = \mathbf{VDIDV}^* = \mathbf{VDDV}^*$$

Similarly, using the same approach, we obtain $\mathbf{HH}^* = \mathbf{UDDU}^*$.According to **Lemma 2.3**, $\mathbf{H}^*\mathbf{H}$ and $\mathbf{HH}^*$ have the same λ_r, and both are self-adjoint, thus

$$\dim C(\mathbf{H}^*\mathbf{H}) = C(\mathbf{HH}^*) = \dim C(\mathbf{H}) = r.$$

Considering $\tilde{\mathbf{D}} = \mathbf{DD} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r)$ both can be unitarily diagonalized as follows.

$$\mathbf{H}^*\mathbf{H} = \mathbf{V}\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_r \end{bmatrix}\mathbf{V}^* \quad \text{and} \quad \mathbf{HH}^* = \mathbf{U}\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_r \end{bmatrix}\mathbf{U}^*$$

with the same eigenvalues and possibly different unitary matrices $\mathbf{U}$ *and* $\mathbf{V}$, it is obtained that:

$$\mathbf{V}^*\mathbf{H}^*\mathbf{HV} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_r \end{bmatrix} = \mathbf{UHH}^*\mathbf{U}^*$$

Thus,

$$\mathbf{H}^*\mathbf{H} = \mathbf{V}^*\mathbf{UHH}^*\mathbf{U}^*\mathbf{V}$$

with $\mathbf{V}^*\mathbf{U}$ and $\mathbf{U}^*\mathbf{V}$ being unitary matrices, if we move one of these unitary matrices to the right-hand side, it is obtained that:

$$\mathbf{V}^*\mathbf{UH}^*\mathbf{H} = \mathbf{V}^*\mathbf{UHH}^* \rightarrow \mathbf{H}^*\mathbf{H} = \mathbf{HH}^*$$

($\leftarrow$)
Let $\mathbf{HH}^* = \mathbf{U}\tilde{\mathbf{D}}\mathbf{U}^*$ with $\mathbf{U}$ being a unitary matrix that diagonalizes $\mathbf{HH}^*$. Based on (4), it is obtained that $\mathbf{H}^*\mathbf{U} = \mathbf{V}\tilde{\mathbf{D}}$. If $\mathbf{D} = \frac{1}{\sqrt{\lambda_i}}$ is considered and

$$v_i = \frac{1}{\sqrt{\lambda_i}}\mathbf{H}^*u_i$$

is chosen, it will be checked whether v_i, v_j *are orthogonal for* $i \neq j$. It should be noted that:

$$\langle v_i, v_j \rangle = v_j^* v_i = \frac{1}{\sqrt{\lambda_j}}\frac{1}{\sqrt{\lambda_i}}u_j^*\mathbf{HH}^*u_i$$

Since

$$\mathbf{HH}^* = \mathbf{U}\tilde{\mathbf{D}}\mathbf{U}^* \rightarrow \mathbf{HH}^*u_i = u_i\lambda_i.$$

Additionally,

$$\langle v_i, v_j \rangle = \frac{1}{\sqrt{\lambda_j}}\frac{1}{\sqrt{\lambda_i}}u_j^*\mathbf{HH}^*u_i = \frac{\sqrt{\lambda_i}}{\sqrt{\lambda_j}}\langle u_i, u_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

based on the unitary property of $\mathbf{U}$. Since $\{v_1, v_2, \ldots, v_r\}$ forms an orthonormal set, a unitary matrix

$$\mathbf{V} = [v_1 \quad v_2 \quad \ldots \quad v_r \quad v_{r+1} \quad \ldots \quad v_W]$$

can be constructed. Thus, for a matrix $\mathbf{H} \in \mathbb{C}^{L\times W}$, a diagonal matrix $\mathbf{D} \in \mathbb{R}^{r\times r}$ a unitary matrix $\mathbf{U} \in \mathbb{C}^{L\times L}$ and $\mathbf{V} \in \mathbb{C}^{W\times W}$ exist. □

Referring to **Lemma 2.2**, **Lemma 2.3** *and* 2.4, there are general consequences that apply to the $\mathbf{H}$ as follows:

Corollary 2.5. *The bounds for* $L; W$ *in* $\mathbf{H} \in \mathbb{C}^{L\times W}$ *are*

$$2 \leq L; W \leq \frac{T}{2}.$$

The commonly used value of $L; W$ is often about half the length of the time series, namely $L; W \approx T/2$. However, in practice, for very large T, it is believed that selecting $L; W$ according to **Corollary 2.5** is not effective. Therefore, to explore the possibility of an optimal limit $L; W$ on time series data, especially with the AR(1) and AR(2) model through a sequence of singular values on SVD, the following procedure is developed for further analysis.

3. EXPERIMENTAL PROCEDURE

The simulation process of AR(1) and AR(2) model is executed through an algorithm designed to generate time series data closely resembling the actual AR process. The generated simulation data undergoes analysis using the SVD technique, which decomposes the data matrix into fundamental components, facilitating insights into the intrinsic structure of the data. The entire simulation process is implemented using the R programming language, structured into three procedural components.

3.1. **Monte Carlo Simulation to Construct AR(1) and AR(2) model.**

First, initialization is performed by setting a seed for data replication to ensure consistency in repeated experiments. The number of data values used in the simulation is defined as T and the AR(1) and AR(2) parameter value (ϕ_i) is assigned to the model. In addition, the standard deviation of the white noise (σ) added to the model is determined and the number of repetitions (M) of the Monte Carlo experiment is specified.

