

Progress and future directions on predictive data-driven reduced-order modeling for Electric Propulsion digital twins

IEPC-2024-501

*Presented at the 38th International Electric Propulsion Conference
Pierre Baudis Convention Center • Toulouse, France
June 23-28, 2024*

Maryam Reza¹, Farbod Faraji², Aaron Knoll³

Plasma Propulsion Laboratory, Imperial College London, London, United Kingdom

and

J. Nathan Kutz⁴

University of Washington, Seattle, United States

Abstract: Reduced-order models (ROMs) are today an indispensable element of contemporary science and engineering. ROMs surpass the capabilities of traditional high-fidelity simulations by striking a unique, optimal balance between cost efficiency and accuracy, opening the door to fast and effective numerical analyses. This transforms the applicability of modeling with regard to both scientific discovery and engineering developments. The paradigm of physics-informed/physics-inspired machine learning (ML) serves to further transcend the ROMs in terms of their interpretability and generalizability. Predictive, interpretable, and generalizable ML-enabled ROMs facilitate scientific research into yet-unresolved complex phenomena. On the applied side, one of the strong motivators for such ROMs is the transformative technology of digital twins. The Electric Propulsion (EP) sector and the broader field of plasma physics can benefit significantly from advances in ML-enabled physics-inspired ROMs that feature low-dimensionality and sparsity. This paper aims to cast light on reduced-order modeling within the context of plasma systems relevant to EP. Following on introductory discussions on the practical relevance of ROMs, particularly in the framework of envisioned EP digital twins, we describe the two central elements of ROM development, i.e., finding a proper reduced coordinate system, and describing the system's dynamics within that reduced coordinate. We emphasize how ML and data-driven (DD) approaches can majorly aid each of these elements. Taking the example of a plasma discharge that represents a radial-azimuthal Hall thruster configuration, we demonstrate how autoencoder networks can find a highly low-dimensional space that fully represents the high-dimensional plasma state. We discuss multiple possibilities for ML/DD learning of the time dynamics in the low-dimensional latent space of the autoencoder. We finally present dynamics forecasting results from a fast and predictive ROM obtained using the optimized dynamic mode decomposition method, which amounts to a special case of combining a linear shallow autoencoder with a linear operator for describing the time dynamics.

¹ Electric Propulsion Researcher, Department of Aeronautics, m.reza20@imperial.ac.uk.

² Electric Propulsion Researcher, Department of Aeronautics, f.faraji20@imperial.ac.uk.

³ Reader in Spacecraft Engineering, Department of Aeronautics, a.knoll@imperial.ac.uk.

⁴ Professor, Department of Applied Mathematics and Electrical and Computer Engineering, kutz@uw.edu.

I. Cost-efficient, predictive models as enabling computational capabilities for EP digital twins

PLASMAS span across vast length scales, from the cosmological down to the laboratory size. Plasmas manifest in many natural phenomena, from interstellar nebulae to lightning, and play a central role in a myriad of human-made technologies, such as fusion reactors and plasma propulsion devices for spacecrafts.

Despite the widespread occurrence and applications of plasmas, and decades of rigorous research since the mid-20th century, many unknowns yet persist regarding the intricate and complex plasma behaviors [1]-[5]. A primary reason of this enduring challenge is that the necessary predictive and cost-efficient simulation models and analysis toolset are still lacking to enable comprehensive studies of the multiscale, multidimensional, and often multiphysics phenomena in plasmas. The absence of these numerical capabilities restricts our ability to justify and generalize theoretical predictions and experimental observations, further hindering the advancement in plasma science. The knowledge gap in plasma physics represents a significant shortcoming, as the unresolved questions have great implications on our ability to understand, predict, and control the plasmas. Answering the persisting questions will revolutionize the fundamental science and will have major applied significance as well.

Due to the complexities of plasmas in many applications, their detailed studies often require employing high-fidelity kinetic simulations based on, e.g., the particle-in-cell (PIC) approach. In theory, a 3D PIC simulation can offer comprehensive insights into the plasma processes, in part, by the merit of its ability to capture all relevant phenomena. Practically, however, a 3D PIC simulation has an enormous computational cost [6], which makes most studies of scientific and applied relevance using this approach impractical or computationally intractable. The computational cost of the conventional PIC approach [7] is significant for studying real-size plasma systems, even in 2D configurations [8]. In this regard, practically meaningful simulation times are most often achieved by incorporating non-physical speedup scaling factors.

It follows that the quest for cost-efficient and self-consistently predictive plasma models is well warranted and indeed crucial to further advance plasma science and engineering.

Reduced-order models serve as viable solutions to the challenges of computational plasma modeling. A reduced-order model (ROM) can be defined as a simplified representation of a full-order model (FOM) that retains essential dynamics of the original system while significantly reducing computational complexity. This simplification is achieved by reducing the number of degrees of freedom (dimensionality), variables, or equations describing the system.

Traditionally, ROMs aim to balance predictive accuracy against computational cost. Nonetheless, useful plasma models that can effectively contribute to the scientific and engineering advances in plasma physics need to not only be predictive and cost-efficient but also generalizable and interpretable, i.e., models that perform well across different scenarios and that their structure and predictions can be easily comprehensible. We will elaborate in the next section on how machine learning has the potential to aid achieving ROMs that meet all these criteria.

Before moving on, we would elaborate on some specific ways by which the ROMs can contribute to the progress in plasma science and engineering.

First, from a fundamental research perspective, predictive, generalizable, and interpretable ROMs will be the long-desired versatile analysis toolset that opens the door to novel approaches and frameworks to study plasmas, with the aim of answering the so-far-unresolved questions. One such approach, which is enabled by the plasma ROMs, consist of performing methodical parametric investigations of plasma systems. In fact, given the intricate nature of the plasmas, an effective pathway to fully comprehend the complex global picture of plasmas behavior can entail breaking down the overall complex picture in order to study and understand the contributions that each specific parameter influencing the plasma may have to the overall behavior and plasma response. Piecing these understandings together, we may solve the mysteries that still surround various plasma phenomena, some remained unanswered from the 1960s, such as the problem of “anomalous” electron transport [5].

