mathematics

MDPI

Article

Reduced Order Data-Driven Twin Models for Nonlinear PDEs by Randomized Koopman Orthogonal Decomposition and Explainable Deep Learning

Diana Alina Bistrian

Department of Electrical Engineering and Industrial Informatics, University Politehnica Timisoara, 300006 Timisoara, Romania; diana.bistrian@upt.ro

Abstract

This study introduces a data-driven twin modeling framework based on modern Koopman operator theory, offering a significant advancement over classical modal decomposition by accurately capturing nonlinear dynamics with reduced complexity and no manual parameter adjustment. The method integrates a novel algorithm with Pareto front analysis to construct a compact, high-fidelity reduced-order model that balances accuracy and efficiency. An explainable NLARX deep learning framework enables real-time, adaptive calibration and prediction, while a key innovation—computing orthogonal Koopman modes via randomized orthogonal projections—ensures optimal data representation. This approach for data-driven twin modeling is fully self-consistent, avoiding heuristic choices and enhancing interpretability through integrated explainable learning techniques. The proposed method is demonstrated on shock wave phenomena using three experiments of increasing complexity accompanied by a qualitative analysis of the resulting data-driven twin models.

Keywords: data-driven twin model; Koopman randomized orthogonal decomposition; nonlinear PDE; explainable deep learning

MSC: 00A71; 46N40; 34A45

check for updates

Academic Editors: Leonid Piterbarg and Carlo Bianca

Received: 7 July 2025
Revised: 8 August 2025
Accepted: 2 September 2025
Published: 5 September 2025

Citation: Bistrian, D.A. Reduced Order Data-Driven Twin Models for Nonlinear PDEs by Randomized Koopman Orthogonal Decomposition and Explainable Deep Learning. *Mathematics* **2025**, *13*, 2870. https://doi.org/10.3390/math13172870

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, digital twin models have emerged as powerful tools for understanding, simulating, and predicting the behavior of complex physical systems across various domains [1–4]. Despite their success in engineered systems with well-defined models and sensing infrastructures, the application of digital twins to nonlinear dynamical systems remains limited due to the inherent complexity of such dynamics. Moreover, conventional approaches often rely on full physical modeling and high computational demands, constraining their scalability and suitability for data-driven or partially observed systems.

To address key limitations in modeling nonlinear dynamical systems, this work introduces a novel theoretical formulation and an algorithmic framework within the emerging domain of data-driven twin modeling. This new paradigm is inspired by the digital twin concept but diverges fundamentally in its construction and purpose: rather than embedding a full physics-based model, the data-driven twin model is designed to learn and replicate the nonlinear behavior of a complex dynamical system directly from data, achieving high fidelity with significantly reduced model complexity.

At the core of this contribution is a new mathematically rigorous framework that decouples system evolution into stationary and dynamic components using the principles of modal decomposition [5]. Based on this theoretical foundation, this research presents the first mathematical introduction of a data-driven twin model (DTM) tailored with a novel algorithm, which is specifically formulated for unsupervised nonlinear dynamical systems.

The proposed framework defines the DTM as a reduced-order model capable of faithfully mirroring the full system behavior with significantly reduced computational complexity. This approach addresses a critical gap in existing modeling approaches, specifically the integration of real-time adaptability with the accurate simulation of nonlinear phenomena.

1.1. Modal Decomposition Literature Review and Previous Work

Modal decomposition [5] represents a generic mathematical framework for the characterization of nonlinear dynamical systems in the context of reduced-order modeling [6–12].

Proper Orthogonal Decomposition (POD) and Dynamic Mode Decomposition (DMD) are widely used modal decomposition methods for reducing computational complexity in modeling complex dynamical systems. POD, rooted in the Karhunen–Loève expansion independently developed by Loève [13] and Karhunen [14], was adapted to fluid mechanics by Lumley [15] to extract dominant coherent structures from turbulence. Since then, POD has been applied across turbulent and convection-dominated flows [16–19], engineering, and oceanography [20–22], often combined with data assimilation [23], control [24] and machine learning, including neural networks [25–29]. However, POD's energy-based truncation can miss low-energy but crucial modes, reducing accuracy in complex or nonlinear systems [30,31], and its projection nature may fail to capture fine-scale or transient dynamics absent from training data, especially in long-term or highly nonlinear scenarios.

Schmid and Sesterhenn [32,33] introduced Dynamic Mode Decomposition (DMD), a numerical technique based on Koopman operator theory [34,35], which has been widely applied across science and engineering [6,10,36,37]. J. Nathan Kutz's group significantly advanced this area by developing robust Koopman-based, data-driven modeling, and system identification methods [8,38–41]. Many DMD variants have since emerged to enhance robustness and applicability, including optimized [42], exact [38], sparsity-promoting [43], multi-resolution [40], extended [44], recursive [45], control-enabled [46], randomized low-rank [47], core sketch [48], bilinear [49], and higher-order DMD [50].

DMD captures spatiotemporal dynamics through modes with distinct frequencies and growth rates; however, its accuracy depends on proper mode selection. Identifying relevant modes often requires additional criteria involving manual tuning, domain expertise, or heuristics, increasing computational cost and limiting its use in fully unsupervised settings.

Koopman operator theory, coupled with DMD, has become a foundational tool in contemporary dynamical systems analysis and is widely regarded as a key mathematical framework of the 21st century. The author has notably advanced Koopman-based algorithms, particularly addressing challenges posed by high-dimensional data and complex temporal dynamics.

Significant contributions include the development of mode selection strategies that leverage amplitude-weighted growth rates and energy–Strouhal number criteria to systematically identify the most dynamically relevant modes [51,52]. To overcome the computational challenges of standard DMD, the author introduced the adaptive randomized dynamic mode decomposition (ARDMD) [53], enhancing efficiency and scalability in reduced-order modeling, especially within fluid dynamics applications [54].

The techniques developed in these prior studies tackled the enduring challenge of identifying dominant modes in nonlinear and unsteady systems, which is essential for precise reconstruction and forecasting. Additionally, the author has expanded the application

of DMD into emerging fields such as epidemiological modeling [55,56], demonstrating its wider societal impact.

1.2. Problem Statement and Novelty of the Present Research

Acknowledging the inherent limitations of classical DMD, namely, the non-orthogonality of its modes and the substantial computational overhead linked to manual mode selection, as well as the constraints of POD, which, despite yielding orthogonal modes, still depends on heuristic or user-specified criteria for mode selection, the proposed work introduces a novel approach for constructing a data-driven twin model that accurately replicates the full dynamics of nonlinear systems while significantly reducing model complexity.

Grounded in modern Koopman operator theory, the proposed framework offers a self-consistent, interpretable, and scalable solution that eliminates manual intervention and is well suited for unsupervised learning. By integrating explainable deep learning within a rigorous mathematical setting, it ensures transparency in the modeling process, an essential feature absent in conventional black-box machine learning models.

This novel framework significantly advances Koopman-based modal decomposition. Its key contributions include the following:

1. A fully self-consistent algorithm incorporating Pareto front analysis to eliminate heuristic parameter tuning;
2. The integration of explainable deep learning techniques to enable transparent, adaptive, and interpretable modeling;
3. A compact, high-fidelity data-driven twin model balancing accuracy and computational efficiency.

The structure of the article is organized as follows: Section 2 outlines the mathematical foundations of reduced-order modeling grounded in Koopman operator theory. Section 3 addresses the computational aspects and introduces a novel modal decomposition algorithm, which is supported by a rigorous mathematical justification. Section 4 applies the proposed method to shock wave phenomena through three experiments of increasing complexity, while Section 5 presents the corresponding numerical results. Section 6 concludes with a summary of key findings and potential directions for future research. Appendix A provides the full derivation of the analytical and numerical solutions, and Appendix B includes the pseudocode of the proposed algorithm for reproducibility and implementation. The key mathematical notations used in this article are summarized in Table 1.

Table 1. Key notations used throughout the article.

Mathematical Notation	Description
$\mathcal{K}^t$	Koopman operator at time t, indicating its dependence on the temporal variable in continuous-time dynamical systems.
$\{\mathcal{K}^t\}_{t\geq 0}$	The semigroup of Koopman operators.
$\mathcal{K}$	Koopman operator for discrete-time dynamical system.
N_x	Number of spatial grid points.
N_t	Number of time steps or snapshots, excluding the initial state.
$\|\cdot\|$	The norm for functions in the Hilbert space $L^2(\Omega)$.
$\|\cdot\|_2$	The Euclidean (or ℓ^2) norm for vectors.
$u_i = u(x, t_i)$, $i = 1, \ldots, N_t$	Measurements of the PDE solution sampled uniformly in time, with step size Δt.
$u_i^{(k)} = u^{(k)}(x, t_i)$	Reduced-order approximation of the PDE solution at time t_i, using the k-th order model.

Table 1. *Cont.*

Mathematical Notation	Description
N_{DTM}	Rank of the optimal Koopman basis, computed by the algorithm.
$\phi_j(x)$, $j = 1, \ldots, N_{DTM}$	Leading Koopman modes computed by the algorithm, representing dominant spatial patterns.
$a_j(t_i)$	The modal growing amplitudes.
$\hat{a}(t)$	The system temporal dynamics in the reduced-order space, modeled by the deep learning architecture.
$u^{DTM}(x,t)$	The reduced-order data-driven twin model for the PDE solution.
H	Conjugate transpose (Hermitian transpose); for a matrix A, $A^H = \overline{A}^T$.
$\langle \cdot, \cdot \rangle$	Inner product operator; for vectors $x, y \in \mathbb{C}^n$, $\langle x, y \rangle = x^H y$.
$X_{.,j}$	The jth column of the matrix X.
$u(x,t)$	Solution of the PDE, representing the system state at spatial location x and time t.

2. Mathematical Background

2.1. Computational Complexity Reduction of PDEs by Reduced-Order Modeling

Partial differential equations (PDEs) serve as a fundamental framework for accurately modeling nonlinear dynamical phenomena, owing to their capacity to represent spatial–temporal variations in complex systems. Nevertheless, the numerical solution of PDEs is often computationally demanding, particularly in high-dimensional and strongly nonlinear settings. This computational burden necessitates strategies for complexity reduction to enable practical simulation and analysis. In this context, by reducing the dimensionality of the system while retaining its essential dynamics, reduced-order modeling significantly lowers computational cost and facilitates efficient simulation and analysis.

Suppose that $\Omega \subset \mathbb{R}^d$ represents the bounded computational domain, and let the Hilbert space $L^2(\Omega)$ of square integrable functions on Ω be defined as follows:

$$L^2(\Omega) = \left\{ \phi : \Omega \to \mathbb{R} \middle| \int_\Omega |\phi|^2 d\Omega < \infty \right\} \tag{1}$$

to be endowed with the inner product

$$\langle \phi_i, \phi_j \rangle = \int_\Omega \phi_i \overline{\phi_j}\, d\Omega \quad for\ \phi_i, \phi_j \in L^2(\Omega) \tag{2}$$

and the induced norm $\|\phi\| = \sqrt{\langle \phi, \phi \rangle}$ for $\phi \in L^2(\Omega)$.

Consider a nonlinear PDE for state variable $u(x,t)$, $x \in \Omega_x \subset \mathbb{R}^d$, $t \in \Omega_t = [t_0, T]$, on domain $D = \Omega_x \times \Omega_t$, of the form

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{N}[u(x,t)], \tag{3}$$

where $\mathcal{N}$ is a nonlinear differential operator, with the initial condition

$$IC: \quad u(x, t_0) = u_0(x). \tag{4}$$

The Robin boundary conditions, defined by two time-continuous functions $g_a(t)$ and $g_b(t)$, specified at the endpoints of the spatial domain, take the general form defined below:

$$BC: \quad \begin{cases} \alpha_1 u(a,t) + \alpha_2 u_x(a,t) = g_a(t), \\ \beta_1 u(b,t) + \beta_2 u_x(b,t) = g_b(t), \end{cases} \tag{5}$$

where α_1, α_2, β_1, and β_2 are real prescribed coefficients.

Over the past half-century, partial differential equations have been predominantly solved using numerical techniques such as the finite difference and finite element methods. While these computational tools have significantly advanced our ability to approximate complex PDE solutions, they also introduce substantial implementation complexity and often require a solid understanding of the underlying analytical structure of the equations. In modern computational mathematics, reduced-order modeling has emerged as a powerful approach to mitigate computational demands by simplifying high-dimensional PDE systems, enabling more efficient and tractable solutions without sacrificing essential accuracy.

Definition 1. *The principle of reduced-order modeling aims to find an approximate solution of the form*

$$\begin{cases} u(x,t) = \lim\limits_{p\to\infty} \sum\limits_{j=1}^{p} a_j(t)\psi_j(x), \quad t \in \mathbb{R}_{\geq 0} \\ \frac{da_j(t)}{dt} = \mathcal{M}\big(a_j(t), t\big), \; a_j(t_0) = a_j^0, \end{cases} \tag{6}$$

where $\{\psi_j(x)\}_{j=1}^{p} \subset L^2(\Omega)$ *is a set of basis functions, and* $\mathcal{M}$ *is a locally Lipschitz continuous map governing the temporal dynamics of the coefficients* $a_j(t)$*. The initial condition is prescribed on the coefficient vector at the initial time* t_0*, with* $a_j(t_0) = a_j^0$*.*

The series is assumed to converge in the $L^2(\Omega)$ *norm for each fixed time* t*, that is,*

$$\lim_{p\to\infty} \left\| u(x,t) - \sum_{j=1}^{p} a_j(t)\psi_j(x) \right\| = 0.$$

This reduced-order model aims to approximate the true solution of the full-order system defined by the PDE (3) while ensuring a small global approximation error and computational stability.

Expression (6) is called the reduced-order model *of the PDE (3).*

The basis functions Ψ are called modes or shape functions; therefore, the representation (6) is called modal decomposition. Two primary goals are achieved through this modal decomposition approach: (1) the decoupling of spatial structures from the unsteady (time-dependent) components of the solution, and (2) the modeling of the original system's nonlinearity through a lower-dimensional temporal dynamical system. This separation not only reduces computational complexity but also preserves the essential dynamics of the original PDE system in a more efficient and interpretable form.

2.2. Koopman Operator Framework for Modal Decomposition

Introduced by Koopman in 1931, Koopman operator theory [34,35,57] provides a rigorous mathematical framework for the modal decomposition of nonlinear dynamical systems by representing them through linear, infinite-dimensional operators acting on observables. Despite its early formulation, the theory remained largely theoretical for decades.

In 1985, Lasota and Mackey [58] brought renewed scientific research interest to the original approach proposed by Koopman. Notably, they were the first to explicitly introduce and formalize the term "Koopman operator" in the literature, thereby establishing a standardized nomenclature for this concept. This terminology was further reinforced in

the second edition of their book [59], where the use of the Koopman operator as the adjoint to the Frobenius–Perron operator was maintained and elaborated upon functional analysis in space L^∞.