In this procedure, $T = 100$, $\sigma = 1$ are used, and M is set to 1500 repetitions. ϕ_1 is varied from $0.1 \leq |\phi_1| \leq 0.9$ for the AR(1) model, and several choices of ϕ_1 and ϕ_2 are considered for the AR(2) model according to its stationarity conditions. During the simulation process, Z_t is generated using the model $Z_t = \phi_1 Z_{t-1} + e_t$ for AR(1) and $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + e_t$ for AR(2) with $e_t \sim \mathcal{N}(0, 1)$ for each time point $t = 1, 2, \ldots, T$. This process is repeated M times to obtain a series $Z_t^{(m)}$ for

each repetition $m = 1, 2, \ldots, M$. The mean at each time point is calculated from all repetitions, i.e.,

$$\bar{Z}_t = \frac{1}{M} \sum_{m=1}^{M} Z_t^{(m)}$$

for $t = 1, 2, \ldots, T$. The output of this procedure includes $Z_t^{(m)}$ for $t = 1, 2, \ldots, T$ and each time repetition $m = 1, 2, \ldots, M$ (each m depicted with different color), along with $\bar{Z}_t$ (depicted with dashed red lines) for $M = 1500$ in the example of $\phi_1 = -0.9$ as illustrated in Figure 1. This process accommodates randomness and produces data centered around the mean with $\mathbb{E}(\bar{Z}_t) = 0$. The resulting data is then structured into $\mathbf{H}$, which serves as a crucial step for subsequent analysis

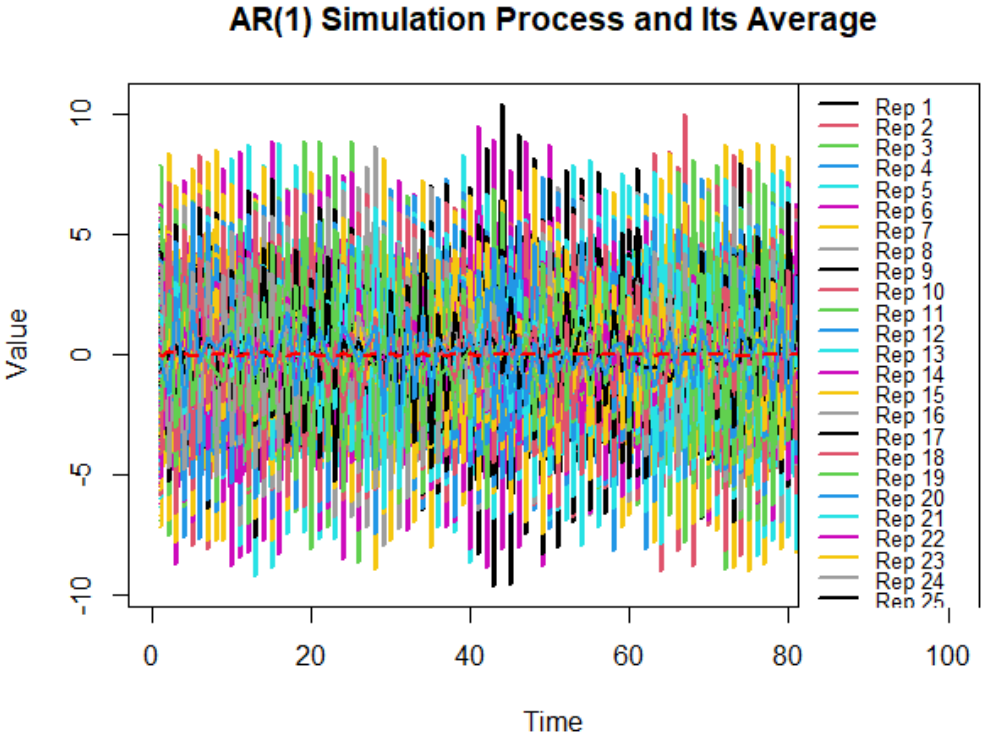

FIGURE 1. Illustration $z_t^{(m)}$ for $t = 1, 2, \ldots, T$ and $m = 1, 2, \ldots, M$ (each m depicted with different color), along with $\bar{z}_t$ for $t = 1, 2, \ldots, T$ (depicted with dashed red lines) for $M = 1500$ in the example of $\hat{\phi}_1 = -0.9$. The data $z_t^{(m)}$ will then be structured into $\mathbf{H}$, which serves as the basis for further analysis.

The AR(1) model simulation algorithm produces 18 data models with ϕ_i presented in Table 1. Meanwhile, 4 simulated data from the AR(2) model, with ϕ_i shown in Table 2, will be analyzed. Each of these AR(2) simulation data will also be referred to as Data A, B, C, and D. These results show how the chosen ϕ_i influence the time series characteristics from the simulation.

TABLE 1. The estimated parameter values for the AR(1) model with $0.1 \leq |\phi_1| \leq 0.9$.

ϕ_1	$\hat{\phi}_0$	$\hat{\phi}_1$	ϕ_1	$\hat{\phi}_0$	$\hat{\phi}_1$
-0.9	-0.0007	-0.9143	0.9	0.0258	0.9039
-0.8	-0.0077	-0.8049	0.8	0.0216	0.8128
-0.7	-0.0018	-0.7111	0.7	0.0179	0.7092
-0.6	-0.0006	-0.5986	0.6	0.0105	0.5120
-0.5	-0.0033	-0.5092	0.5	0.0096	0.4986
-0.4	-0.0005	-0.4108	0.4	0.0033	0.3946
-0.3	-0.0026	-0.3108	0.3	0.0023	0.2919
-0.1	0.0042	-0.0921	0.1	0.0021	0.0917

TABLE 2. The estimated parameter values $\hat{\phi}_0, \hat{\phi}_1$, and $\hat{\phi}_2$ for the AR(2) model with $\phi_1 = \pm 0.5$ and $\phi_2 = \pm 0.4$.