The ROM-enabled parametric studies are also essential to complement the experimental research. The simulations provide a more controllable environment of study in which various effects that normally influence the plasma behavior simultaneously can be isolated and investigated more closely. In addition, the simulations enable access to information regarding the underlying plasma processes that might not be readily accessible in an experimental setting due to the spatiotemporal resolution limitations of the diagnostics and/or the physical accessibility constraints related to the device being tested. As a result, performing parametric sweeps of different plasma discharge parameters using predictive models can aid explaining the experimental observations and/or enriching the results of the test campaigns. Such parametric simulations also serve as a viable method to stress-test the available theoretical descriptions of the plasma to determine the extent to which the theoretically driven correlations between the characteristics of the plasma phenomena and the discharge properties are consistent with the numerical observations. This endeavor may lead to the development of better, more comprehensive theories for plasmas.

On the side of the applied (industrial) applications, ROMs hold immense potential to revamp the design and development processes, yielding groundbreaking advances. They are also cornerstones of catalyzing the transition of the plasma industries to more efficient digital engineering practices [9], particularly based on the digital twin technology. As is discussed at length in the accompanying paper to this work [10], there are several ways by which digital twins of plasma systems formed on the foundation of predictive ROMs can contribute to plasma engineering. These range from enabling computer-enabled optimization-based design approaches to informing optimal experimental design methodologies [11]. The latter application aids identifying the critical parameters and experimental conditions, leading in turn to more cost-efficient and informative test campaigns. Evidently, this can equally support the advancement within the fundamental scientific research.

Furthermore, digital twins are particularly effective for “what-if scenario” analyses, enabling the simulation of the plasma systems’ behavior over unobserved conditions or scenarios, hence, enhancing the comprehensiveness and reliability of the experimental data.

Digital twins provide pathways for continual, bidirectional exchange of information between the models and the physical device [10][12]. This transcends the predictive capabilities of the underpinning ROMs as well as their representativeness to the real world through dynamic adaptation. A specially relevant use case here involves controlling plasmas, which can for instance pave the path toward harnessing the fusion energy in addition to enabling fully autonomous operation of in-space plasma propulsion technologies.

In this paper, we delve into the reduced-order modeling framework, particularly as enabled by the machine-learning methodologies inspired and/or informed by physics, to ensure the accuracy, generalizability, and interpretability of the models. We discuss the progress achieved to date at Imperial Plasma Propulsion Laboratory (IPPL) on predictive, physics-inspired plasma ROMs and share some perspectives from the ongoing research directions.

II. Overview of the reduced-order modeling framework and the role of machine learning

Reduced-order modeling is a powerful technique, which is used to simplify the descriptions of complex systems by representing their essential dynamics in a lower-dimensional space, typically by focusing on dominant coherent structures underlying the system’s evolution.

To aid further elaborations on reduced-order modeling, we have presented in Figure 1 a graph of how various possible plasma physics models spread across the spectrum of computational cost vs accuracy and generalizability. As depicted, ROMs are defined in reference to a full-order model (FOM) against which the reduction is performed. In our context here, plasmas generally evolve within a six-dimensional phase space and over time. Consequently, as was discussed in the preceding section, the reference FOM in plasma physics can be considered to be kinetic simulations, which offer the highest fidelity level but also incur the highest computational cost.

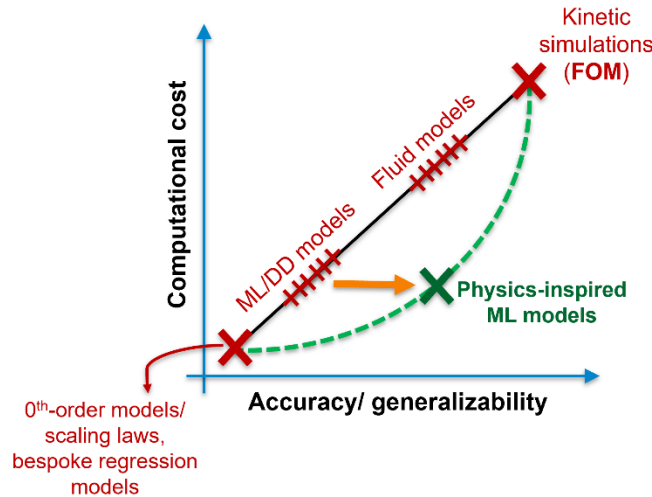


Figure 1. Graph of the computational cost vs accuracy and generalizability for various possible plasma physics models. The kinetic simulations represent full order models (FOMs), whereas fluid models and ML/DD models are reduced-order models (ROMs). Incorporation of physics-inspired constraints into learning promotes accuracy and generalizability while maintaining computational cost at the level of ML/DD models purely based on data.

Noting the high dimensionality of the plasmas' evolution, if one succeeds in establishing a closed 3D system of partial differential equations for a given plasma system, such system of PDEs can still be interpreted as a ROM. With this statement, plasma fluid models are a kind of reduced-order model because, to derive these models, we take velocity moments [13] of the plasma kinetic equation [14], which reduces the dimensionality of the problem in the most general case from 3D3V (three dimensional in configuration (physical) space and three dimensional in velocity (kinetic) space) to 3D. However, solving a 3D system of PDEs still entails significant computational cost, hence, placing fluid models at a rather intermediate position on the computational cost-accuracy plane.