However, at that time, computational resources were still insufficient for the numerical treatment of large-scale complex systems. The practical potential of Koopman operator theory began to be realized in the early 2000s.

Its modern scientific revival began with the pioneering work of Mezić and Banaszuk [60,61] in fluid dynamics applications, establishing the foundation for practical, data-driven modal decomposition analysis. The seminal work by Mezi'c [61] introduced a rigorous spectral framework for analyzing nonlinear dynamical systems via Koopman operator theory, formalizing the spectral decomposition of observables into Koopman eigenfunctions and modes and establishing the foundation for operator-theoretic analysis of nonlinear dynamics. The spectral properties of the Koopman operator have been the subject of ongoing investigation, as exemplified in [39,42,62].

This paper proposes an adaptation of the Koopman framework to the context of nonlinear partial differential equations, extending its applicability to infinite-dimensional dynamical systems governed by PDEs.

Let $\mathcal{U}(t)$ denote the evolution operator of (3) such that

$$u(x,t) = \mathcal{U}(t)u(x,t_0), \tag{7}$$

i.e., $\mathcal{U}(t)$ maps the initial condition $u(x,t_0)$ to the state at time t.

Definition 2. *An observable of the PDE (3) is defined as a possibly nonlinear functional on the state space, typically a Hilbert space, such as $\mathcal{H} = L^2(\Omega)$:*

$$g : \mathcal{H} \to \mathbb{C}^m, \tag{8}$$

which maps the infinite-dimensional state u to a measurable output quantity. Here, $m \in \mathbb{N}$ denotes the dimension of the observable space, representing the number of scalar measurements or features extracted from the state u. These measurements may vary over time, thus capturing the temporal dynamics of the system in a finite-dimensional representation.

Remark 1. *The choice of observable g in the Koopman framework is flexible and can range from simple to highly structured representations. A common example is the identity observable, where $g(u) = u$, mapping the system state directly to itself. Alternatively, one may consider an infinite-dimensional observable vector, comprising a set of functions of the state, such as*

$$g(u) = [u,\ \kappa u,\ u^2,\ u^3,\ \sin(u), \ldots]^T, \tag{9}$$

where κ denotes a linear operator (e.g., a differential operator), and the higher-order terms capture nonlinear features of the system. This extended observable space facilitates a richer linear representation of the underlying nonlinear dynamics through the Koopman operator.

Definition 3. *The semigroup of Koopman operators $\{\mathcal{K}^t\}_{t\geq 0}$ acts on observable functions by composition with the evolution operator of the states as follows:*

$$\left(\mathcal{K}^t g\right)(u_0(x)) = g(\mathcal{U}(t)u_0(x)), \tag{10}$$

where $u_0(x) = u(x,t_0)$ is the initial condition. As a result, the Koopman operator is also known as a composition operator [34].

In practice, the Koopman operator is considered to be the dual, or more precisely, the left adjoint, of the Ruelle–Perron–Frobenius transfer operator [63–65], which acts on the space of probability density functions.

Remark 2. *The Koopman operator is linear, even though the underlying dynamics may be nonlinear [34]:*

$$\mathcal{K}^t(ag_1 + bg_2) = a\mathcal{K}^t g_1 + b\mathcal{K}^t g_2. \tag{11}$$

Remark 3. *The infinitesimal generator $\mathcal{L}$ of the Koopman semigroup governs the evolution of observables via*

$$\frac{d}{dt}g(u) = \mathcal{L}g(u), \tag{12}$$

where, for deterministic dynamics $\dot{u} = \mathcal{N}[u]$, the generator acts as

$$\mathcal{L}g(u) = \left\langle \frac{\delta g}{\delta u}, \mathcal{N}[u] \right\rangle, \tag{13}$$

with $\frac{\delta g}{\delta u}$ denoting the functional (Fréchet) derivative of g with respect to u. The inner product is taken over the spatial domain [58].

Following Koopman operator theory, the solution $u(x,t)$ can be expressed as a modal decomposition of the form

$$u(x,t) = \sum_{j=1}^{\infty} a_j(t)\,\Phi_j(x), \tag{14}$$

where $\Phi_j(x)$ are spatial Koopman modes, and $a_j(t)$ are time-dependent coefficients.

The evolution of the modal coefficients $\mathbf{a}(t) = [a_1(t), a_2(t), \dots]^T$ is governed by a (possibly nonlinear) dynamical system defined as follows:

$$\frac{d\mathbf{a}(t)}{dt} = \mathcal{M}(\mathbf{a}(t), t), \tag{15}$$

where $\mathcal{M}$ is a nonlinear operator (or vector field) that defines the temporal evolution of the coefficients $\mathbf{a}(t)$. Thus, the reduced-order solution of form (6) to the PDE is achievable form the perspective of Koopman mode decomposition.

Koopman modes are generally not unique, and their characterization depends on several interrelated factors. First, the definition of Koopman modes is inherently linked to the choice of observables, as the Koopman operator acts on a space of functions rather than directly on the state space. Different selections of observables can lead to different spectral representations, thereby affecting the resulting modes.

This intrinsic non-uniqueness is further compounded in practical settings, where Koopman modes are often approximated using numerical techniques such as DMD [32,33]. Such methods introduce additional variability due to factors, including data noise, finite sample sizes, and specific algorithmic implementations. Moreover, in cases where the Koopman operator exhibits a degenerate spectrum [10,66], the associated eigenspaces are typically multi-dimensional, allowing for multiple valid decompositions and further contributing to the non-uniqueness of the Koopman modes.

This work proposes a new algorithm, rooted in Koopman operator theory but fundamentally different in its formulation, specifically designed for the construction of a data-driven twin model. This new method computes Koopman modes and their associated temporal coefficients by leveraging the Koopman propagator operator projected onto a Krylov subspace of optimally selected rank. In contrast to classical DMD, the proposed approach facilitates a more structured and computationally efficient extraction of spectral

components by constraining the dynamics within a finite-dimensional subspace tailored to capture the most relevant features of the system's evolution.

3. Computational Aspects

This section introduces a novel methodology for constructing a data-driven twin model of a partial differential equation using Koopman mode decomposition. The proposed approach comprises two primary stages. The first stage, referred to as the offline phase, involves the extraction of Koopman modes and the corresponding modal decomposition coefficients that characterize the state evolution of the system. The second stage, known as the online phase, focuses on identifying a nonlinear continuous-time dynamical model that governs the temporal behavior of the data twin, thereby enabling real-time simulation and prediction. A detailed account of the computational implementation of the proposed methodology is presented hereinafter.

3.1. Offline Phase: Koopman Randomized Orthogonal Decomposition Algorithm

The data $u_i = u(x, t_i),\ t_i = i\Delta t,\ i = 0, \ldots, N_t$ represent measurements of the PDE solution at the constant sampling time Δt, with x representing the Cartesian spatial coordinate.

The data matrix whose columns represent the individual data samples is called *the snapshot matrix*:

$$V = \begin{bmatrix} u_0 & u_1 & \ldots & u_{N_t} \end{bmatrix} \in \mathbb{R}^{N_x \times (N_t+1)}. \tag{16}$$

Each column $u_i \in \mathbb{R}^{N_x}$ is a vector with N_x components, representing the spatial measurements corresponding to the $N_t + 1$ time instances.

We aim to find the data-driven twin representation of the solution at every time step $\{t_1, \ldots, t_{N_t}\}$ according to the following relation:

$$u_i = u(x, t_i) = \sum_{j=1}^{N_{DTM}} \underbrace{a_j(t_i)}_{\text{Modal amplitudes}} \overbrace{\phi_j(x)}^{\text{Leading Koopman modes}}, \quad i = 1, \ldots, N_t, \tag{17}$$

where $\Phi = \{\phi_j\}_{j=1}^{N_{DTM}}$ represents the extracted Koopman modes base functions, which we call *the leading Koopman modes*, $N_{DTM} \ll \min(N_x, N_t + 1)$ represents the number of terms in the representation (17), which we impose to be minimal, and $a_j(t_i)$ represents the modal growing amplitudes.

In classical Dynamic Mode Decomposition, the Koopman modes are not orthogonal, which often necessitates a large number of modes to achieve an accurate modal decomposition. This lack of orthogonality represents a significant drawback, as it may lead to redundancy and reduced computational efficiency.

It is more convenient to seek an orthonormal base of Koopman modes as

$$\Phi = \{\phi_1, \phi_2, \ldots\}, \quad \langle \phi_i(x), \phi_j(x) \rangle = \delta_{ij}, \quad \|\phi\| = 1, \tag{18}$$

where δ_{ij} is the Kronecker delta symbol, consisting of a minimum number of Koopman modes $\phi_i(x)$—such that the approximation of $u(x, t)$ through this base is as good as possible—in order to create a data twin model to the PDE of reduced computational complexity.

The objective of the present technique is to represent with highest precision the original data through the data twin model (17) having as few terms as possible. To achieve this goal, several major innovative features are introduced in the proposed algorithm:

- The method yields orthonormal Koopman modes, guaranteeing mutual orthogonality and thereby enabling a more compact representation of the system dynamics (see Theorem 1 and its proof).
- To avoid the computational burden associated with traditional high-dimensional algorithms, the proposed method employs a randomized singular value decomposition (RSVD) technique for dimensionality reduction. The incorporation of randomization offers a key advantage: It eliminates the need for additional selection criteria to identify shape modes, which is typically required in classical approaches such as DMD or POD. The resulting algorithm efficiently identifies the optimal reduced-order subspace that captures the dominant Koopman mode basis, thereby ensuring both computational efficiency and representational fidelity (see Appendix B).
- The methodology aims to achieve maximal correlation and minimal reconstruction error between the data-driven twin model and the exact solution of the governing PDE (see Section 5).

The theoretical foundation of RSVD was laid in the seminal work of Halko, Martinsson, and Tropp (2011) [67] as part of a broader class of probabilistic algorithms for matrix approximation. A subsequent study of Erichson and Donovan (2016) demonstrated its practical efficiency in motion detection [47]. Bistrian and Navon (2017) extended its use for the first time in fluid dynamics [53,54]. Recently, the theory of randomization has been further employed in the development of reduced-order modeling techniques, particularly in conjunction with projection learning methods [68], block Krylov iteration schemes [69] or deep learning [70].

Proposition 1 (*k*-RSVD: Randomized Singular Value Decomposition of rank *k*)**.** *Let* $V_0 \in \mathbb{R}^{N_x \times N_t}$ *be a real-valued data matrix, and let* $k \in \mathbb{N}$ *with* $k < \min(N_x, N_t)$ *be the target rank.*

Then, the rank-k Randomized Singular Value Decomposition (k-RSVD) of V_0 *yields an approximate factorization as follows:*

$$V_0 \approx U \Sigma W^H,$$

where $U \in \mathbb{R}^{N_x \times k}$ *contains approximate left singular vectors,*
$\Sigma \in \mathbb{R}^{k \times k}$ *is diagonal with approximate singular values,*
and $W \in \mathbb{R}^{N_t \times k}$ *contains approximate right singular vectors.*
This decomposition is computed via the k-RSVD algorithm described in Appendix B.

Theorem 1 (Koopman Randomized Orthogonal Decomposition)**.** *Let* $\{u_i\}_{i=0}^{N_t} \subset \mathbb{R}^{N_x}$ *denote a sequence of solution snapshots of a PDE, arranged into two time-shifted data matrices, defined as*

$$\begin{aligned} V_0 &= \begin{bmatrix} u_0 & u_1 & \dots & u_{N_t-1} \end{bmatrix} \in \mathbb{R}^{N_x \times N_t}, \\ V_1 &= \begin{bmatrix} u_1 & u_2 & \dots & u_{N_t} \end{bmatrix} \in \mathbb{R}^{N_x \times N_t}. \end{aligned} \tag{19}$$

Let $k \in \mathbb{N}$*, with* $k < N_t$*, be a target rank. Then, there exists a reduced k-order approximation* $u_i^{(k)}$ *of the full state* u_i *such that*

$$u_i^{(k)} = u^{(k)}(x, t_i) = \sum_{j=1}^{k} a_j(t_i)\, \phi_j(x), \quad \forall i \in \{1, \dots, N_t\}, \tag{20}$$

where $\{\phi_j\}_{j=1}^{k}$ *is an orthonormal set of Koopman modes spanning the reduced subspace defined as*

$$\operatorname{span}\{\phi_1, \phi_2, \dots, \phi_k\}, \quad \phi_j = \frac{\langle U, X_{\cdot,j} \rangle}{\left\| \langle U, X_{\cdot,j} \rangle \right\|_2}, \quad j = 1, \dots, k. \tag{21}$$

where

$U \in \mathbb{R}^{N_x \times k}$ *is the matrix of the first k left singular vectors of* V_0 *obtained via k-RSVD;*

$X \in \mathbb{R}^{k \times k}$ *contains the eigenvectors of the Gram matrix associated with the finite-dimensional approximation of the Koopman operator* $\mathcal{K}$ *defined by the time evolution*

$$u_{N_t} = \mathcal{K}^{N_t} u_0.$$

The corresponding modal coefficients $a_j(t_i)$ *are computed via the inner product between the data matrix* V_0 *and Koopman modes defined as*

$$a_j(t_i) = \langle V_0, \phi_j \rangle, \quad \forall j \in \{1, \ldots, k\}, \quad \forall i \in \{1, \ldots, N_t\}. \tag{22}$$

The proposed modal representation method is hereafter referred to as Koopman Randomized Orthogonal Decomposition (KROD).

Proof. We consider a sequence of solution snapshots $\{u_i\}_{i=0}^{N_t} \subset \mathbb{R}^{N_x}$ obtained from a discrete-time dynamical system governed by an underlying (possibly nonlinear) evolution law. Under the Koopman operator framework [34], we adopt the *identity observable* $g(u) = u$ so that the state vectors themselves are treated as observables (see Remark 1).

There exists a linear Koopman operator $\mathcal{K}$ acting on observables such that

$$u_i = \mathcal{K}^i u_0, \quad \text{for } i = 0, 1, \ldots, N_t. \tag{23}$$

This implies that each state u_i is obtained by iterated application of the Koopman operator to the initial state u_0 under the identity observable. The operator $\mathcal{K}$ thus advances the system in time linearly in the (possibly infinite-dimensional) observable space.

The sequence (23) lies in the Krylov subspace

$$\mathcal{K}_{N_t+1}(\mathcal{K}, u_0) = span\Big\{u_0, \mathcal{K}u_0, \mathcal{K}^2 u_0, \ldots, \mathcal{K}^{N_t} u_0\Big\}, \tag{24}$$

that captures how the observable u_0 evolves under the Koopman dynamics.