Data	ϕ_1	ϕ_2	$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\phi}_2$
A	0.5	0.4	-0.0006	0.5584	0.3180
B	-0.5	-0.4	0.0014	-0.5093	-0.4814
C	0.5	-0.4	-0.0075	0.4854	-0.4133
D	-0.5	0.4	-0.0031	-0.5059	0.3380

3.2. **Descriptive Statistics and Testing of the Model.**

This procedure is designed to ensure that the constructed model meets the desired assumptions, including stationarity, normal distribution, and the absence of autocorrelation among residuals. Consequently, the model can be considered appropriate for the data as long as all residual assumptions are satisfied. Additionally, statistical tests are conducted to ensure that the constructed model is acceptable according to the data, as no residual assumptions are violated. Furthermore, by examining the correlogram of ACF and PACF, autocorrelation behavior can be investigated.

3.3. **SVD and Its Convergence.**

To initiate this analysis, $\{\bar{z}_t\}$ is converted into a vector form. This sequence is then arranged into a matrix $\mathbf{H}$ and the SVD is then constructed in Eq. (1) where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and $\mathbf{\Sigma}$ contains $\{\lambda_i^*\}$. According to **Corollary 2.5**,

the λ_i^* remain unchanged for $2 \leq L \leq 50$ and $51 \leq L \leq 99$. Therefore, setting the upper limit of L to 50 suffices, λ_i^* can be arranged in descending order as

$$\lambda_1^* \geq \lambda_2^* \geq \cdots \geq \lambda_{50}^*$$

in the form of a scree plot graph. The cumulative proportions of each λ_i^* is calculated as

$$\%\lambda_i^* = \frac{\lambda_i^*}{\sum \lambda_i^*}$$

providing insights into the contribution of each λ_i^* to the overall data structure. Scree plots of $\{\lambda_i^*\}$ and $\{\%\lambda_i^*\}$ are observed to assess the significance of each λ_i^* and to identify any convergence patterns.

Identifying convergence in$\{\lambda_i^*\}$ is used to determine the optimal bound of window length (L_ε) in simulated data constructed from an AR model. Let $\underline{L}$ be the lower limit of L and $\bar{L}$ the upper limit of L, then the difference in the $\%\lambda_i^*$ is expressed as

$$|\Delta_s| = |\%\lambda_i^* - \%\lambda_s^*| \leq \varepsilon$$

where $i = \underline{L}, \ldots, \bar{L} - 1$ and $s = i + 1, \ldots, \bar{L}$. By choosing ε, determining L_ε for the case of simulated AR(1) model data can be made more efficient.

4. NUMERICAL RESULT

The autocorrelation behavior and $\{\lambda_i^*\}$ from SVD in the simulated data model are analyzed numerically through ACF and PACF. Figure 2 and Figure 3 illustrate the correlogram patterns for AR(1) and AR(2) models, respectively. The ACF plot exhibits a mix of damped exponential and sinusoidal waves, forming an organized pattern. However, ACF provides limited insight into autocorrelation across lag intervals and is more useful for identifying the MA model order 'q'. In contrast, PACF helps determine the AR model order 'p'.

PACF reveals partial correlations after accounting for shorter time intervals. In an AR model, PACF typically cuts off beyond a certain lag. For simulated AR(1) data, this cutoff occurs after lag 1 unless $|\widehat{\phi}_1| \sim 0$. Ensuring data accuracy becomes challenging as past values heavily influence predictions. Thus, alternative models or external factors beyond the AR(1) model may need further investigation.

The ACF and PACF in the AR(2) model simulation data is analyzed to observe correlogram patterns with varying values of $\hat{\phi}_1$ and $\hat{\phi}_2$, rather than correlating it with parameter magnitudes. When $\hat{\phi}_1$ and $\hat{\phi}_2$ are both positive (Data A), the ACF pattern is more regular and systematic, starting with positive autocorrelation followed by negative values, repeating in a consistent manner. Conversely, when both parameters are negative (Data B), the ACF pattern alternates between positive and negative values. Further examination shows that when $\hat{\phi}_1$ is positive and $\hat{\phi}_2$ is negative (Data C), the ACF pattern resembles Data A. On the other hand, when $\hat{\phi}_1$ is negative and $\hat{\phi}_2$ is positive (Data D), the ACF pattern resembles Data B

but is more systematic. The PACF correlogram, however, shows a cut-off pattern after lag 2, corresponding to the signs of $\hat{\phi}_1$ and $\hat{\phi}_2$.