It is noteworthy that zeroth-order models, scaling laws, and bespoke regression models are all some form of ROM as well. These models provide fast and computationally inexpensive solutions but lack the accuracy of more complex models and are hardly generalizable beyond the perimeters of their simplifying assumptions, problem descriptions, and the data for which they are developed.

Machine learning (ML) based and data-driven (DD) models fall somewhere between the extremes of accuracy/generalizability and computational cost. These models are faster alternatives to more expensive simulations and can, thus, enable rapid predictions for tasks such as design optimization and control. Traditionally, the high computational efficiency of purely data-based ML/DD ROMs comes at the expense of their generalizability. These models have been usually effective only within the specific geometries or parameter regimes on which they were trained.

We would point out that there are also classical non-ML-based approaches toward reduced-order modelling. One widely employed approach is the “POD-Galerkin” method. POD, or proper orthogonal decomposition, learns a low-dimensional coordinate system from data, enabling the representation of the dominant structures. A dynamical system is obtained within the POD-identified reduced subspace via performing a Galerkin projection of the governing equations (the Navier-Stokes equations for fluids e.g.), onto these modes. While the POD step is data-driven, Galerkin projection requires a functional numerical implementation of the full governing equations [15]. In general, classical projection-based reduced-order modeling methods necessitate custom modifications to the numerical solver, making the approaches rather intrusive [15].

To address the challenges, we need to explore alternative pathways for ROM development, particularly via ML. The primary objective would be to maintain the computational cost at the level of ML/DD models based purely on data while enhancing the accuracy and generalizability of the models. Toward this end, research has demonstrated that embedding prior or partial knowledge of physics can significantly enhance the performance and generalizability of the ML/DD ROMs while also reducing the amount of training data required [15]. This approach is often referred to as “physics-informed” [16] or, more broadly, “physics-consistent” ML modeling [17]. Examples of physics-based constraints include conservation laws, invariances, and symmetries, which can be incorporated as regularization terms in the loss function of the ML algorithm to penalize deviations from known physical behaviors.

The approach that we are envisioning to pursue is a nuanced version of the above. We refer to this as “physics-inspired” ML because, rather than strictly incorporating physical principles, we intend to focus on ensuring that the model respects two key constraints: (i) low dimensionality, and (ii) sparsity (parsimony).

With regard to low dimensionality, one of the primary ideas in ML/DD ROM is that even within very high-dimensional complex systems, there are underlying domain patterns that are crucial. For instance, in a megapixel snapshot of a simulation, not all individual pixels are important – only the dominant coherent structures matter. The number of these coherent modes is most often much lower than the total dimensionality of the system in the physical space (number of pixels, e.g.). In addition, the concept of sparsity or parsimony is inspired by the observation that, in reality, dynamical systems can be represented by a few key terms and variables. Indeed, all fundamental physical laws and equations currently known to humanity are inherently sparse.

The general philosophy behind the framework of “physics-inspired ML” is to develop models that do more than merely fitting the data. While numerous ML algorithms can accurately fit the data, these models are often regarded as black boxes, i.e., they lack interpretability and generalizability beyond their training envelope. We aim for dynamical systems that are sparse and possess as few degrees of freedom as necessary to describe the behavior of the system.

Following on the above discussions, we would now look into the two main steps that are, generally speaking, associated with ROM development: first, finding a suitable reduced coordinate system (set of dominant modes) and, second, discovering a dynamical system model (a set of ordinary differential equations (ODEs) for instance) in the identified reduced coordinates. For example, from high-fidelity, high-dimensional simulation data, a dimensionality-reduction technique, such as POD, can extract the dominant modes (structures) underlying the data. Then, a dynamics model can be constructed that describes how the amplitudes of these dominant modes evolve over time. Such a model is reduced order because instead of needing to resolve a large number of degrees of freedom, it now just resolves the time evolution of the amplitudes of a few spatial coherent structures.

We note that the above two steps for reduced-order modeling can be carried out consecutively or simultaneously. The simultaneous discovery of the reduced coordinates and the dynamics offers the advantage of constraining the learning algorithm to identify a coordinate system in which the dynamics can be described in a parsimonious and, hopefully, generalizable manner.

Coordinate transformation and, rather equivalently, dimensionality reduction are crucial elements of developing predictive ROMs that hold the highest potential for generalizability and interpretability. To illustrate the importance of these concepts, we consider two specific examples, visualized in Figure 2 and Figure 3.

First, referring to Figure 2, we consider the comparison between the geocentric and heliocentric views of the solar system. In the “naïve” geocentric coordinate system (Figure 2(a)), where Earth is placed at the center, the resulting dynamics appear complex, and no obvious laws describe the planetary motions in a straightforward way. However, when the coordinate system is transformed to place the Sun at the center (the heliocentric view, Figure 2(b)), the dynamical system becomes much simpler. Within this new framework, it is possible to derive simple, sparse models – Kepler’s laws – that accurately describe planetary motion. Interestingly, Kepler’s laws are, in fact, an early example of data-driven models derived from observational data.

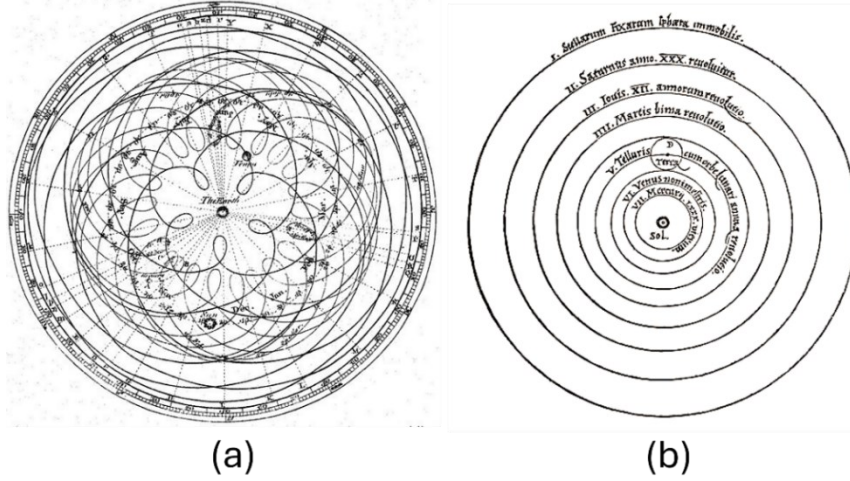


Figure 2. Visual demonstration of the concept of coordinate transformation for interpretable and generalizable physics modeling; models of the planetary motion: (a) the Geocentric Ptolemy’s model, (b) the Heliocentric Kepler’s model.