The Krylov subspace (24) gives a finite-dimensional proxy to study the action of the infinite-dimensional Koopman operator. For a sufficiently long sequence of the snapshots, suppose that the last snapshot u_{N_t} can be written as a linear combination of previous N_t vectors, on the Krylov subspace (24), such that

$$u_{N_t} = c_0 u_0 + c_1 u_1 + \cdots + c_{N_t-1} u_{N_t-1} + \mathcal{R}, \tag{25}$$

in which $\mathrm{c}_i \in \mathbb{R}, \mathrm{i} = 0, \ldots, \mathrm{N_t} - 1$, and $\mathcal{R}$ is the residual vector. The following relations are true:

$$\{u_1, u_2, \ldots u_{N_t}\} = \mathcal{K}\{u_0, u_1, \ldots u_{N_t-1}\} = \{u_1, u_2, \ldots, V_0 c\} + \mathcal{R}, \tag{26}$$

where $c = \begin{pmatrix} c_0 & c_1 & \ldots & c_{N_t-1} \end{pmatrix}^T$ is the unknown column vector.

Assume there exists a diagonal matrix $\mathcal{S}$ whose eigenvalues asymptotically approximate those of $\mathcal{K}$ as $\|\mathcal{R}\|_2 \to 0$ approaches zero. Solving the eigenvalue problem

$$V_1 = \mathcal{K} V_0 = V_0 \mathcal{S} + \mathcal{R}, \tag{27}$$

is equivalent to solve the minimization problem

$$\min_{\mathcal{S}} \ \mathcal{R} = \|V_1 - V_0 \mathcal{S}\|_2, \tag{28}$$

where $\|\cdot\|_2$ is the L_2-norm of $\mathbb{R}^{N_x \times N_t}$.

For $1 \leq k \leq \min(N_x, N_t)$, we identify the k-RSVD of V_0 that yields the following factorization:

$$V_0 = U\Sigma W^H, \tag{29}$$

where $U \in \mathbb{R}^{N_x \times k}$ and $W \in \mathbb{R}^{N_t \times k}$ are orthogonal matrices that contain the eigenvectors of $V_0 {V_0}^H$ and ${V_0}^H V_0$, respectively, $\Sigma = diag(\sigma_1, \ldots, \sigma_k) \in \mathbb{R}^{k \times k}$ is a square diagonal matrix containing the singular values of V_0, and H means the conjugate transpose.

Relations $\mathcal{K}V_0 = V_1 = V_0 S + \mathcal{R}, \|\mathcal{R}\|_2 \to 0$ and $V_0 = U\Sigma W^H$ yield

$$\mathcal{K}U\Sigma W^H = V_1 = U\Sigma W^H \mathcal{S} \;\Rightarrow\; U^H \mathcal{K} U\Sigma W^H = U^H U\Sigma W^H \mathcal{S} \;\Rightarrow\; \mathcal{S} = U^H \mathcal{K} U.$$

As a consequence, the solution to the minimization problem (28) is the matrix operator

$$\mathcal{S} = U^H \left(V_1 W \Sigma^{-1} \right). \tag{30}$$

As a direct result, the eigenvalues of $\mathcal{S}$ will converge toward the eigenvalues of the Koopman propagator operator $\mathcal{K}$.

Let $X \in \mathbb{C}^{k \times k}$ and $\Lambda \in \mathbb{C}^{k \times k}$ be the eigenvectors, respectively the eigenvalues of the Hermitian positive semi-definite Gram matrix $\mathcal{G} = \mathcal{S}^H \mathcal{S}$ as

$$\mathcal{G}X = X\Lambda. \tag{31}$$

From the k-RSVD of V_0, it follows that the column space of V_0 is the same as the column space of U, i.e.,

$$\text{Range}(V_0) = \text{Range}(U). \tag{32}$$

Using the definition $V_1 = V_0 \mathcal{S}$ and substituting into the expression for $\mathcal{S}$,

$$\begin{aligned} \mathcal{S} &= U^H V_1 W \Sigma^{-1} \\ &= U^H (V_0 \mathcal{S}) W \Sigma^{-1} \\ &= U^H (U\Sigma W^H \mathcal{S}) W \Sigma^{-1} \\ &= (U^H U)\Sigma (W^H \mathcal{S} W) \Sigma^{-1} \\ &= \Sigma (W^H \mathcal{S} W) \Sigma^{-1}, \end{aligned}$$

since $U^H U = I$.

Since $\mathcal{S}$ is diagonal and W is unitary, $W^H \mathcal{S} W$ represents a unitary similarity transform [71] of matrix $\mathcal{S}$, i.e., $\mathcal{S}$ acts in a reduced coordinate system defined by the k-RSVD.

Let

$$\Phi = UX. \tag{33}$$

Since the columns of X are vectors in $\mathbb{C}^k$, and U is an orthonormal basis for the column space of V_0, the columns of Φ are linear combinations of the columns of U.

Thus,

$$\text{Range}(\Phi) \subseteq \text{Range}(U) = \text{Range}(V_0), \tag{34}$$

i.e., the set of vectors $\Phi = UX$ lies in the column space of V_0. Therefore, the span of the eigenvectors Φ forms a subspace of the space containing the data in V_0.

It follows that there exists a subspace of the space containing the data in V_0 spanned by the following sequence of Koopman modes:

$$span\{\phi_1, \phi_2, \ldots, \phi_k\},\ \{\phi_j\}_{j=1}^{k} = \langle U, X_{\cdot,j} \rangle, \tag{35}$$

where U represents the matrix of left singular vectors produced by k-RSVD of data matrix V_0.

We proceed to demonstrate that the columns of $\Phi = UX$ are orthonormal. Since U comes from the k-RSVD of V_0, it has orthonormal columns defined as

$$U^H U = I. \tag{36}$$

Now, compute the inner product of Φ as

$$\Phi^H \Phi = (UX)^H (UX) = X^H U^H U X = X^H X. \tag{37}$$

So, the orthonormality of Φ depends on $X^H X$. The matrix X contains the eigenvectors of the Hermitian Gram matrix $\mathcal{G} = \mathcal{S}^H \mathcal{S}$. Since $\mathcal{G}$ is Hermitian, it has a complete set of orthonormal eigenvectors. Thus, by choosing the eigenvectors to be orthonormal, we have

$$X^H X = I \quad \Rightarrow \quad \Phi^H \Phi = X^H X = I, \tag{38}$$

i.e., columns of Φ are orthonormal.

It follows that

$$\langle \phi_i, \phi_j \rangle = \delta_{ij}, \quad 1 \leq i \leq j \leq k \tag{39}$$

i.e., Koopman modes Φ form an orthnormal base to the data space.

We continue by establishing the formula for the modal coefficients.

Assume that the data evolve linearly in reduced k-order representation (20) and are written in matrix formulation as

$$V_1 = \Phi A. \tag{40}$$

Multiplying both sides on the left by Φ^H results in

$$\Phi^H V_1 = \Phi^H \Phi A. \tag{41}$$

Since Φ has orthonormal columns, we get

$$\Phi^H \Phi = I \quad \Rightarrow \quad \Phi^H V_1 = A. \tag{42}$$

Thus,

$$A = \Phi^H V_1. \tag{43}$$

If $V_1 = V_0 \mathcal{S}$, and $\Phi = UX$, then we can alternatively write

$$A = \Phi^H V_1 = \Phi^H V_0 \mathcal{S}. \tag{44}$$

We choose to project V_0 onto the subspace spanned by the orthonormal basis Φ to reduce computational complexity. Consequently, the coefficient matrix

$$A = \Phi^H V_0 \tag{45}$$

provides a representation with a smaller approximation error in the least-squares sense. □

Proposition 2 (Multi-objective twin-model selection via Pareto front analysis)**.** *Let* $\{k_1, k_2, \ldots, k_N\} \subset \mathbb{N}$ *be a strictly increasing sequence of target ranks satisfying*

$$1 < k_1 < k_2 < \cdots < k_N \leq N_t.$$

For each target rank k_i*, define a Koopman-based model triplet* $T_i = (\Phi_i, A_i, k_i)$ *consisting of*
Φ_i*: a Koopman mode basis of size* k_i*;*
A_i*: the corresponding modal amplitude matrix;*

k_i: *the reduced-order rank.*

Define the candidate set of model triplets as

$$\mathcal{T} = \{T_1, T_2, \ldots, T_N\},$$

and let $u^{(T_i)}$ *denote the reduced-order solution associated with* T_i*, which is given by*

$$u^{(T_i)}(x,t) = \sum_{j=1}^{k_i} a_j(t)\,\phi_j(x), \quad \text{for all } t \in \{t_1, \ldots, t_{N_t}\}.$$

Given the full-order PDE solution u*, the goal is to identify the optimal triplet by solving the following constrained multi-objective optimization problem:*

$$\begin{aligned} &\text{Find } \mathcal{T}^* \subseteq \mathcal{T}, \quad \text{such that:} \\ &\text{Objective } 1: \quad \min_{T_i \in \mathcal{T}} \; E(T_i) := \|u - u^{(T_i)}\|_2, \\ &\text{Objective } 2: \quad \max_{T_i \in \mathcal{T}} \; C(T_i) := \frac{\|u^H u^{(T_i)}\|_2^2}{\|u^H u\|_2 \cdot \|(u^{(T_i)})^H u^{(T_i)}\|_2}, \end{aligned} \tag{46}$$

where $\|\cdot\|_2$ *denotes the discrete Euclidean norm over all spacetime points, and* H *denotes the Hermitian (conjugate transpose).*

The Pareto-optimal subset of model triplets is defined as

$$\mathcal{T}^* = \big\{ T_i \in \mathcal{T} \;\big|\; \nexists T_j \in \mathcal{T} \text{ such that } E(T_j) \le E(T_i), \; C(T_j) \ge C(T_i), \text{ and at least one inequality is strict} \big\}.$$

Let the Pareto-optimal set of the Koopman model triplet be denoted by the solution that represents the best trade-offs between competing objectives as

$$\mathcal{T}^* = (\Phi_{\text{DTM}}, A_{\text{DTM}}, N_{\text{DTM}}),$$

where the number of modes in the final reduced model is N_{DTM}*.*

The resulting data-driven twin model (DTM) is then expressed as

$$u_i^{\text{DTM}}(x) = \sum_{j=1}^{N_{\text{DTM}}} a_j(t_i)\,\phi_j(x), \quad i = 1, \ldots, N_t, \tag{47}$$

where $a_j(t_i) = (A_{\text{DTM}})_{i,j}$*, and* $\phi_j(x) = (\Phi_{\text{DTM}})_{:,j}$*.*

This framework is designed to simultaneously balance two competing objectives: (Objective 1) ensuring high fidelity by minimizing the pointwise discrepancy between the true solution and data twin model prediction and (Objective 2) preserving structural alignment by maximizing the normalized cosine similarity or correlation between their respective dynamics.

The optimization problem (46) is resolved numerically using Pareto front analysis [72], which is a multi-objective optimization strategy that identifies the set of non-dominated solutions balancing competing objectives. The implementation details are provided in Appendix B.

3.2. Online Phase: Modeling the Data Twin Temporal Dynamic by Explainable Deep Learning

The calculation step of the modes in the offline phase is performed only once. The data twin model is not complete without the temporal component. These computations are carried out in the second step, during which the mathematical model is integrated over

time. Additionally, if needed, the dynamics of the reduced-order model can be evaluated at any specific time point. For this reason, this step is referred to as the online phase, as it can be executed on demand to simulate the behavior of the original process.

Unlike conventional machine learning models, which typically function as opaque black boxes operating on entire datasets without disentangling underlying factors, a key innovation of this study lies in the explicit modeling of temporal dynamics decoupled from spatial dependencies. This is achieved through the use of explainable deep learning techniques, which offer interpretable insights into the data twin model's parameters.

In this work, Nonlinear AutoRegressive models with eXogenous inputs (NLARX) are employed as a novel deep learning-based approach within the domain of nonlinear system identification. Since the emergence of artificial neural networks as numerical tools [73,74], NLARX models have been used for various purposes, ranging from simulation [75] to nonlinear predictive control [76] or higher-order nonlinear optimization problems [77,78].

The objective is to identify a dynamical system that effectively captures the nonlinear temporal evolution of the underlying process expressed in the form

$$\begin{cases} \frac{d\hat{a}_j(t)}{dt} = \mathcal{M}\big(\hat{a}_j(t), a_j(t), t\big), \quad t \in \mathbb{R}_{\geq 0} \\ \hat{a}_j(t_0) = \hat{a}_j^0, \end{cases} \tag{48}$$

based on modal temporal coefficients $a_j(t_i),\ j = 1, \ldots, N_{DTM},\ \ t_i \in \{\mathrm{t}_1, \ldots, \mathrm{t}_{N_t}\}$, computed by the KROD algorithm in the offline phase, and initial condition obtained by projecting the first snapshot onto the Koopman mode basis as

$$\hat{a}_j^0 = \langle u_0, \phi_j \rangle, j = 1, \ldots, N_{DTM}. \tag{49}$$

To approximate the system dynamics, the evolution of the state $\hat{a}(t)$ is modeled using the NLARX model structure, which is expressed as

$$\hat{a}(t) = \mathcal{F}[\hat{a}(t-1), \ldots, \hat{a}(t-n_a), a(t-n_k), \ldots, a(t-n_k-n_b+1)] + e(t), \tag{50}$$

where $\mathcal{F}$ is a nonlinear estimator realized by a cascade-forward neural network, capturing the underlying nonlinear temporal dependencies and input effects. Here, n_a denotes the number of past output terms, n_b represents the number of past input terms used to predict the current output, and n_k corresponds to the input delay, while $e(t)$ captures the modeling error. Each output of the NLARX model described in (50) is determined by regressors, which are transformations of previous input and output values. Typically, this mapping consists of a linear component and a nonlinear component, with the final model output being the sum of these two contributions.

Thus, the NLARX model serves as an empirical approximation of the continuous dynamics $\mathcal{M}$, with the neural network mapping $\mathcal{F}$ implicitly representing the evolution operator of the system such that

$$\hat{a}(t) \approx \hat{a}(t-1) + \Delta t \cdot \mathcal{M}(\hat{a}(t-1), a(t-1), t-1) \approx \mathcal{F}[\hat{a}(t-1), \ldots, a(\cdot)] + e(t), \tag{51}$$

where Δt is the discrete time step, and $e(t)$ denotes the modeling error.

Training the NLARX model can be formulated as a nonlinear unconstrained optimization problem of the form

$$\theta(n_a, n_b, n_k) = \arg\min \frac{1}{2N_t} \sum_{i=1}^{N_t} \|a(t_i) - \widehat{a}(t_i)\|_2. \tag{52}$$

In this formulation, the training dataset comprises the measured input $a(t)$, while $\widehat{a}(t)$ denotes the output generated by the NLARX model. The symbol $\|\cdot\|_2$ refers to the L_2 norm, and $\theta(n_a, n_b, n_k)$ represents the parameter vector associated with the nonlinear function $\mathcal{F}$.