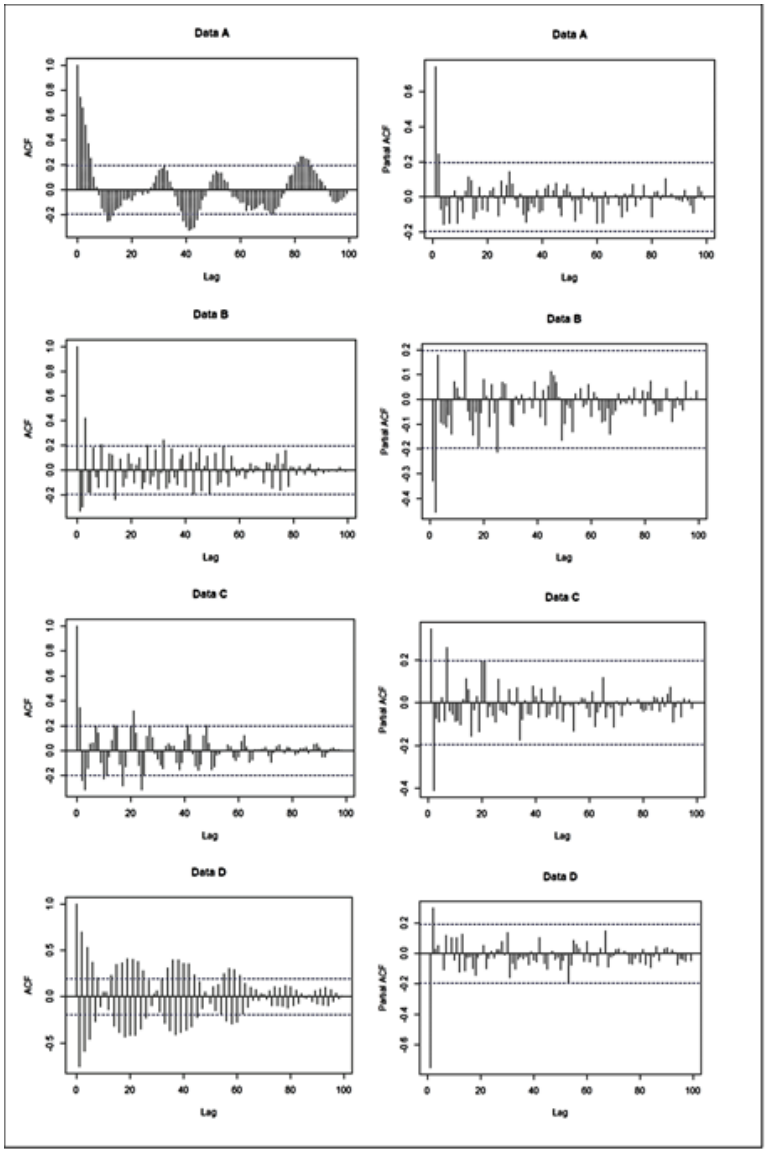

FIGURE 2. Correlogram illustration using ACF and PACF for several simulated AR(1) data, highlighting significant autocorrelation patterns to be identified. The PACF plot serves as a reference for comparison with the scree plot SVD of $\{\lambda_i^*\}$ obtained from $\mathbf{H}$.

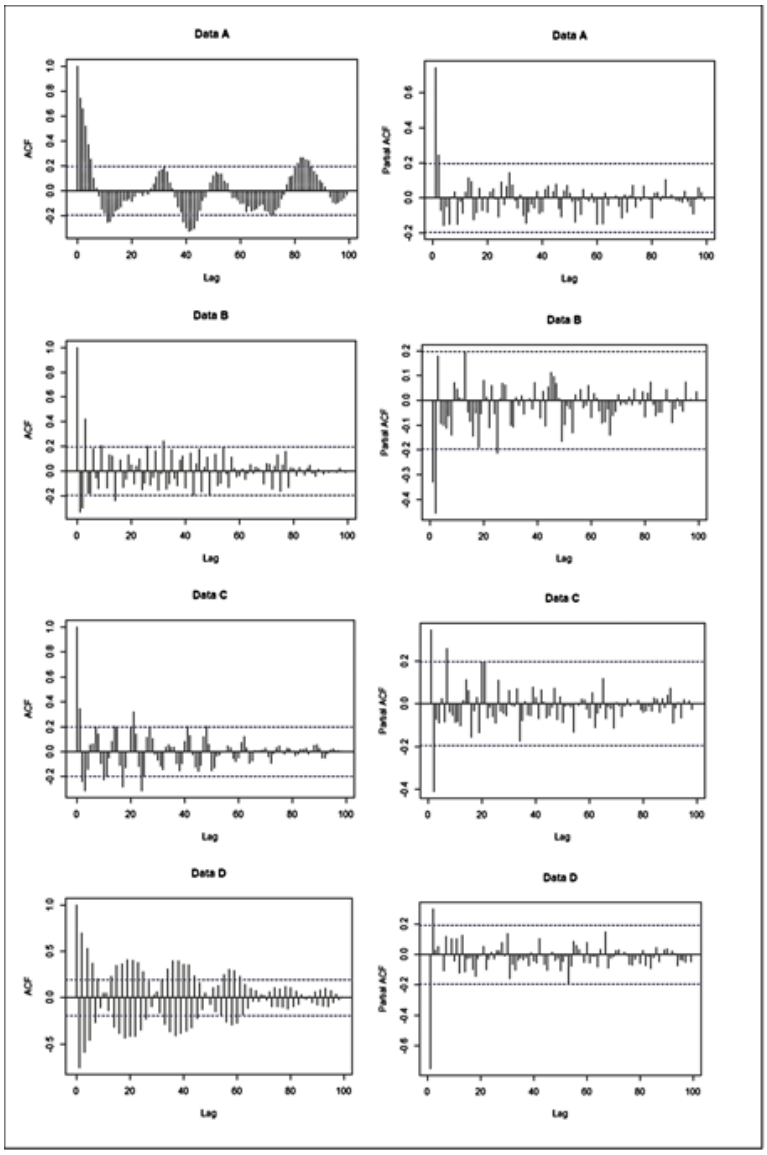

FIGURE 3. ACF (left) and PACF (right) plots for simulated AR(2) model data with parameters $\phi_1 = \pm 0.5$ and $\phi_2 = \pm 0.4$ (Data A–D on Table 2). The PACF plot serves as a reference for comparison with the scree plot SVD of $\{\lambda_i^*\}$ obtained from $\mathbf{H}$.