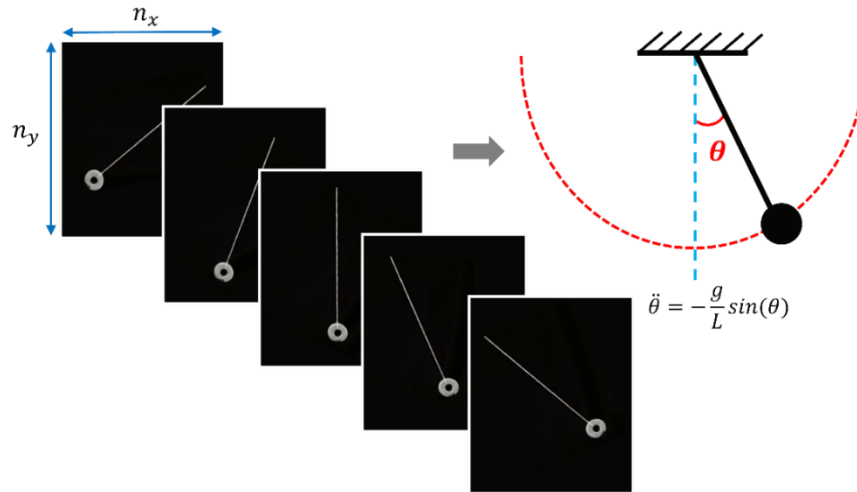


Figure 3. Visual demonstration of the role of dimensionality reduction to describe the underlying dynamics; example of a swinging pendulum as seen from high-dimensional video snapshots of the motion vs as described by one essential variable, the displacement angle (θ).

Second, looking at Figure 3, we imagine having a high-resolution video of a swinging pendulum. An ideal data-driven reduced-order model should be able to determine from the video that the minimum variable required to describe the motion of the pendulum is the angular displacement θ , and subsequently discover the governing equation for the

evolution of θ without any prior knowledge of the underlying dynamical system. However, learning both the coordinates and the dynamics from such high-dimensional data (i.e., the video) is not trivial. Nonetheless, by promoting sparsity and constraining the model to be as low-dimensional as possible, we can achieve this goal.

The above examples underscore the importance of coordinate transformation and dimensionality reduction on the path to developing effective and generalizable ROMs. By identifying the appropriate low-dimensional coordinate system, we can simplify the underlying dynamics, making it feasible to derive sparse, interpretable models that capture the essential behavior of complex systems.

A. Prior examples of ML ROM development for plasmas from literature

The interest in ML/DD applications for plasma modelling and simulation has surged over the past years. This rise in community attention has been accompanied by the emergence of a range of research efforts.

Despite this trend, it is important to acknowledge that ML applications in plasma physics are still nascent and have remained relatively limited in scope and extent, particularly compared to the adjacent fields to plasma physics such as fluid mechanics. High-temperature fusion plasma research has seen more prevalence in ML applications, whereas scientific machine learning has been more limited in low-temperature plasma physics and especially for plasmas relevant to EP technologies. One main reason for this status quo can be the scarcity of large enough volumes of data from plasma systems to support rigorous ML research for plasma physics. Computational cost issues of conventional modeling capabilities, as well as challenges associated with plasma experiments, are possible root causes for plasma physics still being distant from the realm of “big data” [18], where ML can be exploited to its fullest potential.

Nevertheless, community-wide efforts and coordination initiated within plasma physics communities in recent years to address the computational cost issue of high-fidelity plasma simulations, as well as to improve the employed diagnostics, have been changing the circumstances. This implies that ML research for plasma physics is today greatly timely and indeed essential for further advancements of the field.

Accordingly, in consideration of the context of this paper, a review of some prominent examples of ML-enabled ROM development efforts in the field is warranted before proceeding further.

Dynamic mode decomposition [19] has been used by several researchers toward reduced-order plasma modeling. Sasaki et al. [20] applied DMD to the simulation data of the Kelvin-Helmholtz instability [21] in order to extract the underlying dominant plasma structures and develop models for the time evolution of these structures (modes). Taylor et al. [22] and Kaptanoglu et al. [23] applied the DMD method to the experimental and simulation data from the HIT-SI fusion experiment. Faraji et al. [24][25] applied the optimized variant of the DMD method to data from different cross-field plasma configurations [1], demonstrating that this DMD variant enables reliable identification of dominant instability modes underlying plasma systems and the development of predictive plasma ROMs, even in the presence of high noise levels in the data.