The NLARX architecture is well suited for capturing system dynamics by feeding previous network outputs back into the input layer. It also allows for the specification of how many past input and output time steps are necessary to accurately represent the system's behavior. A critical aspect of effectively applying an NLARX network lies in the careful selection of inputs, input delays, and output delays. The objective is to adjust the network parameters $\theta(n_a, n_b, n_k)$ across the entire trajectory to minimize the objective function defined in (52).

The reduced-order data-driven twin model for the PDE solution is obtained in the form

$$\begin{cases} u^{DTM}(x,t) = \sum_{j=1}^{N_{DTM}} \hat{a}_j(t)\phi_j(x), \ t \in R_{\geq 0}, \\ \hat{a}(t) = \mathcal{F}[\hat{a}(t-1), \ldots, \hat{a}(t-n_a), a(t-n_k), \ldots, a(t-n_k-n_b+1)], \quad \hat{a}_j(t_0) = a_j^0, \end{cases} \tag{53}$$

where N_{DTM} is the rank of the optimal Koopman basis, computed by the algorithm, $\phi_j(x), j = 1, \ldots, N_{DTM}$ are the leading Koopman modes representing dominant spatial patterns computed by the algorithm, and $\hat{a}(t)$ is the system temporal dynamics in the reduced-order space—modeled by the deep learning architecture with initial condition given by Equation (49).

3.3. Qualitative Analysis of the Data-Driven Twin Model

A qualitative assessment of the orthogonality between the Koopman modes derived using the KROD algorithm proposed in this study can be carried out through the Modal Assurance Criterion (MAC) [79].

The discrete MAC value between two mode vectors $\phi_i, \phi_j \in \mathbb{C}^n$ is defined as

$$\mathrm{MAC}_{ij} = \frac{\left|\langle \phi_i, \phi_j \rangle\right|^2}{\langle \phi_i, \phi_i \rangle \cdot \langle \phi_j, \phi_j \rangle}, \tag{54}$$

where $\langle \phi_i, \phi_j \rangle = \phi_i^H \phi_j$ denotes the Hermitian (complex-conjugate) inner product in $\mathbb{C}^n$.

The orthogonality matrix $\mathrm{MAC} \in \mathbb{R}^{N_{DTM} \times N_{DTM}}$ is symmetric and real-valued. The MAC_{ij} values lie within the interval $[0,\ 1]$, as the leading Koopman modes obtained from the KROD algorithm are normalized vectors. In this context, Koopman modes are considered orthogonal if

$$MAC_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases} \tag{55}$$

The Pearson correlation coefficient is employed to quantify the correlation between the data twin model solution and the true solution of the partial differential equation, providing a measure of how closely the data twin model approximates the true dynamics across the spacetime domain.

Let $u(x,t)$ be the true solution and $u^{DTM}(x,t)$ be the data twin model solution of the PDE, sampled at N_x spatial points and N_t time instances. Define the flattened vectors as

$$\mathbf{U} = [u(x_1,t_1), \ldots, u(x_{N_x},t_1), u(x_1,t_2), \ldots, u(x_M,t_{N_t})] \in \mathbb{R}^{N_x N_t},$$
$$\mathbf{U^{DTM}} = [u^{DTM}(x_1,t_1), \ldots, u^{DTM}(x_{N_x},t_1), u^{DTM}(x_1,t_2), \ldots, u^{DTM}(x_M,t_{N_t})] \in \mathbb{R}^{N_x N_t}.$$

Let U_k denote the k-th element of the vector $\mathbf{U}$ and U_k^{DTM} denote the k-th element of the vector $\mathbf{U^{DTM}}$.

The Pearson correlation coefficient ρ between $\mathbf{U}$ and $\mathbf{U_{DTM}}$ is given by

$$\rho = \frac{\sum_{k=1}^{N_x N_t} (U_k - \bar{U})\left(U_k^{DTM} - \overline{U^{DTM}}\right)}{\sqrt{\sum_{k=1}^{N_x N_t} (U_k - \bar{U})^2}\sqrt{\sum_{k=1}^{N_x N_t} \left(U_k^{DTM} - \overline{U^{DTM}}\right)^2}}, \tag{56}$$

where the quantities

$$\bar{U} = \frac{1}{N_x N_t}\sum_{k=1}^{N_x N_t} U_k, \quad \overline{U^{DTM}} = \frac{1}{N_x N_t}\sum_{k=1}^{N_x N_t} U_k^{DTM}$$

represent the mean values of the vectors $\mathbf{U}$ and $\mathbf{U}^{\mathrm{DTM}}$, respectively.

In addition to the Pearson correlation coefficient, the mean absolute error (MAE) is employed as a complementary metric to assess the average magnitude of pointwise discrepancies over the spacetime domain, providing a more direct measure of predictive accuracy irrespective of correlation structure, as

$$MAE = \frac{1}{N_x N_t}\sum_{k=1}^{N_x N_t} \left|U_k - U_k^{DTM}\right|. \tag{57}$$

In addition to these metrics, to provide fine-grained, pointwise insight into how well the data twin model approximates the true solution of PDE at specific locations in space and time, we define the absolute local error matrix $E \in \mathbb{R}^{N_x \times N_t}$ as

$$E_{ij} = \left|u(x_i, t_j) - u^{DTM}(x_i, t_j)\right| \tag{58}$$

for all $i = 1, \ldots, N_x$ and $j = 1, \ldots, N_t$.

3.4. Time Simulation and Validation of the Data-Driven Twin Model

The data-driven twin model was subjected to time-domain simulation, followed by a two-step input–output validation of its predicted response. In the first validation step, the model was trained using the initial two-thirds of the available snapshots, with the remaining one-third reserved for validation. In the subsequent step, the training dataset was further reduced to the first one-third of the snapshots, and the remaining two-thirds were employed to assess the model's predictive performance.

4. Data-Driven Twin Modeling of Shock Wave Phenomena Using KROD

This section demonstrates the application of the proposed KROD algorithm to shock wave modeling using the viscous Burgers' equation.

4.1. Governing Equations of the Mathematical Model

The viscous Burgers equation model is considered in the form

$$\begin{cases} \frac{\partial}{\partial t}u(x,t) + \frac{\partial}{\partial x}\left(\frac{u(x,t)^2}{2}\right) = \nu\frac{\partial^2}{\partial x^2}u(x,t), \quad t > 0, \quad \nu > 0, \\ u(x,0) = u_0(x), \quad x \in \mathbb{R}, \end{cases} \tag{59}$$

where $u(x,t)$ is the unknown function of time t, and ν is the viscosity parameter.

Three experiments are considered, exhibiting progressively more complex dynamics, with each characterized by an initial condition of the following form:

$$Experiment\ 1: \quad u_0(x) = -\sin(\pi x), \tag{60}$$

$$Experiment\ 2: \quad u_0(x) = \begin{cases} u_L, & x \leq 0 \\ u_R, & x > 0 \end{cases}, \tag{61}$$

$$Experiment\ 3: \quad u_0(x) = -\cos(1.5\pi x)^2. \tag{62}$$

Homogeneous Dirichlet boundary conditions are imposed for all three examples and specified as follows:

$$u(0,t) = u(L,t) = 0 \tag{63}$$

Equation (60), which defines the initial condition for Experiment 1, produces a sinusoidal pulse characterized by an abrupt change in slope at the boundaries of the domain. The initial value problem with the discontinuous initial condition given by (61) is referred to as a Riemann problem, which leads to the formation of a shock wave. For Experiment 3, the initial condition described by (62) induces a more complex evolution of the pulse.

All three experiments involve phenomena that pose significant challenges for standard numerical methods to accurately capture.

4.2. Analytical and Numerical Derivation of the Exact Solution

The Cole–Hopf transformation, independently introduced by Hopf (1950) [80] and Cole (1951) [81], converts the Burgers' equation into the heat equation, thereby enabling the derivation of its exact solution. The analytical solution of the Burgers equation model (59) is derived using the Cole–Hopf transformation, as detailed in Appendix A.1.

Using the Cole–Hopf transformation, the analytic solution to the problem (59) is obtained in the following form:

$$u(x,t) = \frac{\int_{-\infty}^{\infty} \frac{x-\xi}{t} \varphi_0(\xi) e^{-\frac{(x-\xi)^2}{4\nu t}} d\xi}{\int_{-\infty}^{\infty} \varphi_0(\xi) e^{-\frac{(x-\xi)^2}{4\nu t}} d\xi}. \tag{64}$$

To numerically compute the exact solution (64) corresponding to the initial conditions of the three test cases (60)–(62), the Gauss–Hermite quadrature technique is employed [82], which is fully described in Appendix A.2. The following results are thereby obtained for the three experimental cases.

The exact solution to the Burgers equation model (59) with the initial condition (60) is equivalent to the following form:

$$u(x,t) = \frac{\int_{-\infty}^{\infty} 4\nu z\, e^{-\frac{1}{2\nu\pi}\cos\left[\pi\left(x - z\sqrt{4\nu t}\right)\right]} e^{-z^2} dz}{\int_{-\infty}^{\infty} \sqrt{4\nu t}\, e^{-\frac{1}{2\nu\pi}\cos\left[\pi\left(x - z\sqrt{4\nu t}\right)\right]} e^{-z^2} dz}. \tag{65}$$

In the case of Experiment 2, the exact solution to the Burgers equation model (59) with the initial condition (61) is

$$u(x,t) = \frac{\int_{-\infty}^{\infty} 4\nu z\, e^{-\frac{u_R}{2\nu}\left(x - z\sqrt{4\nu t}\right)} e^{-z^2} dz}{\int_{-\infty}^{\infty} \sqrt{4\nu t}\, e^{-\frac{u_R}{2\nu}\left(x - z\sqrt{4\nu t}\right)} e^{-z^2} dz}. \tag{66}$$

In the case of Experiment 3, the exact solution to the Burgers equation model (59) with the initial condition (62) is

$$u(x,t) = \frac{\int_{-\infty}^{\infty} 4\nu z \, e^{\frac{1}{4\nu}\left(\frac{1}{3\pi}\sin\left(3\pi(x-z\sqrt{4\nu t})\right)+x-z\sqrt{4\nu t}\right)} e^{-z^2} dz}{\int_{-\infty}^{\infty} \sqrt{4\nu t} \, e^{\frac{1}{4\nu}\left(\frac{1}{3\pi}\sin\left(3\pi(x-z\sqrt{4\nu t})\right)+x-z\sqrt{4\nu t}\right)} e^{-z^2} dz}. \tag{67}$$

5. Numerical Results

The algorithm underlying the Koopman Randomized Orthogonal Decomposition method is described in detail in Appendix B. This section presents numerical results highlighting the computational efficiency of the proposed KROD method combined with deep learning (DL) for temporal simulation. The findings demonstrate that the approach produces a high-fidelity, reduced-complexity model that accurately captures the dynamics of the original system and enables real-time simulation, effectively functioning as a data-driven twin model.

The computational domain is defined as $[0, L]$, with $L = 2$, and the time interval considered is $[0, T]$, where $T = 3$. The viscosity parameter in the Burgers equation was set to $\nu = 10^{-2}$ for all three experiments. The domain was uniformly discretized using $N = 100$ grid points, resulting in a mesh size of $\Delta x = 0.02$.

The Riemann problem parameters were assumed to be $u_L = 0.1$ and $u_R = 0.5$.

By applying the Cole–Hopf transformation (A1), the analytical solutions to the Burgers problem (59) were obtained for all three experiments, corresponding, respectively, to the forms (65), (66), and (67), as detailed in Appendix A.1. Subsequently, the exact solutions for the three test cases were computed numerically using the Gauss–Hermite quadrature method (A8) with $n = 100$ nodes (see Appendix A.2).

The following figure presents the exact solution to the viscous Burgers equation model. Figure 1a presents the exact solution of the viscous Burgers' equation corresponding to the initial condition given in Equation (60), as derived in the analytical form shown in Equation (65). The exact solution of the viscous Burgers' equation for the second experiment, based on the initial condition specified in Equation (61), is depicted in Figure 1b. This solution is obtained in the analytical form provided in Equation (66). Figure 1c shows the exact solution of the viscous Burgers' equation for the third experiment, corresponding to the initial condition defined in Equation (62), and derived in the analytical form given in Equation (67).

The data-driven twin models of the form (47) are generated using the KROD algorithm as high-fidelity surrogates for the three Burgers problems examined in this study. For this purpose, the training dataset consists of a total of $N_t = 300$ snapshots collected at uniformly spaced time intervals of $\Delta t = 0.01$ and $N_x = 101$ spatial measurements recorded for each time snapshot.

KROD algorithm determines the optimal dimension N_{DTM} of the subspace spanned by the dominant Koopman modes by solving the multiobjective optimization problem with nonlinear constraints stated in (46). This problem is solved by employing a genetic algorithm to identify the Pareto front corresponding to the two fitness functions.

Figures 2a,c,e depict the optimization objectives defined in Equation (46) with respect to the modal space dimension in the range $[5, 200]$. Figures 2b,d,f present the Pareto front solution for each of the three experiments, respectively.

For Experiment 1, the optimal dimension of the subspace spanned by the leading modes is $N_{DTM} = 10$ (Figure 2b). In Experiment 2, the data twin model representing the Riemann problem dynamics has an optimal subspace dimension of $N_{DTM} = 20$ (Figure 2d). For Experiment 3, the increased complexity of the pulse is captured by a data twin model with a higher dimension of $N_{DTM} = 45$ (Figure 2f). In all three cases, the computational

time required by the KROD algorithm in offline phase remains extremely low, as illustrated in Figure 3.

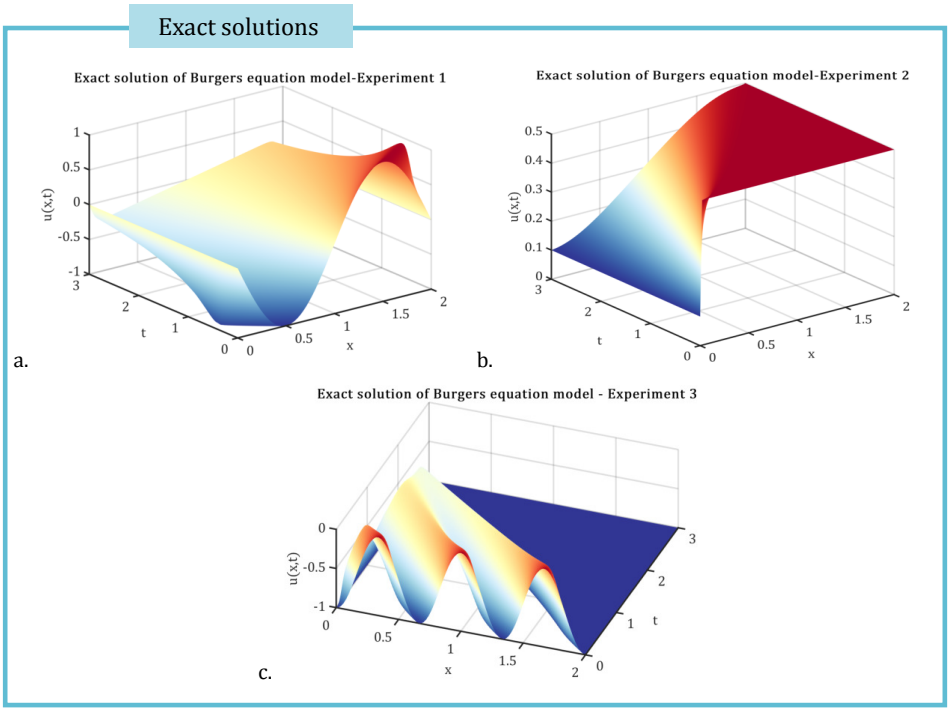


Figure 1. (**a**) Exact solution of the viscous Burgers equation for the initial condition in Equation (60), derived analytically as in Equation (65). (**b**) Exact solution of the viscous Burgers' equation for the second experiment with initial condition from Equation (61), derived analytically as in Equation (66). (**c**) Exact solution of the viscous Burgers equation for the third experiment with initial condition from Equation (62), derived analytically as in Equation (67).