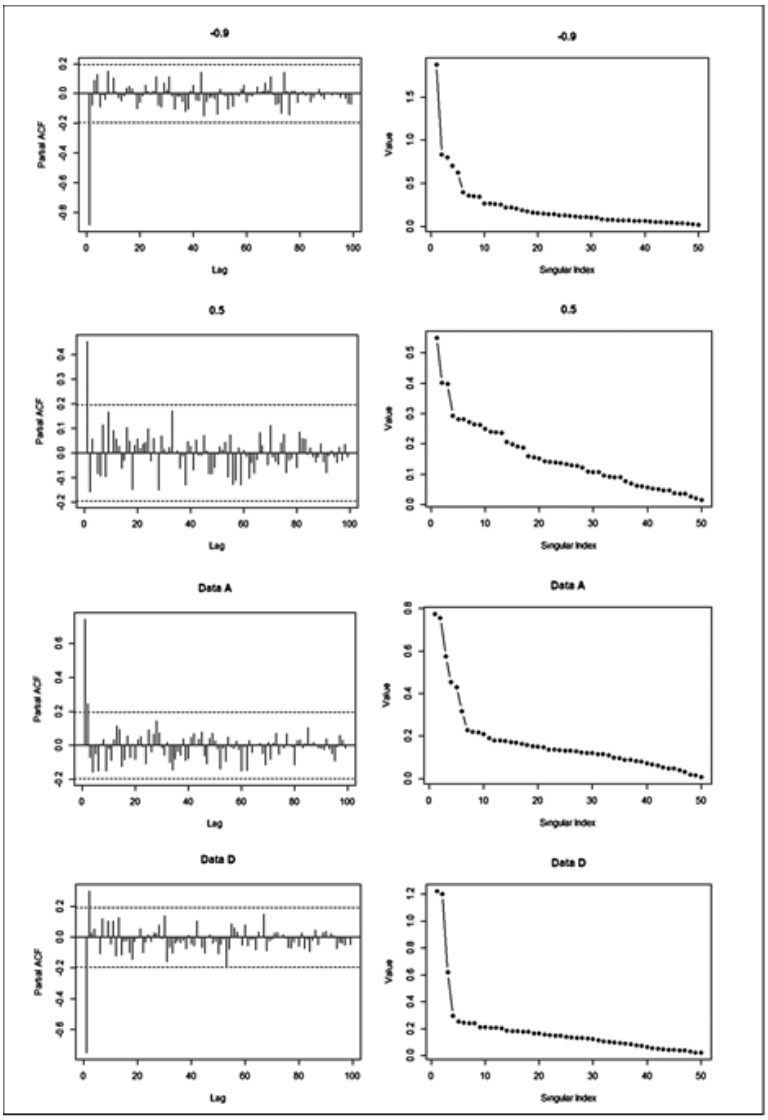

FIGURE 4. PACF of $\bar{z}_t$ (left) and scree plots SVD of $\{\lambda_i^*\}$ obtained from **H** (right) for simulated AR(1) and AR(2) model data show similar patterns

Interestingly, the PACF patterns in the simulated AR(1) and AR(2) models show a cut-off pattern similar to the scree plot pattern observed in the SVD shown in Figure 4. In the AR(1) model, the PACF cut-off pattern resembles the 'elbow' in the SVD scree plot, especially when $|\hat{\phi}_1| \sim 1$, the cut-off appears after the 1^{st} lag/λ_1^*. Conversely, when $|\hat{\phi}_1| \sim 0$, the variance contribution from the λ_i^* is more

evenly distributed, the 'elbow' becomes less distinct, indicating weaker and more dispersed correlations within the data. Thus, the data structure in the AR(1) model changes with different values, highlighting the importance of selecting an appropriate $\hat{\phi}_1$ value to accurately capture significant patterns in the AR(1) model.

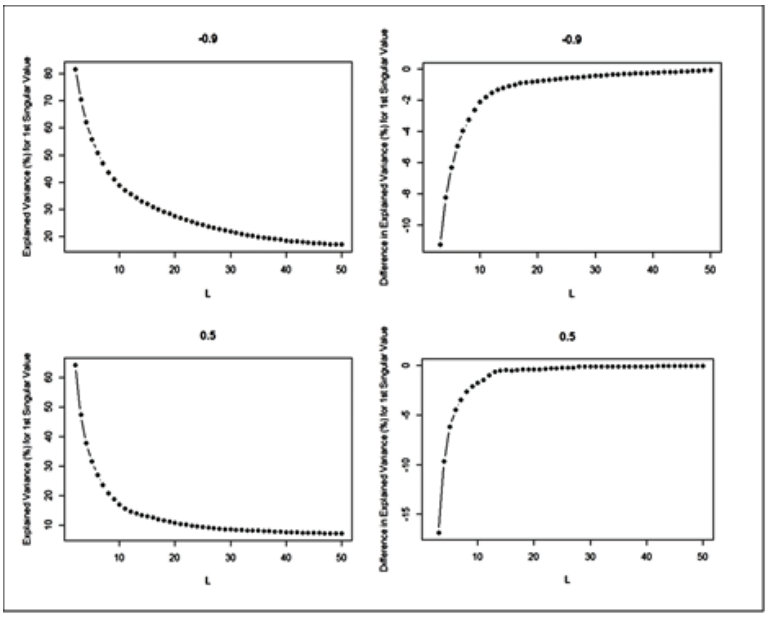

FIGURE 5. An illustration of $\%\lambda_1^*$ (left) and $|\Delta_s^{\lambda_1}|$ obtained from **H** in several AR(1) data model simulations for $2 \leq L \leq 50$. The plot reveals a convergence pattern at a specific value of L.

In the simulated AR(2) model data, λ_1^* and λ_2^* contribute the most significantly to explaining the variance in the data compared to other λ_i^*. This 'elbow' pattern is similar to the PACF pattern, which shows a cut-off after the 2^{nd} lag. The cut-off pattern observed in the PACF, which mirrors the pattern in the SVD, suggests that these methods can complement each other for a comprehensive view of TSA. Meanwhile, convergence is evaluated by determining if the $\%\lambda_i^*$ remains consistent or varies with the addition of more L. As L increases, there is a greater absorption of data information. The L_t represents the maximum number of L to be considered in the analysis. This helps in understanding the data structure and ensures that the SVD analysis captures enough information from the simulated AR(1) and AR(2) model data. Observations focus on changes in variance proportions ($|\Delta_s^{\lambda_i}|$) explained by $\%\lambda_i^*$, as shown in Figures 5 and 6.