Kaptanoglu et al. investigated the first-principles and data-driven approaches toward reduced-order modeling of a class of MHD plasmas, comparing the performance of the classical POD-Galerkin method against POD-SINDy (Sparse Identification of Nonlinear Dynamics [26]) and evaluating the effects of imposing physics constraints on the ROMs [27]. Recently, Kaptanoglu et al. have examined in Ref. [28] the applications of sparse regression methods such as SINDy [refer to section III.C for more details] to various tasks in plasma physics, including system identification, compressed sensing [29], and design optimizations. The SINDy algorithm has been used for plasma dynamics discovery by several other authors as well. The weak (integral) form of the SINDy algorithm was employed by Alvez and Fiuza in Ref. [30] to recover from PIC simulation data the known differential equations such as the Vlasov and the MHD equations. In Ref. [31], Abramovic et al. applied SINDy to data from the simulations of drift-wave turbulence based on the Hasegawa-Wakatani and modified Hasegawa-Wakatani models. Donaghy and Germaschewski [32] employed SINDy to derive from kinetic PIC simulation data of a magnetic reconnection configuration – the “collisionless Harris sheet reconnection problem” – to find a ten-moment fluid system of equations for the plasma with a closure model for the tenth moment term, aiming to improve the “Hammett-Perkins closure”.

Finally, in the context of Hall thruster plasmas, Jorns followed symbolic regression to find a functional relationship for the “anomalous” electron collision frequency in terms of the local plasma properties [33]. The derived data-driven symbolic function showed a predictive capability as well as a quantitative improvement over five first-principles models of the anomalous collision frequency [33]. The limitations of the approach, partly in terms of the extensibility of the derived functional forms to parameter spaces notably outside the training range, were examined and discussed [33].

According to the above brief and non-exhaustive overview, the integration of ML techniques into plasma modeling has demonstrated notable potential. The advances highlight the transformative impact of ML for developing predictive reduced-order plasma models, justifying further progress and consolidated efforts on this front.

III. Progress and perspectives on ML-enabled reduced-order plasma modelling

In this section, we discuss more in detail the two components of ROM development, namely, identifying a suitable reduced coordinate system, and discovering a dynamical system model within this reduced space. Alongside these discussions, we provide results from our progress toward ML-enabled predictive plasma ROMs and present some perspectives on promising directions for follow-on research.

A. Dimensionality reduction and identification of reduced coordinates

As was mentioned in section II, the concept of reduced-order modeling is fundamentally based on the premise that data generated by/for physical systems are typically low-rank. This means that within the often high-dimensional data, there exist a few dominant patterns or structures that are sufficient to explain the system's behavior. These dominant patterns capture the essential dynamics of the system, allowing for a significantly reduced representation without substantial loss of accuracy.

In the following, we will discuss two key approaches for dimensionality reduction: (i) singular value decomposition (SVD), and (ii) autoencoders.

1. Background

- I. **Singular Value Decomposition:** SVD is a powerful method in linear algebra for extracting low-rank features and patterns from data. In fluid dynamics, this method is often referred to as POD, while in statistics, it is known as PCA. These methods are fundamentally similar and provide a systematic way to decompose data into its most important components (underlying modes). SVD (POD) underpins many ROM techniques such as DMD and is also employed for data compression and denoising.

To apply the SVD method, we arrange our time series data collected from the system into a data matrix X . Each column of X contains the state of the system at a given time. This could be a snapshot of simulation data, reshaped or flattened into a high-dimensional column vector. Thus, $X \in \mathbb{C}^{n \times m}$, where n is the number of spatial points and m is the number of time steps (or snapshots).

Applying SVD to our data matrix X , we decompose it into three matrices as per Eq. 1.

$$X = U \Sigma V^*, \quad (\text{Eq. 1})$$

where the superscript $*$ denotes conjugate transpose. The three matrices in Eq. 1 are defined as follows: $U \in \mathbb{C}^{n \times m}$ is an orthogonal matrix whose columns (u_i) are the left singular vectors, representing the dominant spatial modes or patterns in the data, which form the basis of our coordinate system. $V^* \in \mathbb{C}^{m \times m}$ is an orthogonal matrix whose columns (v_i^*) are the right singular vectors, representing the time evolution or amplitudes of the corresponding modes. $\Sigma \in \mathbb{R}^{m \times m}$ is a diagonal matrix whose elements σ_i (the singular values) are ordered from largest to smallest and indicate the energy or variance captured by each SVD mode.

This decomposition provides a hierarchical coordinate system with bases defined by the columns of the matrix U . Depending on the distribution of singular values and the desired variance to retain within the data, we can truncate the SVD expansion at a certain rank r , yielding a low-rank approximation of our data ($\tilde{X}_r$) as per Eq. 2.

$$X \approx \tilde{X}_r = \tilde{U}_r \tilde{\Sigma}_r \tilde{V}_r^* = \sum_{i=1}^r \sigma_i u_i v_i^*. \quad (\text{Eq. 2})$$

The choice of r is typically guided by examining the distribution of the singular values and/or the normalized cumulative sum of the first r σ_i values [24][29]. A sharp drop in the distribution of singular values indicates that a (relatively) low-rank representation of the data exists that will capture most of the variance. This low-rank representation significantly reduces the dimensionality of the data while preserving its essential features.

- II. **Autoencoders:** These are neural network (NN) architectures designed to learn efficient, compressed representations of the input data by first encoding it into a lower-dimensional space and then decoding it back to the original input. Autoencoders are used for tasks such as feature extraction and denoising. Figure 4 shows a schematic representation of an autoencoder with multiple hidden layer (7 in this case). The hidden layers of an autoencoder can include fully connected layers with nonlinear activation functions, convolutional layers, etc.

As can be seen from Figure 4, autoencoders feature an input and output layer that matches in size the high-dimensional system state and a bottleneck hidden layer that reduces the state to a low-dimensional space, consisting of one or more latent variables (ζ). The encoder part (ϕ) of the network maps the high-dimensional state (x_k , k being a time instant for a time-series data, e.g.) to the latent space. Autoencoders belong to the

unsupervised learning category. As a result, even though it is the latent space representation at the bottleneck of the autoencoder that we are interested in, the decoder part that reconstructs an estimate of the high-dimensional state ($\hat{x}_k$) from the latent space is crucial for the unsupervised training of the network. A common loss function for the autoencoder architectures is the L2-norm error between the decoder's state reconstruction and the original input state.