The KROD algorithm identified the optimal dimension of the modal basis and computed the optimal Koopman modes along with their associated time-dependent amplitudes. As a result, the data-driven twin model (47) was constructed, which—by definition—exhibits minimal approximation error and maximal correlation with respect to the original data model.

To facilitate a qualitative assessment of the DTMs, Table 2 reports the outcomes of the multiobjective optimization problem (46) as solved by the KROD. Specifically, the table includes the dimensionality of the leading Koopman modes; the absolute error, as defined in Equation (57); and the correlation coefficient, as given by Equation (56) between the exact solution and the corresponding data twin model for each of the three experiments considered.

An analysis of the results presented in Table 2 reveals that the KROD algorithm produced models exhibiting perfect correlation with the original data, i.e., $\rho = 1$, across all three test cases. Furthermore, the absolute errors remained remarkably low: on the order of $MAE < \mathcal{O}(10^{-7})$.

The selection of modes solely based on their energy ranking proves effective only under specific conditions [45,83]. Some modes, despite contributing minimally to the overall energy, may exhibit rapid temporal growth at low amplitudes or represent strongly damped high-amplitude structures. Such modes can play a significant role in enhancing the fidelity of the data twin model. A key advantage of the KROD algorithm over traditional DMD-based approaches lies in its ability to generate a substantially lower-dimensional subspace that retains the most dynamically significant Koopman modes. As a result, the KROD method avoids the need for an additional—often subjective—mode selection

criterion. Moreover, the orthogonality of the KROD modes contributes to their qualitative superiority, enabling accurate modeling with a reduced number of modes.

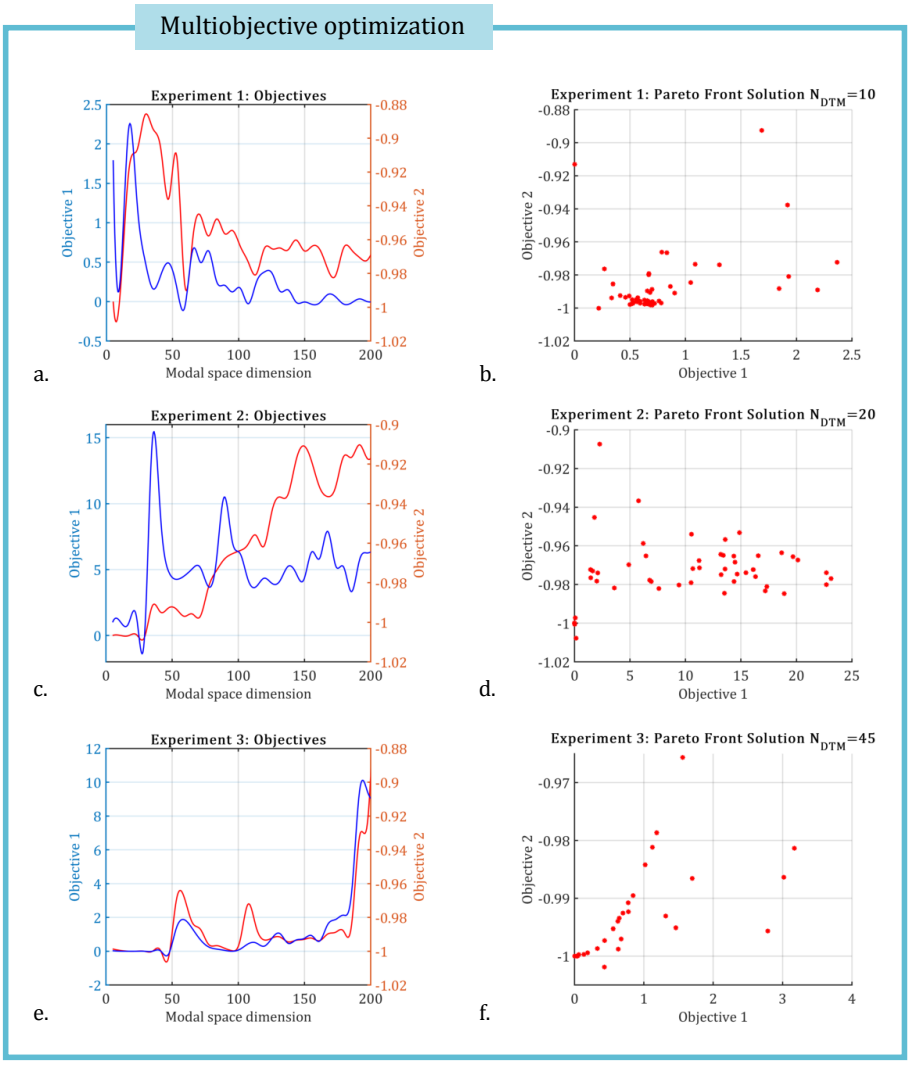


Figure 2. Figures (**a**,**c**,**e**) show the optimization objectives defined in Equation (46) with respect to the modal space dimension in the range $[5, 200]$. Figures (**b**,**d**,**f**) present the corresponding Pareto front solutions for the three experiments obtained using a genetic algorithm.

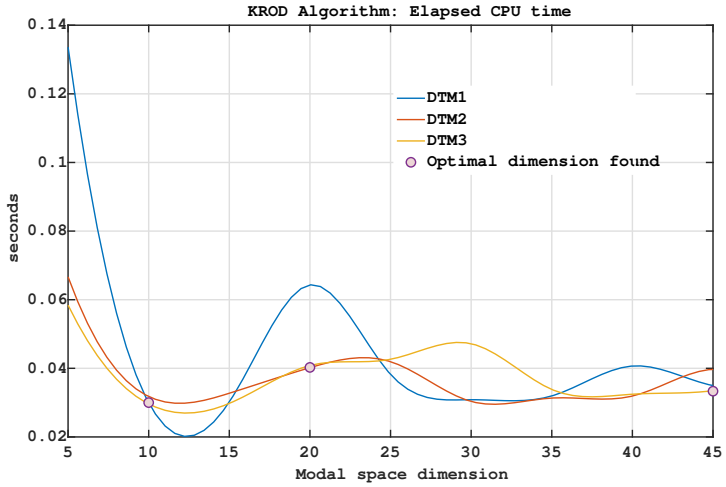


Figure 3. The running CPU time for KROD in offline phase, for the three experiments, respectively.

Table 2. Qualitative analysis of the DTMs.

Test Case	DTM Complexity [1]	DTM Correlation [2]	DTM Absolute Error [3]
Experiment 1	$N_{DTM} = 10$	1.0000	8.7264×10^{-7}
Experiment 2	$N_{DTM} = 20$	1.0000	6.4552×10^{-8}
Experiment 3	$N_{DTM} = 45$	1.0000	2.9684×10^{-7}

[1] Number of retained Koopman modes. [2] DTM correlation given by Equation (56). [3] DTM absolute error given by Equation (57).

To assess the orthogonality of the Koopman modes computed by KROD algorithm, the orthogonality matrix introduced in Section 3.3 was evaluated for each of the three experiments considered in this study. The orthogonality matrices corresponding to the three investigated experiments are shown, respectively, in both 3D and 2D views in Figure 4a–c. The MAC values calculated using Equation (54) validate the mutual orthogonality of the Koopman modes across all cases.

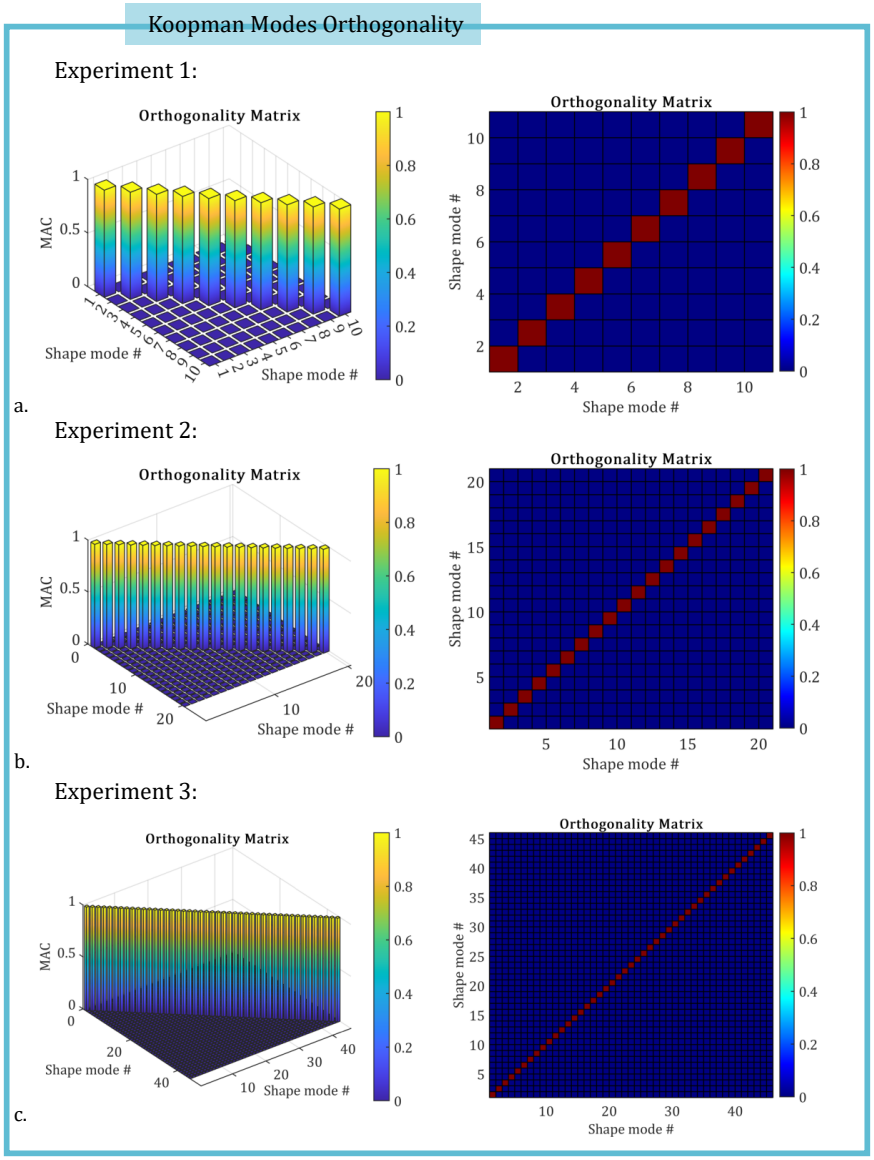


Figure 4. The orthogonality matrices corresponding to the three investigated experiments are shown, respectively, in both 3D and 2D views, in Figures (**a**–**c**). The MAC values, computed according to Equation (54), confirm the mutual orthogonality of the Koopman modes in each case.

The KROD algorithm facilitates the identification of the leading Koopman modes and their corresponding temporal coefficients. To achieve high-fidelity time simulation of the data-driven twin model's temporal parameters, Nonlinear AutoRegressive models with eXogenous inputs (NLARX) were employed (revise Section 3.2. The data-driven twin model, formulated as in Equation (53), was evaluated through time-domain simulation followed by a two-step input–output validation procedure. First, it was trained on the initial two-thirds of the snapshots, with the remaining third used for validation. Then, the training set was reduced to the first third, and the remaining two-thirds were used to evaluate predictive performance. Figures 5–7 present the two-fold input–output validation results for three randomly selected temporal coefficients simulated using the optimal NLARX models for Experiments 1, 2, and 3, respectively. These results demonstrate that the deep learning-based NLARX models accurately replicated the coefficient dynamics, even when the training dataset was reduced to only the first third of the available snapshots. Relevant performance metrics for the deep learning simulations are also provided in the figures.

Data-driven twin models, constructed in the form prescribed by Equation (53), were subsequently evaluated against the corresponding exact solutions to assess the absolute local error, as defined by Equation (58), for each of the three experimental cases under consideration. As illustrated in Figure 8, the response of the twin models exhibits near-perfect correlation with the exact solutions, as evidenced by the consistently low magnitude of the local error.

Figure 9 presents the CPU time required during the online phase of the algorithm, in which the deep learning process was executed, as a function of the optimized modal space dimension for the three experimental cases analyzed. The reported CPU times correspond to two successive computational runs. All simulations and visualizations were performed using custom scripts developed in Matlab R2022a and executed on a system equipped with an Intel i7-7700K processor. The observed computational times are considered acceptable and fall within a reasonable range for practical implementation.