In the AR(1) model, the focus will be on $|\Delta_s|$ for λ_1^* (Figure 5), while in the AR(2) model, the focus will be on λ_1^* and λ_2^* denoted by $|\Delta_s^{\lambda_1}|$ and $|\Delta_s^{\lambda_2}|$ respectively (Figure 6). The sequence $\{|\Delta_s|\}$ decreases monotonically with $|\Delta_{s+1}| \leq |\Delta_s|$ as

the size of L increases. This allows for the selection of ε such that $|\Delta_s| \leq \varepsilon$ for $L \geq L_\varepsilon$. By choosing ε, determining L_ε for the simulated data can be done more efficiently. If ε is chosen as 5% or 1%, the L_ε for each simulated AR(1) and AR(2) model data can be summarized as shown in Table 3 and Table 4.

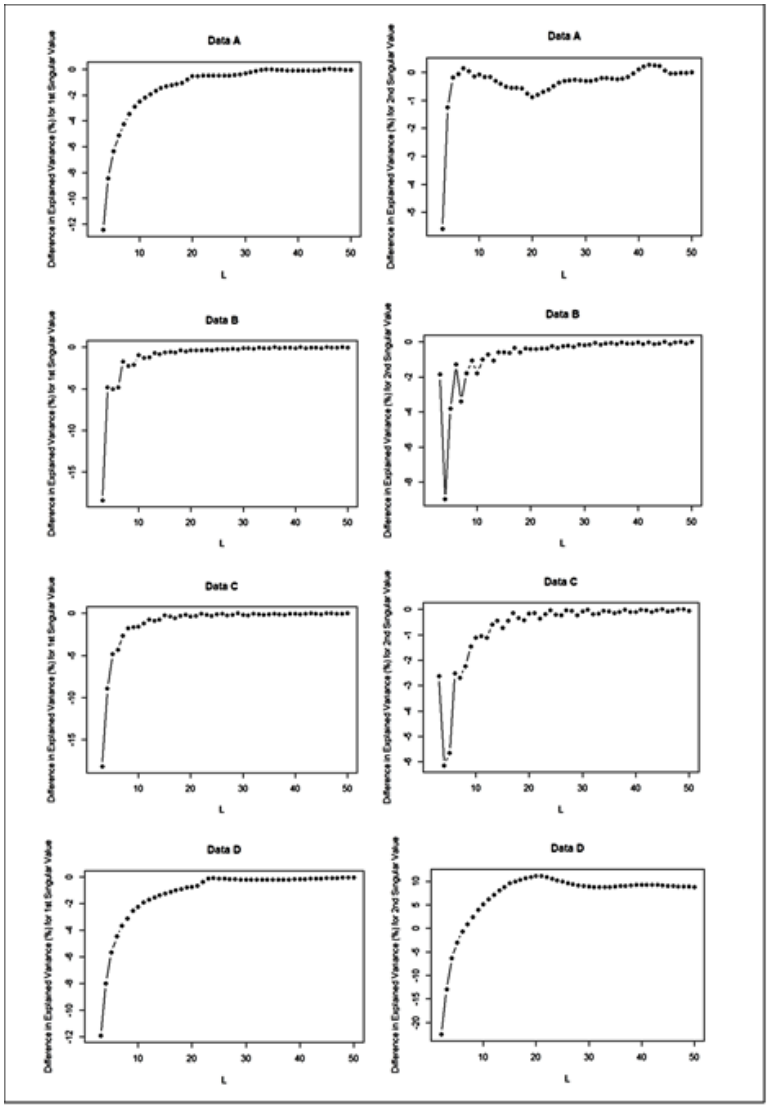

FIGURE 6. $|\Delta\lambda_1^s|$ (left) and $|\Delta\lambda_2^s|$ (right) represent the variance differences in the AR(2) data model simulations for $2 \leq L \leq 50$. These $\{\lambda_i^*\}$ are derived from $\mathbf{H}$, revealing structural variations within the simulated data.

TABLE 3. The optimal bound of window length (L_ε) for the AR(1) model fall within the range of $0.1 \leq \hat{\phi}_1 \leq 0.9$ for ε 5% and 1%.

Left Table			Right Table		
$\hat{\phi}_1$	5%	1%	$\hat{\phi}_1$	5%	1%
-0.9	**7 (-4.50)**	**18 (-0.99)**	0.9	7 (-4.50)	18 (-0.99)
-0.8	6 (-4.73)	16 (-0.98)	0.8	7 (-4.15)	18 (-0.98)
-0.7	6 (-4.87)	14 (-0.98)	0.7	7 (-4.12)	15 (-0.87)
-0.6	6 (-4.05)	13 (-0.86)	0.6	7 (-3.80)	15 (-0.91)
-0.5	6 (-4.21)	14 (-0.97)	0.5	**6 (-4.48)**	**13 (-0.62)**
-0.4	6 (-4.10)	12 (-0.99)	0.4	6 (-4.33)	12 (-0.95)
-0.3	6 (-4.20)	11 (-0.63)	0.3	6 (-3.96)	12 (-0.97)
-0.2	6 (-3.53)	12 (-0.48)	0.2	6 (-3.61)	12 (-0.99)
-0.1	6 (-3.59)	10 (-0.94)	0.1	6 (-2.78)	10 (-0.96)

TABLE 4. The optimal bound of window length (L_ε) for the simulated AR(2) model data based on $|\Delta_s^{\lambda_1}|$ (left) and $|\Delta_s^{\lambda_2}|$ (right) for ε values of 5% and 1%.