The power of autoencoders to learn nonlinear manifold coordinates to represent the data can enable achieving much improved compressions in the latent space [15] compared to linear decomposition methods like SVD/POD. As such, autoencoders play a crucial role in ROM development.

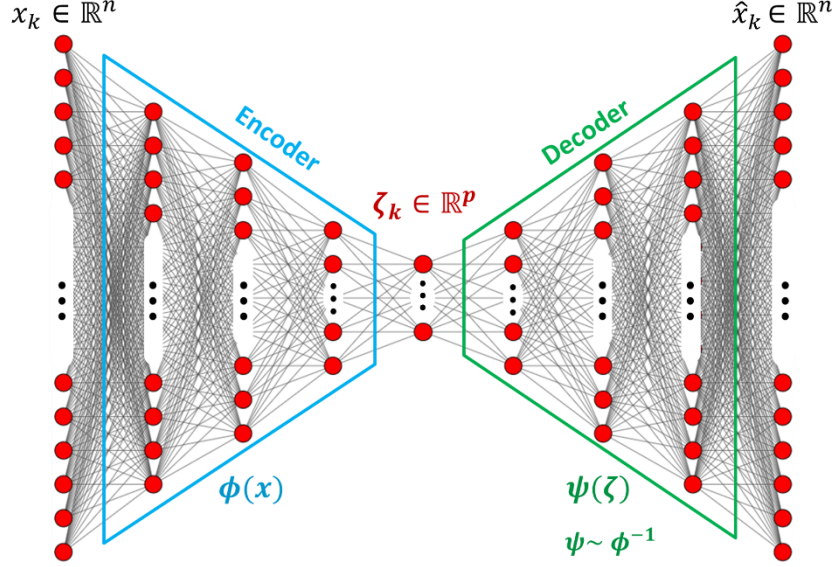


Figure 4. Schematic representation of an autoencoder network. The encoder ($\phi(x)$) reduces the dimensionality of the data ($x_k \in \mathbb{R}^n$), finding the latent space variables (ζ_k) belonging to a reduced subspace $\mathbb{R}^p$, with typically $p \ll n$. The decoder ($\psi(\zeta)$) maps the latent variables of the autoencoder back to the high-dimensional $\mathbb{R}^n$ space. The network is trained by minimizing the loss between the input data to the encoder (x_k) and the output data of the decoder ($\hat{x}_k$).

We would underline here that there is a connection between the SVD and the autoencoder architecture, nonetheless. In fact, the autoencoder is a nonlinear generalization of the SVD method, achieved by extending the number of layers and incorporating nonlinear activation functions. It can be shown that a shallow autoencoder with single-layer encoder and decoder, as well as linear identity activation functions, is equivalent to SVD (POD). Within the following results subsection, we will demonstrate this point using data from a plasma test case.

2. Results

Now that we have introduced two key dimensionality-reduction techniques, we intend to demonstrate their capabilities in providing a low-rank (low-dimensional) representation of data from a complex plasma system.

We compare the performance of SVD against three autoencoder architectures with an increasing number of fully connected hidden layers and a progressively reduced number of latent variables in the autoencoder's bottleneck layer. The three autoencoder architectures are schematically illustrated in Figure 5. Through the comparison of the performance, we aim to evaluate the extent to which we can lower the dimensionality of our data using each of the methods and architectures without losing the ability to have an accurate reconstruction of the ground-truth.

Before presenting the results, we briefly describe the test case and the characteristics of the data. The test case is adopted from Ref. [34], where a 2D benchmarking problem has been presented and discussed. This benchmark is a well-established and well-studied case [35]-[38] and corresponds to a plasma configuration representative of a radial-azimuthal ($x - z$) plane of a typical Hall thruster geometry. The test case setup features externally imposed spatiotemporally constant fields – an axial electric field (E_y) and a radial magnetic field (B_x). The plasma in this test case undergoes complex spatiotemporal evolution dynamics under the influence of the electromagnetic fields and interactions with the radial walls.

The data used for all the results and analyses in this paper are obtained from a high-fidelity reduced-order quasi-2D (Q2D) PIC simulation of the radial-azimuthal benchmark test problem using the IPPL-Q2D code [39]-[42]. The details of the simulation setup, as well as the associated numerical and physical parameters, can be found in Refs.

[34][41]. These details of how the PIC simulation was performed are not relevant to the discussions that will follow. What is important are the characteristics of the data, which we focus on next.