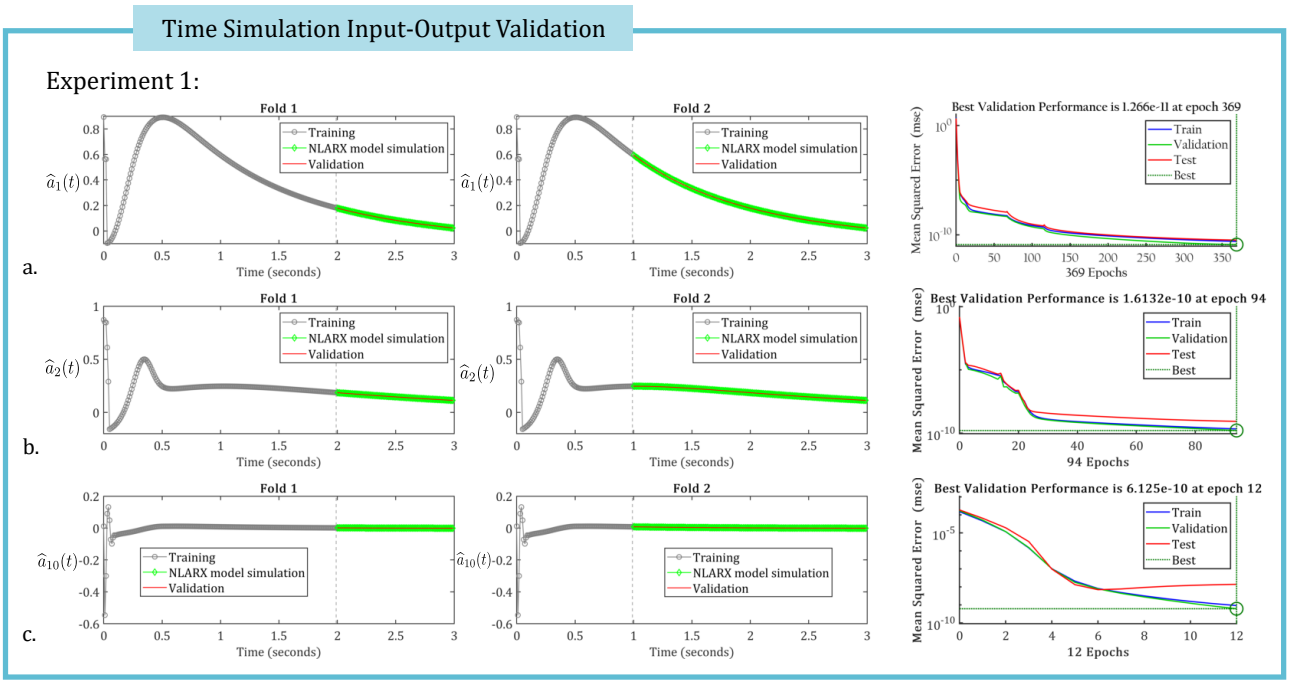


Figure 5. Two-fold input-output validation for the first two (**a**,**b**) temporal coefficients and the tenth (last) (**c**) temporal coefficient during the simulated responses of the optimal NLARX models in Experiment 1. The corresponding deep learning performance metrics are also included in the figure on the right.

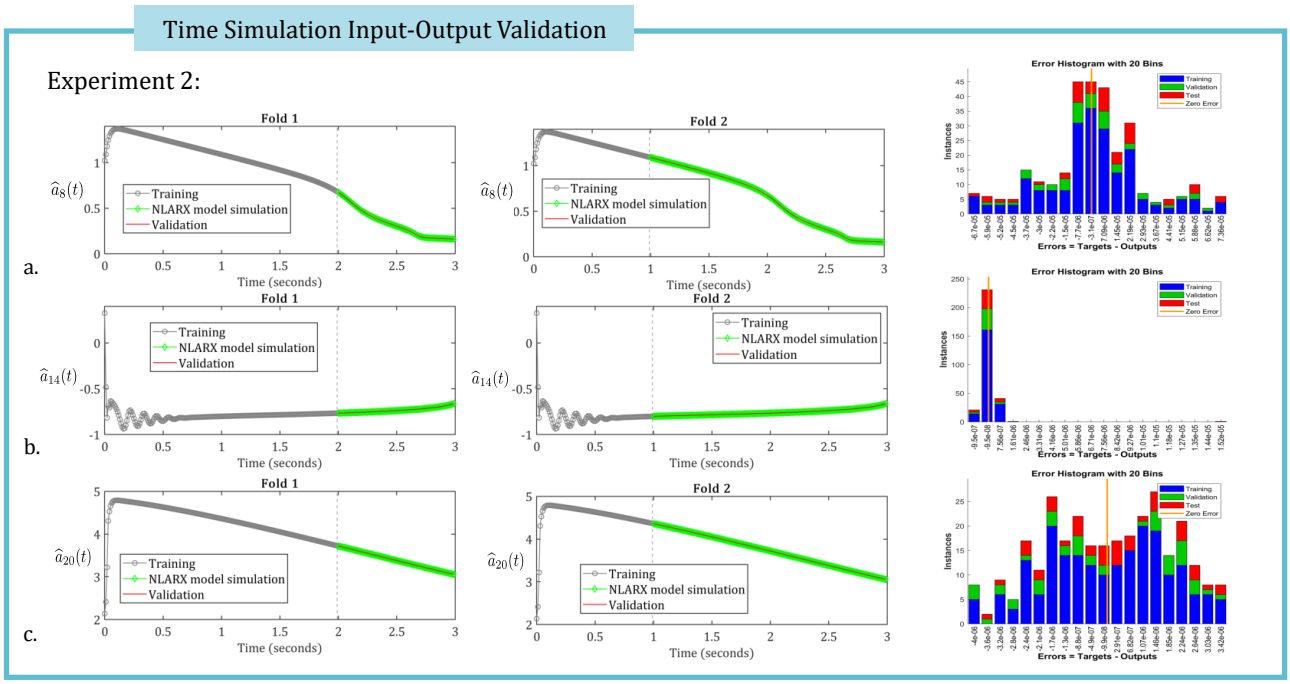


Figure 6. Two-fold input-output validation for the eighth (**a**) temporal coefficient, the fourteenth (**b**) temporal coefficient, and the twentieth (last) (**c**) temporal coefficient during the simulated responses of the optimal NLARX models in Experiment 2. The corresponding deep learning performance metrics are also presented in the figure on the right.

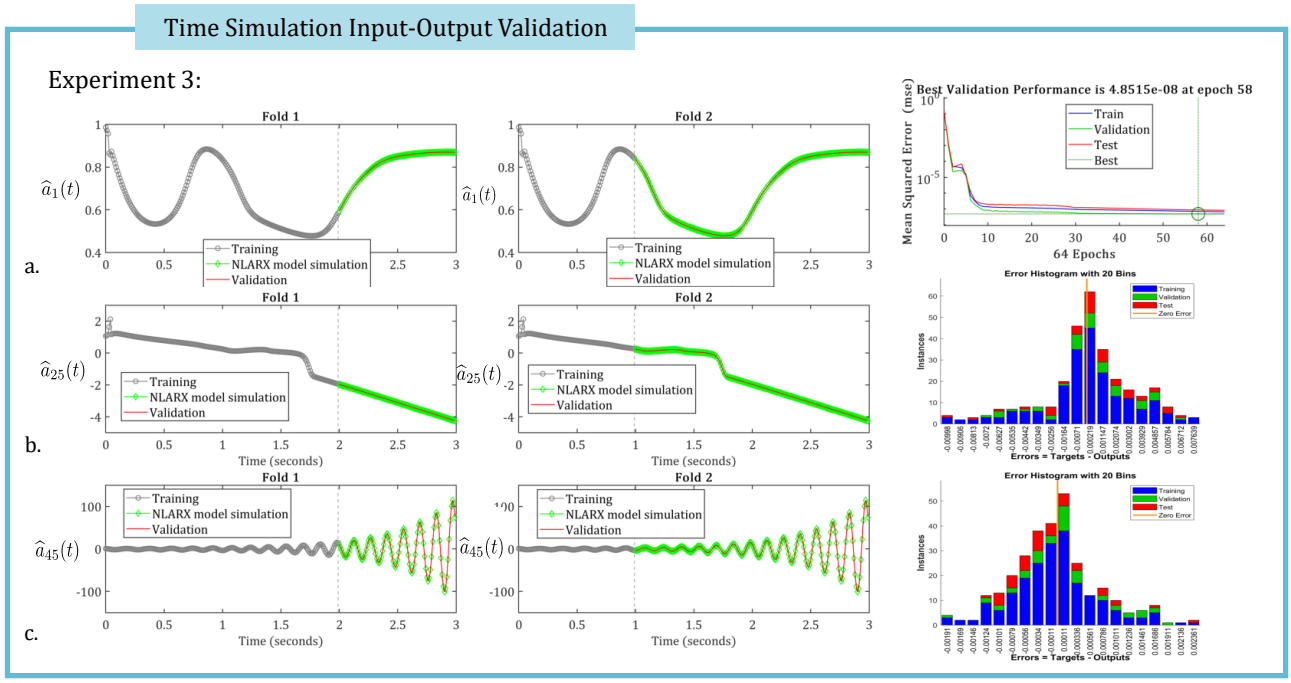


Figure 7. Two-fold input-output validation for the first (**a**) temporal coefficient, the twenty-fifth (**b**) temporal coefficient, and the forty-fifth (last) (**c**) temporal coefficient during the simulated responses of the optimal NLARX models in Experiment 3. Corresponding deep learning performance metrics are presented in the figure on the right.

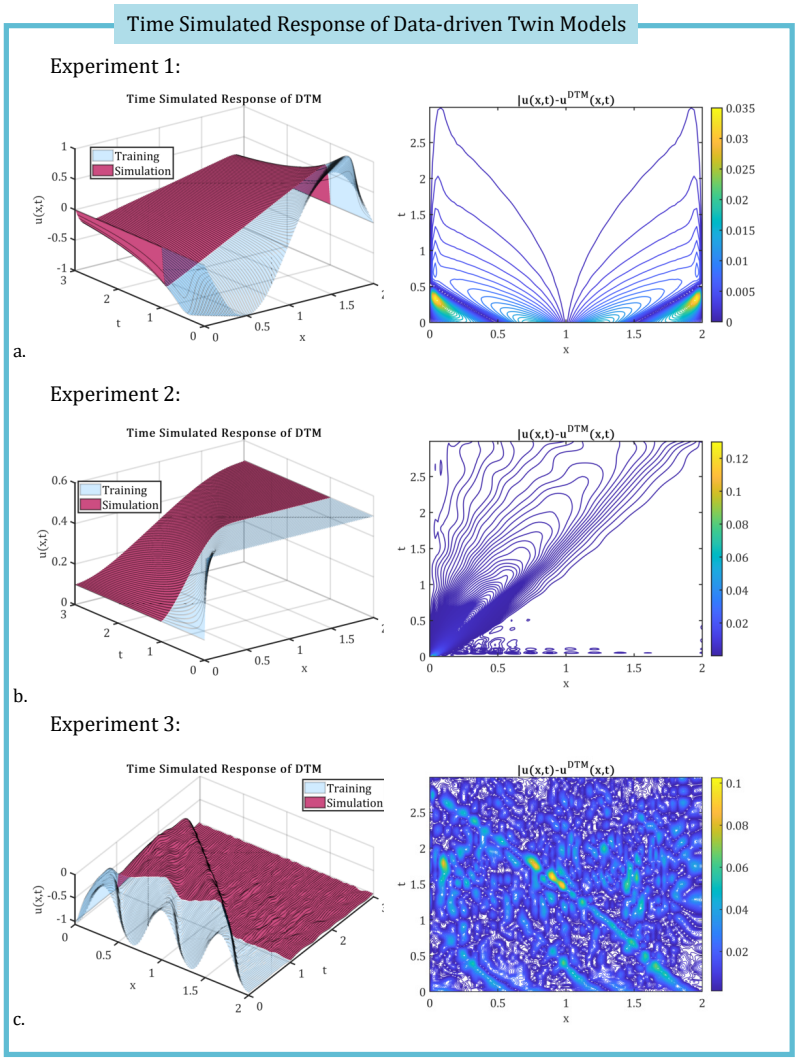


Figure 8. Data-driven twin model responses, constructed via Equation (53), show near-perfect agreement with exact solutions across all experiments. The left subfigures display the time-simulated responses for each experiment, while the right subfigures show the corresponding absolute local errors computed from Equation (58). (**a**) Results for Experiment 1, (**b**) results for Experiment 2, (**c**) results for Experiment 3.

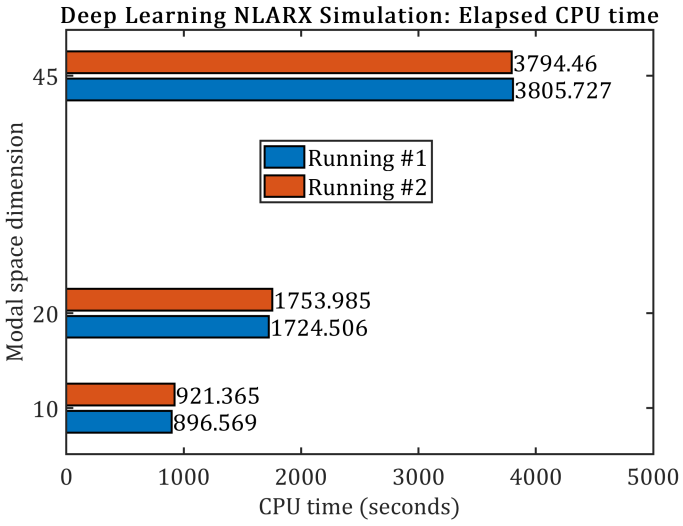


Figure 9. The running CPU time for the online phase of the algorithm with respect to the modal space dimension optimized in the case of the three experiments considered.

6. Conclusions

This study presents a novel framework for developing data-driven twin models with reduced computational complexity for nonlinear dynamical systems governed by partial differential equations, addressing a problem of growing importance in the field of data science. To the best of the author's knowledge, this is the first work to introduce the Koopman Randomized Orthogonal Decomposition (KROD) algorithm in the context of nonlinear PDEs.

This study introduces several methodological advancements that set it apart from classical modal decomposition techniques such as DMD and POD, which typically rely on heuristic or ad hoc criteria for mode selection. In contrast, the proposed KROD framework reduces computational cost while fully automating the identification of dominant modes.

The proposed approach comprises two primary phases. In the first phase, a novel numerical algorithm is developed to construct a reduced-order data twin model, selected through Pareto front analysis, to balance accuracy and complexity. In the second phase, an explainable NLARX deep learning framework is employed to enable real-time, adaptive calibration of the twin model and to predict its dynamic response. The resulting data-driven twin model achieves high fidelity by exhibiting minimal approximation error and strong correlation with the original dataset, thereby providing an accurate and computationally efficient surrogate for the underlying physical system.

A key novelty lies in the first quantitative framework for computing orthogonal Koopman modes using randomized orthogonal projections, which ensures optimal representation in the data space. Moreover, the proposed method integrates explainable deep learning techniques within the modal decomposition process, enabling interpretable and real-time calibration of the reduced-order data twin model.

Importantly, the proposed framework provides a pathway for future extensions into unsupervised learning paradigms, functioning effectively on raw observational data without requiring predefined models. This characteristic promotes broader applicability and generalization across a wide range of fluid dynamics problems and complex dynamical systems. The specific application of the KROD methodology to fluid dynamics scenarios will be the subject of future investigations.