Data	$\Delta_s^{\lambda_1}$		$\Delta_s^{\lambda_2}$	
	5%	1%	5%	1%
A	**7 (-4.26)**	**19 (-0.81)**	**4 (-1.25)**	**5 (-0.19)**
B	6 (-4.81)	13 (-0.67)	5 (-3.81)	14 (-0.59)
C	5 (-4.81)	12 (-0.76)	6 (-2.52)	13 (-0.60)
D	6 (-4.48)	18 (-0.91)	4 (-1.54)	6 (-0.75)

Referring to Table 3, it can be inferred that to effectively use the SVD scree plot pattern in the matrix $\mathbf{H}$ for simulated AR (1) model data with $0.1 \leq |\hat{\phi}_1| \leq 0.9$, selecting L_ε 6 or 7 for $\varepsilon = 5\%$ and L_ε between 10 and 18 for $\varepsilon = 1\%$ would be appropriate. For example, in the case of the AR(1) model with $\hat{\phi}_1 = -0.9$, if $\varepsilon = 5\%$ is chosen, the optimal bound L_ε is 7 with $|\Delta_s^{\lambda_1}| = 4.50$. Meanwhile, for $\varepsilon = 1\%$, the optimal bound L_ε is 18 with $|\Delta_s^{\lambda_1}| = 0.99$. Similarly, for $\hat{\phi}_1 = 0.5$, if $\varepsilon = 5\%$ is chosen, the size L_ε is 6 with $|\Delta_s^{\lambda_1}| = 4.48$, and if $\varepsilon = 1\%$ is chosen, the size L_ε is 13 with $|\Delta_s^{\lambda_1}| = 0.62$.

In the simulation of the AR(2) data model, the plots of $|\Delta_s^{\lambda_1}|$ and $|\Delta_s^{\lambda_2}|$ based on Figure 6 exhibit similar patterns, both showing convergence to 0 as L approaches $\bar{L}$, with the sequence $\{|\Delta_s|\}$ decreasing monotonically. For example, referring to Table 4, in Data A, if ε is set at 5%, the optimal boundary size is $L_\varepsilon = 7$ with $|\Delta_s^{\lambda_1}| = 4.26$, and $L_\varepsilon = 4$ with $|\Delta_s^{\lambda_2}| = 1.2$. In contrast, when ε is set at 1%, the optimal boundary size is $L_\varepsilon = 19$ with $|\Delta_s^{\lambda_1}| = 0.81$, and $L_\varepsilon = 5$ with $|\Delta_s^{\lambda_2}| = 0.19$.

5. RESULT AND DISCUSSION

Matrix-based analysis becomes particularly relevant when applied to established TSA methods such as the Box-Jenkins approach, particularly within the AR model framework. In AR(1) and AR(2) models, the PACF exhibits a characteristic cut-off after lag 1 for AR(1) and lag 2 for AR(2), indicating that PACF values remain significant up to the corresponding model order before declining sharply toward zero. This pattern closely resembles the scree plot in SVD, where the largest singular values decrease markedly after the dominant components, forming an "elbow" shape. In SVD, only the first few singular values retain most of the information, analogous to how PACF retains significant values only up to the AR model order.

This similarity suggests that PACF and the SVD scree plot can complement each other in identifying key structures in time-series data. PACF helps determine the order of an AR model by identifying where the partial correlations are cut off, while the scree plot in SVD reveals the number of significant components in the decomposition of the data structure. Furthermore, SVD sensitivity to AR parameters depends on the magnitude of $(\hat{\phi}_1/\hat{\phi}_2)$ due to stationarity requirements, but it is not affected by the sign of the parameter (+/-). This can be explained by considering that stationarity depends on the modulus of the characteristic polynomial roots of the AR model rather than the individual coefficient signs.

Other findings indicate that the lower and upper limits of **Corollary 2.5** are not effective in the adhesion process of HM (**H**), which requires an analytical study of the optimal boundary L_ε that improves efficiency by selecting a specific ε. It is hypothesized that an optimal limit of $L_\varepsilon = 20$ may be sufficient for time series data. The numerical results obtained in this study suggest promising directions for future research by revealing potential patterns and relationships in time series data. While these findings provide valuable insights, reinforcing them with analytical proofs would enhance their theoretical foundation. Therefore, future studies could focus on developing formal mathematical validations to complement these results. By integrating the SVD approach with trajectory matrices into existing TSA methods, this framework can evolve to be more adaptive and objective, strengthening its applicability in TSA.

Acknowledgement. This research was supported by the Statistics Research Division, Faculty of Mathematics and Natural Science, ITB. The author also would like to thank LPDP for giving the scholarship.

REFERENCES

[1] C. E. Montagnon, "Forecasting by splitting a time series using singular value decomposition then using both ARMA and a Fokker-Planck equation," *Physica A: Statistical Mechanics and*

its Applications, vol. 567, p. 125708, 2021. https://www.doi.org/10.1016/j.physa.2021.125708.

[2] J. Gillard and K. Usevich, "Hankel low-rank approximation and completion in time series analysis and forecasting: a brief review," *Statistics and Its Interface*, vol. 6, no. 1, pp. 123–145, 2022. https://doi.org/10.48550/arXiv.2206.05103.

[3] R. A. Horn and C. R. Johnson, *Matrix analysis*. New York: Cambridge University Press, 2nd ed., 2013.

[4] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of time series structure: SSA and related techniques*. London: Chapman & Hall/CRC, 2001.

[5] N. Golyandina and A. Zhigljavsky, *Singular spectrum analysis for time series*. Springer, 2013.

[6] N. Golyandina, "Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing," *WIREs Computational Statistics*, vol. 12, no. 3, p. e1483, 2020. https://doi.org/10.1002/wics.1487.