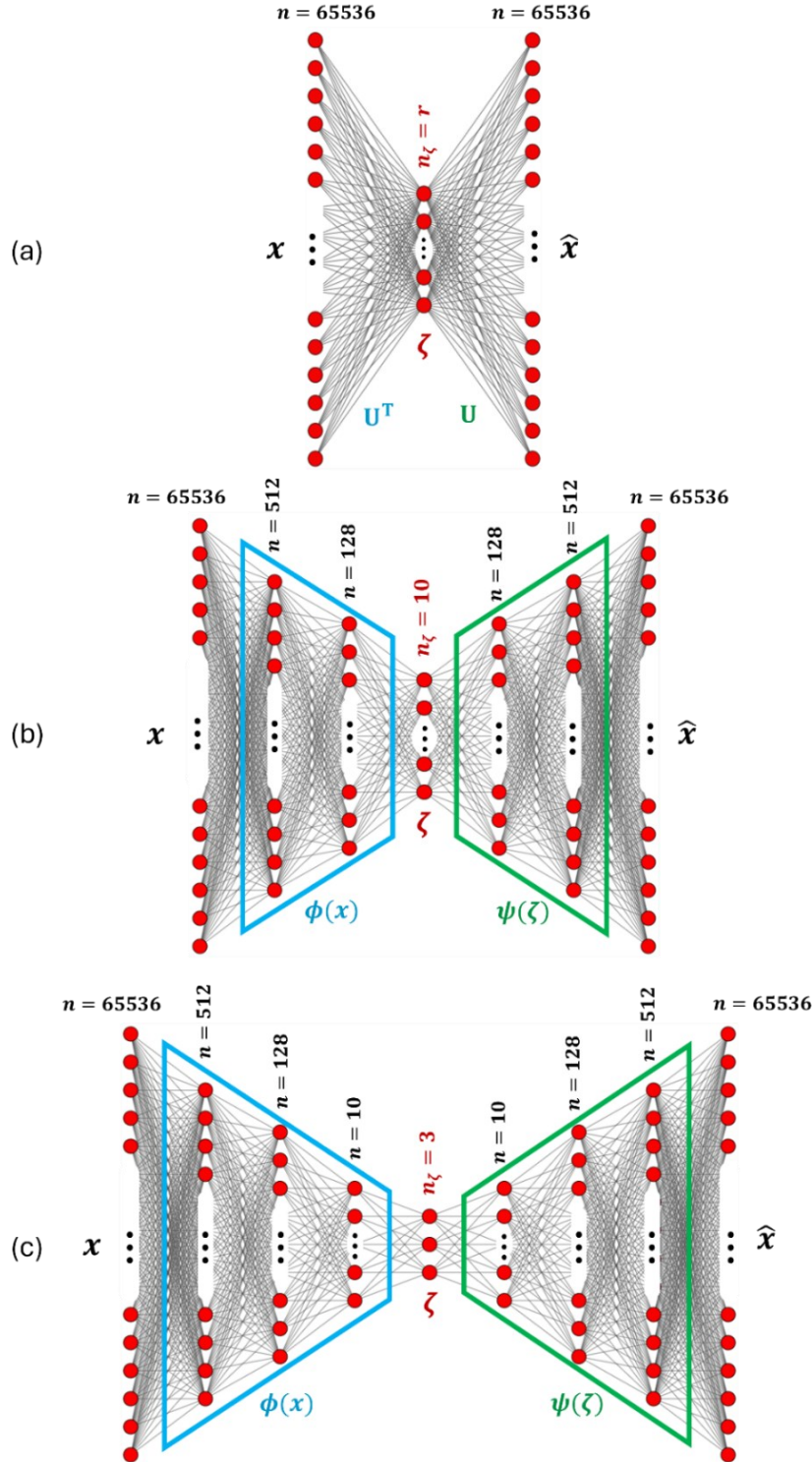


Figure 5. Schematic representation of the autoencoder networks applied to the data from the radial-azimuthal plasma test case; (a) shallow linear autoencoder, which is equivalent to SVD, (b) autoencoder with 5 hidden layers and 10 latent variables, (c) autoencoder with 7 hidden layers and 3 latent variables.

The dataset corresponds to a $135 \mu\text{s}$ -long simulation of the radial-azimuthal test case. This is equivalent to about 9000 output snapshots of the plasma property of interest in this paper, which is the electrons' axial current density

(J_{ey} , representing electrons' cross-magnetic-field transport). Each J_{ey} snapshot represents an average over 1000 simulation timesteps. The Q2D output data of the J_{ey} were reconstructed on a 2D grid with the size of 256×256 along each direction of the domain (radial and azimuthal). Hence, the total dimensionality of the “ground-truth” snapshots is $n = 256 \times 256 = 65536$.

The first 400 snapshots, which correspond to the initial transient of the system before reaching a quasi-steady state, were excluded from the dataset. Out of the remaining 8500 snapshots, about 3500 were used for training and about 5000 for testing.

In the context of dimensionality reduction, testing implies the assessment of how well unseen data can be represented by the identified reduced bases from the dimensionality-reduction algorithm (here, SVD or autoencoder). Accordingly, after determining the reduced coordinates using the training data, the test data are projected onto the reduced coordinates, and then the accuracy of the low-rank reconstructions of the unseen data is evaluated. For SVD, this can be mathematically presented as in Eq. 3, in which subscript k indicates the timestep. For autoencoders, the described projection and reconstruction amount to feed-forward passes of the test data through the trained network.

$$X_{k,r} = U_r U_r^T X_k. \quad (\text{Eq. 3})$$

Figure 6 shows the distribution of the singular values (σ) vs the number of ranks from applying SVD to the radial-azimuthal test case data. The rank number is equivalent to the number of latent variables (n_ζ) in the bottleneck layer of the shallow linear autoencoder in Figure 5(a). From the distribution, a truncation rank (r) of 100 seems reasonable for a low-dimensional representation of the data. This is confirmed from Figure 7, which shows that the reconstructed J_{ey} snapshots using 100 SVD rank are almost exactly reproducing the ground-truth PIC-simulation snapshots at several random time instants within the test window. A rank-100 representation implies that the dimensionality of the data is reduced by a factor of about 600.

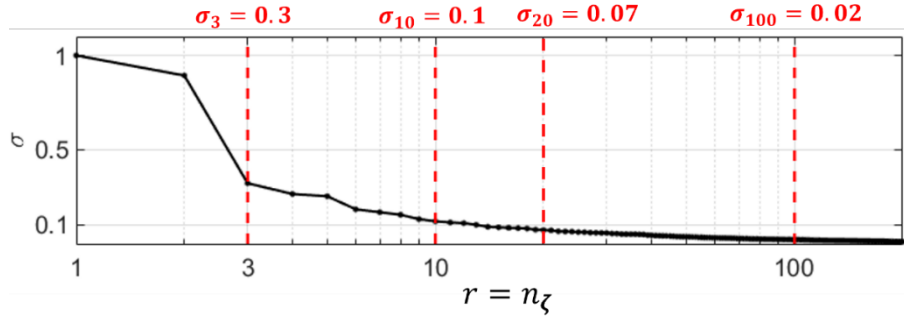


Figure 6. Distribution of the singular values (σ) from the SVD of the data. r indicates the number of SVD ranks, which can be equivalent to the number of variables (n_ζ) in the latent space (bottleneck) of an autoencoder.