In addition, the framework is readily extensible beyond the current scope. In particular, it holds promise for adaptation to stochastic partial differential equations, systems with sparse or incomplete measurements, and higher-dimensional problems involving 2D and 3D PDEs, directions the author aims to pursue in future research.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The author would like to express sincere gratitude to the anonymous reviewers for their in-depth and constructive feedback, which significantly contributed to improving the clarity and quality of this manuscript. The author also extends heartfelt thanks to Michael I. Navon for over a decade of invaluable collaboration and mentorship in the field of reduced-order modeling.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PDE	Partial Differential Equation
DTM	Data-driven Twin Model
POD	Proper Orthogonal Decomposition

DMD	Dynamic Mode Decomposition
RSVD	Randomized Singular Value Decomposition
k-RSVD	Randomized Singular Value Decomposition of rank k
KROD	Koopman Randomized Orthogonal Decomposition
NLARX	Nonlinear AutoRegressive models with eXogenous inputs
MAC	Modal Assurance Criterion

Appendix A

Appendix A.1. Derivation of the Analytical Solution Using the Cole–Hopf Transformation

The Cole–Hopf transformation is defined by

$$u = -2\nu \frac{1}{\varphi}\frac{\partial \varphi}{\partial x}. \tag{A1}$$

Through analytical treatment, it is found that

$$\frac{\partial u}{\partial t} = \frac{2\nu}{\varphi^2}\left(\frac{\partial \varphi}{\partial t}\frac{\partial \varphi}{\partial x} - \varphi\frac{\partial^2 \varphi}{\partial x \partial t}\right), \quad u\frac{\partial u}{\partial x} = \frac{4\nu^2}{\varphi^3}\frac{\partial \varphi}{\partial x}\left(\varphi\frac{\partial^2 \varphi}{\partial x^2} - \frac{\partial \varphi}{\partial x}\frac{\partial \varphi}{\partial x}\right), \tag{A2}$$

$$\nu\frac{\partial^2 u}{\partial x^2} = -\frac{2\nu^2}{\varphi^3}\left(2\left(\frac{\partial \varphi}{\partial x}\right)^3 - 3\varphi\frac{\partial^2 \varphi}{\partial x^2}\frac{\partial \varphi}{\partial x} + \varphi^2\frac{\partial^3 \varphi}{\partial x^3}\right). \tag{A3}$$

Substituting these expressions into (59) it follows that

$$\frac{\partial \varphi}{\partial x}\left(\frac{\partial \varphi}{\partial t} - \nu\frac{\partial^2 \varphi}{\partial x^2}\right) = \varphi\left(\frac{\partial^2 \varphi}{\partial x \partial t} - \nu\frac{\partial^3 \varphi}{\partial x^3}\right) = \varphi\frac{\partial}{\partial x}\left(\frac{\partial \varphi}{\partial t} - \nu\frac{\partial^2 \varphi}{\partial x^2}\right). \tag{A4}$$

Relation (A4) indicates that if φ solves the heat equation, then $u(x,t)$ given by the Cole–Hopf transformation (A1) solves the viscous Burgers Equation (59). Thus, the viscous Burgers Equation (59) is recast into the following:

$$\begin{cases} \frac{\partial \varphi}{\partial t} - \nu\frac{\partial^2 \varphi}{\partial x^2} = 0, \quad x \in R, t > 0, \nu > 0, \\ \varphi(x,0) = \varphi_0(x) = e^{-\int_0^x \frac{u_0(\xi)}{2\nu} d\xi}, x \in \mathbb{R}. \end{cases} \tag{A5}$$

Taking the Fourier transform with respect to x for both the heat equation and the initial condition (A5), it is found that

$$\varphi(x,t) = \frac{1}{2\sqrt{\pi\nu t}}\int_{-\infty}^{\infty} \varphi_0(\xi)\, e^{-\frac{(x-\xi)^2}{4\nu t}} d\xi. \tag{A6}$$

Using the Cole–Hopf transformation (A1), the analytic solution to the problem (59) is obtained in the following form:

$$u(x,t) = \frac{\int_{-\infty}^{\infty} \frac{x-\xi}{t}\varphi_0(\xi) e^{-\frac{(x-\xi)^2}{4\nu t}} d\xi}{\int_{-\infty}^{\infty} \varphi_0(\xi) e^{-\frac{(x-\xi)^2}{4\nu t}} d\xi}. \tag{A7}$$

Appendix A.2. Derivation of the Numerical Exact Solution Using the Gauss–Hermite Quadrature

To compute the exact solution (A7) corresponding to the initial conditions of the three test cases (60)–(62), the Gauss–Hermite quadrature technique [82] is employed.

The Gauss–Hermite quadrature approximates the value of integrals of the following kind:

$$\int_{-\infty}^{\infty} f(z)\, e^{-z^2} dz \approx \sum_{i=1}^{n} w_i f(x_i), \tag{A8}$$

where n represents the number of sample points used, x_i are the roots of the Hermite polynomial $H_n(x)$, and the associated weights w_i are given by

$$w_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 (H_{n-1}(x_i))^2}, \quad i = 1, \ldots, n. \tag{A9}$$

In Experiment 1, the initial condition is defined by Equation (60) as

$$\varphi_0(x) = e^{-\frac{1}{2\nu}\int_0^x u_0(\xi)d\xi} = e^{-\frac{1}{2\nu}\int_0^x -\sin(\pi\xi)d\xi} = e^{\frac{1}{2\nu\pi}} \cdot e^{-\frac{\cos(\pi x)}{2\nu\pi}}. \tag{A10}$$

and the exact solution (A7) is obtained in the form

$$u(x,t) = \frac{\int_{-\infty}^{\infty} \frac{x-\xi}{t} e^{-\frac{\cos(\pi\xi)}{2\nu\pi}} \cdot e^{-\left(\frac{x-\xi}{\sqrt{4\nu t}}\right)^2} d\xi}{\int_{-\infty}^{\infty} e^{-\frac{\cos(\pi\xi)}{2\nu\pi}} \cdot e^{-\left(\frac{x-\xi}{\sqrt{4\nu t}}\right)^2} d\xi}. \tag{A11}$$

With the variable change

$$z = \frac{x-\xi}{\sqrt{4\nu t}}, \tag{A12}$$

the exact solution to the Burgers equation model (59) with the initial condition (60) is equivalent to the following form:

$$u(x,t) = \frac{\int_{-\infty}^{\infty} 4\nu z\, e^{-\frac{1}{2\nu\pi}\cos\left[\pi\left(x - z\sqrt{4\nu t}\right)\right]} e^{-z^2} dz}{\int_{-\infty}^{\infty} \sqrt{4\nu t}\, e^{-\frac{1}{2\nu\pi}\cos\left[\pi\left(x - z\sqrt{4\nu t}\right)\right]} e^{-z^2} dz}. \tag{A13}$$

Similarly, by applying the variable transformation (A12), the exact solutions for the subsequent two experiments are derived as follows.

In the case of Experiment 2, the exact solution to the Burgers equation model (59) with the initial condition (61) is

$$u(x,t) = \frac{\int_{-\infty}^{\infty} 4\nu z\, e^{-\frac{u_R}{2\nu}\left(x - z\sqrt{4\nu t}\right)} e^{-z^2} dz}{\int_{-\infty}^{\infty} \sqrt{4\nu t}\, e^{-\frac{u_R}{2\nu}\left(x - z\sqrt{4\nu t}\right)} e^{-z^2} dz}. \tag{A14}$$

In the case of Experiment 3, the exact solution to the Burgers equation model (59) with the initial condition (62) is

$$u(x,t) = \frac{\int_{-\infty}^{\infty} 4\nu z\, e^{\frac{1}{4\nu}\left(\frac{1}{3\pi}\sin\left(3\pi\left(x - z\sqrt{4\nu t}\right)\right) + x - z\sqrt{4\nu t}\right)} e^{-z^2} dz}{\int_{-\infty}^{\infty} \sqrt{4\nu t}\, e^{\frac{1}{4\nu}\left(\frac{1}{3\pi}\sin\left(3\pi\left(x - z\sqrt{4\nu t}\right)\right) + x - z\sqrt{4\nu t}\right)} e^{-z^2} dz}. \tag{A15}$$

To obtain the numerical exact solutions for the three test cases, both the numerator and denominator are evaluated using Gauss–Hermite quadrature (A8).

Appendix B

The following presents Algorithm A1, which outlines the routine employed to compute the randomized singular value decomposition of rank k (k-RSVD). Algorithm A2 details the Koopman Randomized Orthogonal Decomposition (KROD) procedure.

Algorithm A1 (k-RSVD) Randomized Singular Value Decomposition of Rank k

Initial data: $V_0 \in \mathbb{R}^{N_x \times N_t}$, integer target rank $k \geq 2$ and $k \leq N_t$.

1. Generate a Gaussian random test matrix $M \in \mathbb{R}^{N_t \times k}$ with independent and identically distributed (i.i.d.) standard normal entries.
2. Compute the compressed sampled range matrix $Q = V_0 M \in \mathbb{R}^{N_x \times k}$.
3. Produce Gram–Schmidt orthonormalisation of sampling matrix Q.
4. Form the projected matrix $P = Q^H V_0 \in \mathbb{R}^{k \times N_t}$.
5. Compute the economy-size singular value decomposition of P:

$$P = T \Sigma W^H,$$

where $T \in \mathbb{R}^{k \times k}$ is orthogonal, $\Sigma \in \mathbb{R}^{k \times k}$ is diagonal, and $W \in \mathbb{R}^{N_t \times k}$ has orthonormal columns.

6. Compute the approximate left singular vectors as:

$$U = QT \in \mathbb{R}^{N_x \times k}.$$

Output: Procedure returns $U \in \mathbb{R}^{N_x \times k}$, $\Sigma \in \mathbb{R}^{k \times k}$, $W \in \mathbb{R}^{N_t \times k}$.

Algorithm A2 (KROD) Koopman Randomized Orthogonal Decomposition

Initial data:

- A sequence of solution snapshots of the PDE, arranged into two time-shifted data matrices, is defined as

$$V_0 = \begin{bmatrix} u_0 & u_1 & \dots & u_{N_t-1} \end{bmatrix} \in \mathbb{R}^{N_x \times N_t}, \; V_1 = \begin{bmatrix} u_1 & u_2 & \dots & u_{N_t} \end{bmatrix} \in \mathbb{R}^{N_x \times N_t}.$$

- A sequence of integer target ranks $\{k_1, k_2, \dots, k_N\}$ with $2 < k_1 < \cdots < k_N \leq N_t$.

Offline phase:

1. For $k \in \{k_1, k_2, \dots, k_N\}$:
   (a) Produce the randomized singular value approximation of rank k for V_0 as

$$V_0 \approx U \Sigma W^H.$$

   The procedure for computing the k-RSVD is outlined in Algorithm A1.
   (b) Compute the matrix operator $\mathcal{S}$ according to Equation (30) as

$$\mathcal{S} = U^H \left(V_1 W \Sigma^{-1} \right).$$

   (c) Compute the Hermitian positive semi-definite Gram matrix as

$$\mathcal{G} = \mathcal{S}^H \mathcal{S}.$$

   (d) Solve the eigenvalue problem $\mathcal{G}X = X\Lambda$ and obtain dynamic modes as $\Phi = UX$.
   (e) Project V_0 onto the subspace spanned by the orthonormal basis Φ to obtain the coefficient matrix as $A = \Phi^H V_0$.
2. Save corresponding model triplets as $\mathcal{T} = \{T_1, T_2, \dots, T_N\}$, where $T_i = (\Phi_i, A_i, k_i)$.

Algorithm A2 *Cont.*

3. Multi-objective twin-model selection by solving the optimization problem (46) by Pareto front analysis:
   (a) For each $T_i = (\Phi_i, A_i, k_i) \in \mathcal{T}$, compute the reduced-order approximation as
   $$u^{(T_i)} = \sum_{j=1}^{k_i} a_j(t)\phi_j(x).$$
   (b) Compute the reconstruction error as
   $$E(T_i) = \|u - u^{(T_i)}\|_2.$$
   (c) Compute the normalized cosine similarity as
   $$C(T_i) = \frac{\|u^H u^{(T_i)}\|_2^2}{\|u^H u\|_2 \cdot \|(u^{(T_i)})^H u^{(T_i)}\|_2}.$$
   (d) Construct the Pareto front $\mathcal{T}^*$ by selecting all non-dominated triplets as
   $$\mathcal{T}^* = \{T_i \in \mathcal{T} \mid \nexists T_j \in \mathcal{T} \text{ such that } E(T_j) \leq E(T_i), C(T_j) \geq C(T_i), \text{ and strict in at least one}\}.$$
   (e) Return the Pareto-optimal set of Koopman model triplet as
   $$\mathcal{T}^* = (\Phi_{\text{DTM}}, A_{\text{DTM}}, N_{\text{DTM}}),$$
   consisting of the Koopman leading modes, the associated modal amplitudes, and the rank of the Koopman basis, i.e., the number of terms in Koopman randomized orthogonal decomposition in form of Equation (47).

Online phase:

4. Construct the nonlinear mapping for approximating the system dynamics by evolving the temporal coefficients via the optimal deep learning-based NLARX model Equation (51). **Output:** The reduced-order data-driven twin model for the PDE solution is obtained in the form given by Equation (53).

References

1. Grieves, M.; Vickers, J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems*; Springer: Cham, Switzerland, 2017; pp. 85–113.
2. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y.C. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2405–2415. [CrossRef]
3. Glaessgen, E.; Stargel, D. The digital twin paradigm for future NASA and U.S. Air Force vehicles. In Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Honolulu, HI, USA, 23–26 April 2012.
4. Rasheed, A.; San, O.; Kvamsdal, T. Digital Twin: Values, Challenges and Enablers From a Modeling Perspective. *IEEE Access* **2020**, *8*, 21980–22012. [CrossRef]
5. Holmes, P.; Lumley, J.; Berkooz, G. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*; Cambridge University Press: Cambridge, UK, 1996.
6. Rowley, C.W.; Mezić, I.; Bagheri, S.; Schlatter, P.; Henningson, D.S. Reduced-order models for flow control: Balanced models and Koopman modes. In *Seventh IUTAM Symposium on Laminar-Turbulent Transition*; IUTAM Bookseries; Springer: Dordrecht, The Netherlands, 2010; Volume 18, pp. 43–50.
7. Frederich, O.; Luchtenburg, D.M. Modal analysis of complex turbulent flow. In Proceedings of the 7th International Symposium on Turbulence and Shear Flow Phenomena (TSFP-7), Ottawa, ON, Canada, 28–31 July 2011.
8. Kutz, J.; Sashidhar, D.; Sahba, S.; Brunton, S.; McDaniel, A.; Wilcox, C. Physics-informed machine-learning for modeling aero-optics. In Proceedings of the International Conference On Applied Optical Metrology IV, San Diego, CA, USA, 1–5 August 2021; Volume 11817, pp. 70–77.