[7] N. Golyandina and A. Korobeynikov, "Basic singular spectrum analysis and forecasting with R," *Computational Statistics and Data Analysis*, vol. 64, pp. 92–107, 2013. https://doi.org/10.1016/j.csda.2013.04.009.

[8] N. Golyandina, A. Korobeynikov, and A. Zhigljavsky, *Singular spectrum analysis with R*. Springer, 2018. https://doi.org/10.1007/978-3-662-57380-8.

[9] C. Deng, *Time series decomposition using singular spectrum analysis*. Ph.d. dissertation, East Tennessee State University, 2014. https://dc.etsu.edu/etd/2352.

[10] S. Wadekar, A. Mahalkari, A. Ali, and A. Gupta, "A review on singular spectrum analysis," in *Proceedings of the IEEE International Conference on Current Development in Engineering and Technology (CCET)*, 2022. https://doi.org/10.1109/CCET56606.2022.10080082.

[11] R. Kazemi and P. Rodrigues, "Robust singular spectrum analysis: comparison between classical and robust approaches for model fit and forecasting," *Computational Statistics*, vol. 38, no. 1, pp. 87–110, 2023. https://doi.org/10.1007/s00180-022-01322-4.

[12] K. E. Lestari, U. S. Pasaribu, and S. W. Indratno, "Graphical depiction of three-way association in contingency table using higher-order singular value decomposition Tucker3," *Journal of Physics: Conference Series*, vol. 1280, p. 022035, 2019. .

[13] K. E. Lestari, *Structure of contingency table using Cardano and Cardano-Ferrari formulas on correspondence analysis*. Ph.d. dissertation, Institut Teknologi Bandung, 2021.

[14] S. Banerjee and A. Roy, *Linear algebra and matrix analysis for statistics*. CRC Press Taylor and Francis Group, 2014.

[15] G. M. Djakou and X. Jiang, "GDP modelling and forecasting using ARIMA: an empirical study for cameroon," *International Journal of Science and Business*, vol. 22, pp. 41–52, 2023. https://ideas.repec.org/a/aif/journl/v22y2023i1p41-52.html.

[16] F. M. Razali and N. F. Haron, "Forecasting data using box-jenkins procedure: a case study for unemployed people in malaysia," *Gading Journal of Science and Technology*, vol. 6, pp. 31–39, 2023. https://ir.uitm.edu.my/id/eprint/114699/.

[17] T. A. Ojurongbe, T. A. Bello, N. O. Adeboye, H. A. Afolabi, O. O. Aduroja, and K. A. Bashiru, "Time series analysis on the prevalence of malaria in osogbo, nigeria," *Adeleke University Journal of Engineering and Technology (AUTJET)*, vol. 6, pp. 85–95, 2023. http://aujet.adelekeuniversity.edu.ng/index.php/aujet/article/view/289.

[18] T. Makoni, G. Mazuruse, and B. Nyagadza, "International tourist arrivals modelling and forecasting: a case of zimbabwe," *Sustainable Technology and Entrepreneurship*, vol. 2, no. 1, p. 100027, 2023. https://www.doi.org/10.1016/j.stae.2022.100027.

[19] K. A. Kumar, P. Pinto, C. Spulbar, R. Birau, I. T. Hawaldar, S. Vishal, and I. C. Barbacioru, "Arima model to forecast the rsa-1 rubber price in india: a case study for textile industry," *Industria Textila*, vol. 74, 2023. 1https://www.doi.org/10.35530/IT.074.02.2022132.

[20] G. Bezabih, M. Wale, N. Satheesh, S. W. Fanta, and M. Atlabachew, "Forecasting cereal crops production using time series analysis in ethiopia," *Journal of the Saudi Society of Agricultural Sciences*, vol. 22, no. 8, pp. 546–559, 2023. https://www.doi.org/10.1016/j.jssas.2023.07.001.

[21] S. G. Makridakis, S. C. Wheelwright, and V. E. McGee, *Forecasting: Methods and Applications*. New York: Wiley, 2nd ed., 1983.
[22] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day Inc, 1976.
[23] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley and Sons, Inc, 5th ed., 2016.
[24] W. W. S. Wei, *Time series analysis: univariate and multivariate methods*. Pearson-Addison Wesley, 2nd ed., 2006.
[25] J. D. Cryer and K. S. Chan, *Time series analysis with applications in R*. Springer, 2nd ed., 2008. https://link.springer.com/chapter/10.1007/978-0-387-75959-3_15.
[26] W. Palma, *Time series analysis*. John Wiley and Sons Inc, 2016.
[27] M. S. Paolella, *Linear models and time series analysis: regression, ANOVA, ARMA and GARCH*. John Wiley and Sons Inc, 2019.
[28] F. Kamalov, F. Thabtah, and I. Gurrib, "Autocorrelation for time series with linear trend," in *International Conference on Innovation and Intelligence for Informatics, Computing and Technology (3ICT)*, 2021. https://doi.org/10.1109/3ICT53449.2021.9581809.
[29] X. Tan, *Penalized estimation of autocorrelation*. Ph.d. dissertation, Clemson University, 2022.
[30] C. H. Weib, B. Aleksandrov, M. Faymonville, and C. Jentsch, "Partial autocorrelation diagnostics for count time series," *Entropy*, vol. 25, 2023. https://doi.org/10.3390/e25010105.
[31] P. Kwizera, *Matrix singular value decomposition*. Ph.d. dissertation, University of North Florida, 2010.
[32] P. Frame and A. Towne, "Space-time POD and the hankel matrix," *PLoS ONE*, vol. 18, no. 7, p. e0286234, 2023. https://doi.org/10.1371/journal.pone.0289637.