9. Schmid, P.J.; Violato, D.; Scarano, F. Decomposition of Time-resolved Tomographic PIV. In Proceedings of the 16th International Symposium on Applications of Laser Techniques to Fluid Mechanics, Lisbon, Portugal, 9–12 July 2012.
10. Mezić, I. Koopman Operator, Geometry, and Learning of Dynamical Systems. *Not. Am. Math. Soc.* **2021**, *68*, 1087–1105. [CrossRef]
11. Ahmed, S.E.; San, O.; Bistrian, D.A.; Navon, I.M. Sampling and resolution characteristics in reduced order models of shallow water equations: Intrusive vs nonintrusive. *Int. J. Numer. Methods Fluids* **2020**, *92*, 992–1036. [CrossRef]
12. Xiao, D.; Fang, F.; Heaney, C.E.; Navon, I.M.; Pain, C.C. A domain decomposition method for the non-intrusive reduced order modelling of fluid flow. *Comput. Methods Appl. Mech. Eng.* **2019**, *354*, 307–330. [CrossRef]
13. Loéve, M. Sur la covariance d'une fonction aleatoire. *Comptes Rendus Hebd. Séances L'Académie Sci. Paris* **1945**, *220*, 295–296.
14. Karhunen, K. Zur Spektraltheorie stochastischer Prozesse. *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.* **1946**, *34*, 1–7.
15. Lumley, J.L. *Stochastic Tools in Turbulence*; Academic Press: New York, NY, USA, 1970.
16. Iliescu, T. ROM Closures and Stabilizations for Under-Resolved Turbulent Flows. In Proceedings of the 2022 Spring Central Sectional Meeting, Virtual, 26–27 March 2022.
17. Wang, Z.; Akhtar, I.; Borggaard, J.; Iliescu, T. Proper orthogonal decomposition closure models for turbulent flows: A numerical comparison. *Comput. Methods Appl. Mech. Eng.* **2012**, *10*, 237–240. [CrossRef]
18. Xiao, D.; Heaney, C.E.; Fang, F.; Mottet, L.; Hu, R.; Bistrian, D.A.; Aristodemou, D.; Navon, I.M.; Pain, C.C. A domain decomposition non-intrusive reduced order model for turbulent flows. *Comput. Fluids* **2019**, *182*, 15–27. [CrossRef]
19. Gunzburger, M.; Iliescu, T.; Mohebujjaman, M.; Schneier, M. Nonintrusive stabilization of reduced order models for uncertainty quantification of time-dependent convection-dominated flows. *arXiv* **2018**, arXiv:1810.08746.
20. San, O.; Staples, A.; Wang, Z.; Iliescu, T. Approximate deconvolution large eddy simulation of a barotropic ocean circulation model. *Ocean Model.* **2011**, *40*, 120–132. [CrossRef]
21. San, O.; Iliescu, T. A stabilized proper orthogonal decomposition reduced-order model for large scale quasigeostrophic ocean circulation. *Adv. Comput. Math.* **2014**, *41*, 1289–1319. [CrossRef]
22. Osth, J.; Noack, B.R.; Krajnovic, S.; Barros, D.; Boree, J. On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high-Reynolds-number flow over an Ahmed body. *J. Fluid Mech.* **2014**, *747*, 518–544. [CrossRef]
23. Xiao, D.; Du, J.; Fang, F.; Pain, C.; Li, J. Parameterised non-intrusive reduced order methods for ensemble Kalman filter data assimilation. *Comput. Fluids* **2018**, *177*, 69–77. [CrossRef]
24. Sierra, C.; Metzler, H.; Muller, M.; Kaiser, E. Closed-loop and congestion control of the global carbon-climate system. *Clim. Change* **2021**, *165*, 1–24. [CrossRef]
25. Xiao, D.; Heaney, C.E.; Mottet, L.; Fang, F.; Lin, W.; Navon, I.M.; Guo, Y.; Matar, O.K.; Robins, A.G.; Pain, C.C. A reduced order model for turbulent flows in the urban environment using machine learning. *Build. Environ.* **2019**, *148*, 323–337. [CrossRef]
26. San, O.; Maulik, R.; Ahmed, M. An artificial neural network framework for reduced order modeling of transient flows. *Commun. Nonlinear Sci. Numer. Simul.* **2019**, *77*, 271–287. [CrossRef]
27. Kaptanoglu, A.; Callaham, J.; Hansen, C.; Brunton, S. Machine Learning to Discover Interpretable Models in Fluids and Plasmas. In Proceedings of the APS March Meeting 2022, Chicago, IL, USA, 14–18 March 2022.
28. Owens, K.; Kutz, J. Data-driven discovery of governing equations for coarse-grained heterogeneous network dynamics. *SIAM J. Appl. Dyn. Syst.* **2023**, *22*, 2601–2623. [CrossRef]
29. Cheng, M.; Fang, F.; Pain, C.; Navon, I. Data-driven modelling of nonlinear spatio-temporal fluid flows using a deep convolutional generative adversarial network. *Comput. Methods Appl. Mech. Eng.* **2020**, *365*, 113000. [CrossRef]
30. Dumon, A.; Allery, C.; Ammar, A. Proper Generalized Decomposition method for incompressible Navier-Stokes equations with a spectral discretization. *Appl. Math. Comput.* **2013**, *219*, 8145–8162. [CrossRef]
31. Dimitriu, G.; Ştefănescu, R.; Navon, I.M. Comparative numerical analysis using reduced-order modeling strategies for nonlinear large-scale systems. *J. Comput. Appl. Math.* **2017**, *310*, 32–43. [CrossRef]
32. Schmid, P.J.; Sesterhenn, J. Dynamic mode decomposition of numerical and experimental data. In Proceedings of the 61st Annual Meeting of the APS Division of Fluid Dynamics, American Physical Society, San Antonio, TX, USA, 23–25 November 2008; Volume 53.
33. Schmid, P. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **2010**, *656*, 5–28. [CrossRef]
34. Koopman, B. Hamiltonian systems and transformations in Hilbert space. *Proc. Natl. Acad. Sci. USA* **1931**, *17*, 315–318. [CrossRef]
35. Koopman, B.; von Neumann, J. Dynamical systems of continuous spectra. *Proc. Natl. Acad. Sci. USA* **1932**, *18*, 255–263. [CrossRef]
36. Percic, M.; Zelenika, S.; Mezić, I. Artificial intelligence-based predictive model of nanoscale friction using experimental data. *Friction* **2021**, *9*, 1726–1748. [CrossRef]
37. Pant, P.; Doshi, R.; Bahl, P.; Farimani, A.B. Deep learning for reduced order modelling and efficient temporal evolution of fluid simulations. *Phys. Fluids* **2021**, *33*, 107101. [CrossRef]
38. Tu, J.H.; Rowley, C.W.; Luchtenburg, D.M.; Brunton, S.L.; Kutz, J.N. On dynamic mode decomposition: Theory and applications. *J. Comput. Dyn.* **2014**, *1*, 391–421. [CrossRef]

39. Brunton, S.L.; Brunton, B.W.; Proctor, J.L.; Kutz, J.N. Koopman Invariant Subspaces and Finite Linear Representations of Nonlinear Dynamical Systems for Control. *PLoS ONE* **2016**, *11*, e0150171. [CrossRef] [PubMed]
40. Kutz, J.N.; Fu, X.; Brunton, S.L. Multiresolution dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **2016**, *15*, 713–735. [CrossRef]
41. Li, B.; Ma, Y.; Kutz, J.; Yang, X. The Adaptive Spectral Koopman Method for Dynamical Systems. *SIAM J. Appl. Dyn. Syst.* **2023**, *22*, 1523–1551. [CrossRef]
42. Chen, K.K.; Tu, J.H.; Rowley, C.W. Variants of dynamic mode decomposition: Boundary condition, Koopman and Fourier analyses. *J. Nonlinear Sci.* **2012**, *22*, 887–915. [CrossRef]
43. Jovanovic, M.R.; Schmid, P.J.; Nichols, J.W. Sparsity-promoting dynamic mode decomposition. *Phys. Fluids* **2014**, *26*, 024103. [CrossRef]
44. Williams, M.O.; Kevrekidis, I.G.; Rowley, C.W. A data-driven approximation of the Koopman operator: Extending Dynamic Mode Decomposition. *J. Nonlinear Sci.* **2015**, *25*, 1307–1346. [CrossRef]
45. Noack, B.R.; Stankiewicz, W.; Morzynski, M.; Schmid, P.J. Recursive dynamic mode decomposition of transient and post-transient wake flows. *J. Fluid Mech.* **2016**, *809*, 843–872. [CrossRef]
46. Proctor, J.L.; Brunton, S.L.; Kutz, J.N. Dynamic mode decomposition with control. *SIAM J. Appl. Dyn. Syst.* **2016**, *15*, 142–161. [CrossRef]
47. Erichson, N.B.; Donovan, C. Randomized low-rank Dynamic Mode Decomposition for motion detection. *Comput. Vis. Image Underst.* **2016**, *146*, 40–50. [CrossRef]
48. Ahmed, S.E.; Dabaghian, P.H.; San, O.; Bistrian, D.A.; Navon, I.M. Dynamic mode decomposition with core sketch. *Phys. Fluids* **2022**, *34*, 066603. [CrossRef]
49. Goldschmidt, A.; Kaiser, E.; Dubois, J.; Brunton, S.; Kutz, J. Bilinear dynamic mode decomposition for quantum control. *New J. Phys.* **2021**, *23*, 033035. [CrossRef]
50. Le Clainche, S.; Vega, J.M. Higher Order Dynamic Mode Decomposition. *SIAM J. Appl. Dyn. Syst.* **2017**, *16*, 882–925. [CrossRef]
51. Bistrian, D.A.; Navon, I.M. An improved algorithm for the shallow water equations model reduction: Dynamic Mode Decomposition vs POD. *Int. J. Numer. Methods Fluids* **2015**, *78*, 552–580. [CrossRef]
52. Bistrian, D.A.; Navon, I.M. The method of dynamic mode decomposition in shallow water and a swirling flow problem. *Int. J. Numer. Methods Fluids* **2017**, *83*, 73–89. [CrossRef]
53. Bistrian, D.A.; Navon, I.M. Randomized dynamic mode decomposition for nonintrusive reduced order modelling. *Int. J. Numer. Methods Eng.* **2017**, *112*, 3–25. [CrossRef]
54. Bistrian, D.A.; Navon, I.M. Efficiency of randomised dynamic mode decomposition for reduced order modelling. *Int. J. Comput. Fluid Dyn.* **2018**, *32*, 88–103. [CrossRef]
55. Bistrian, D.A.; Dimitriu, G.; Navon, I.M. Modeling dynamic patterns from COVID-19 data using randomized dynamic mode decomposition in predictive mode and ARIMA. In Proceedings of the AIP Conference Proceedings, Albena, Bulgaria, 24–29 June 2020; Volume 2302, p. 080002.
56. Bistrian, D.A.; Dimitriu, G.; Navon, I.M. Processing epidemiological data using dynamic mode decomposition method. In Proceedings of the AIP Conference Proceedings, Albena, Bulgaria, 20–25 June 2019; Volume 2164, p. 080002.
57. Koopman, B. On distributions admitting a sufficient statistic. *Trans. Am. Math. Soc.* **1936**, *39*, 399–400. [CrossRef]
58. Lasota, A.; Mackey, M.C. *Probabilistic Properties of Deterministic Systems*; Cambridge University Press: Cambridge, UK, 1985.
59. Lasota, A.; Mackey, M. *Chaos, Fractals, and Noise. Stochastic Aspects of Dynamics*; Springer: Berlin/Heidelberg, Germany, 1994.
60. Mezić, I.; Banaszuk, A. Comparison of systems with complex behavior. *Phys. D Nonlinear Phenom.* **2004**, *197*, 101–133. [CrossRef]
61. Mezić, I. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dyn.* **2005**, *41*, 309–325. [CrossRef]
62. Rowley, C.W.; Mezić, I.; Bagheri, S.; Schlatter, P.; Henningson, D.S. Spectral analysis of nonlinear flows. *J. Fluid Mech.* **2009**, *641*, 115–127. [CrossRef]
63. Perron, O. Zur Theorie der Matrices. *Math. Ann.* **1907**, *64*, 248–263. [CrossRef]
64. Frobenius, G. Uber Matrizen aus positiven Elementen. *Sitzungsberichte Königlich Preuss. Akad. Wiss.* **1908**, *1*, 471–476.
65. Ruelle, D.P. *Chance and Chaos*; Princeton University Press: Princeton, NJ, USA, 1993; Volume 110.
66. Mezić, I. On numerical approximations of the Koopman operator. *Mathematics* **2022**, *10*, 1180. [CrossRef]
67. Halko, N.; Martinsson, P.G.; Tropp, J.A. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.* **2011**, *53*, 217–288. [CrossRef]
68. Surasinghe, S.; Bollt, E.M. Randomized Projection Learning Method for DynamicMode Decomposition. *Mathematics* **2021**, *9*, 2803. [CrossRef]
69. Qiu, Y.; Sun, W.; Zhou, G.; Zhao, Q. Towards efficient and accurate approximation: Tensor decomposition based on randomized block Krylov iteration. *Signal Image Video Process* **2024**, *18*, 6287–6297. [CrossRef]

70. Bistrian, D.A. High-fidelity digital twin data models by randomized dynamic mode decomposition and deep learning with applications in fluid dynamics. *Modelling* **2022**, *3*, 314–332. [CrossRef]
71. Golub, G.H.; Loan, C.F.V. *Matrix Computations*, 4th ed.; Johns Hopkins University Press: Baltimore, MD, USA, 2013.
72. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; John Wiley & Sons: Hoboken, NJ, USA, 2001.
73. Juditsky, A.; Hjalmarsson, H.; Benveniste, A.; Deylon, B.; Ljung, L.; Sloberg, J.; Zhang, Q. Nonlinear black-box models in system identification: Mathematical foundations. *Automatica* **1995**, *31*, 1725–1750. [CrossRef]
74. Ljung, L. *System Identification: Theory for the User*; Information and System Sciences Series; Prentice Hall: Hoboken, NJ, USA, 1999.
75. Nelles, O. *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*; Springer: Berlin/Heidelberg, Germany, 2001.
76. Liu, G.; Kadirkamanathan, V.; Billings, S. Predictive control for non-linear systems using neural networks. *Int. J. Control* **1998**, *71*, 1119–1132. [CrossRef]
77. Tieleman, T.; Hinton, G. *Divide the Gradient by a Running Average of Its Recent Magnitude*; Neural Networks for Machine Learning; Coursera: Mountain View, CA, USA, 2012.
78. Wang, Z.; Xiao, D.; Fang, F.; Govindan, R.; Pain, C.; Guo, Y. Model identification of reduced order fluid dynamics systems using deep learning. *Int. J. Numer. Methods Fluids* **2018**, *86*, 255–268. [CrossRef]
79. Ewins, D. *Modal Testing: Theory, Practice and Application*; Research Studies Press Ltd.: Totnes Devon, UK, 2000.
80. Hopf, E. The partial differential equation $u_t + uu_x = vu_{xx}$. *Commun. Pure Appl. Math.* **1950**, *3*, 201–230. [CrossRef]
81. Cole, J.D. On a quasi-linear parabolic equation occurring in aerodynamics. *Q. Appl. Math.* **1951**, *9*, 225–236. [CrossRef]
82. Brass, H.; Petras, K. *Quadrature Theory. The Theory of Numerical Integration on a Compact Interval*; American Mathematical Society: Providence, RI, USA, 2011.
83. Noack, B.R.; Morzynski, M.; Tadmor, G. *Reduced-Order Modelling for Flow Control*; Springer: Berlin/Heidelberg, Germany, 2011.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.