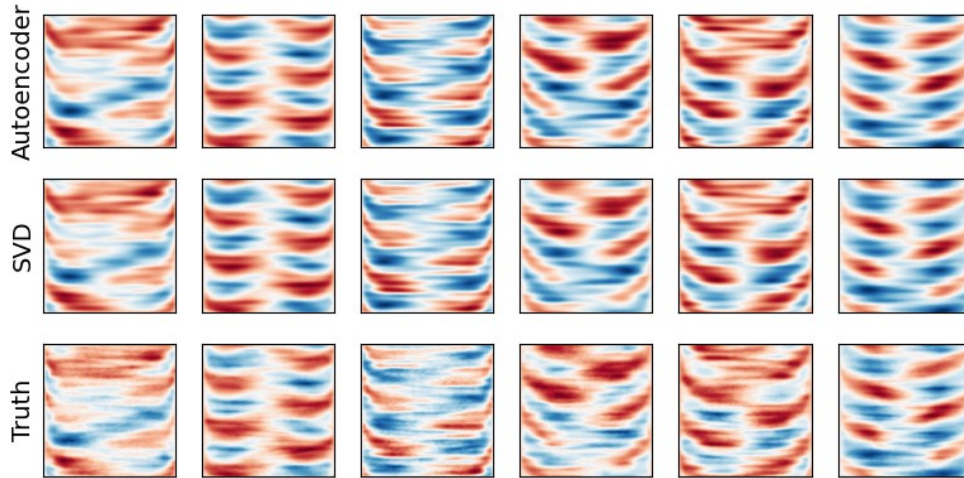


Figure 7. Comparison of the reconstructed normalized 2D snapshots of the electrons' axial current density (J_{ey}) from the shallow linear autoencoder (**top row**) and the SVD (**middle row**) with $r = n_\zeta = 100$ against the ground-truth snapshots from the PIC simulation (**bottom row**). Each column represents a random sample time instant within the test window. The horizontal and vertical axes represent the radial and the azimuthal directions, respectively.

Interestingly, we can see from Figure 7 that the reconstructed J_{ey} snapshots from the shallow linear autoencoder with $n_z = 100$ also provide an excellent reconstruction of the ground-truth and, moreover, represent an identical resemblance to the reconstructed snapshots from SVD. This observation attests the argument mentioned earlier that a shallow linear autoencoder is equivalent to SVD. Further supporting evidence is provided in Figure 8, where we have compared several dominant modes underlying the data as identified by the shallow autoencoder and SVD. It is also evident from these comparisons that the two approaches have equivalent performance.

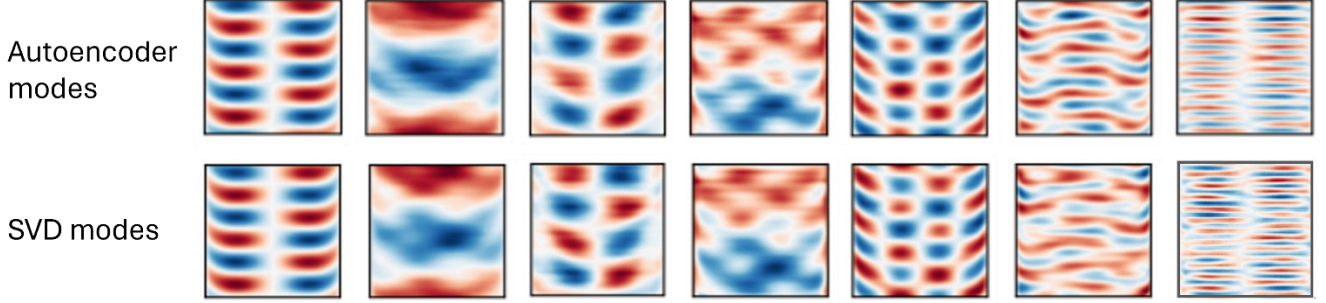


Figure 8. Comparison of the first several dominant modes from the decomposition of the data using the shallow linear autoencoder (**top row**) and the SVD (**bottom row**).

In any case, an important aspect that distinguishes the autoencoders from linear decomposition techniques such as SVD is their modularity, which in turn enables the networks to find nonlinear embeddings of the data. Therefore, it is expected that autoencoders with increased hidden-layer numbers and nonlinear activation functions would be able to represent data using a much lower number of latent space variables.

To demonstrate this, we start by assessing the performance of a shallow linear autoencoder with various number of latent variables at the bottleneck and then examine how the reconstruction quality at the extreme of highly low-dimensional latent space improves when we add nonlinearity to the architecture. The goal is to find out how much further we can reduce the dimensionality of our complex plasma data beyond the factor of 600 obtained using rank-100 SVD representation.

The results of this investigation are illustrated in Figure 9, where we show snapshots of the J_{ey} data from the PIC simulation and from the autoencoder architectures of Figure 5 at several randomly selected time instants within the test window. Comparing the reconstructed snapshots from the shallow linear autoencoder with various numbers of latent space variables against the ground-truth ones, we observe that, with $n_z = 10$, substantial loss of detailed information in the data occurs, and with merely 3 modes, the reconstruction is inadequate and not at all similar to the ground-truth. These comparisons clearly show that the reconstruction quality of the shallow liner autoencoder (essentially SVD) is poor with a low number of latent variables, failing to capture the essential details of the plasma data.

To address this, we utilize deeper nonlinear autoencoders. We see from Figure 9 that a 5-layer nonlinear autoencoder can reduce the dimensionality to 10 latent variables while maintaining an accurate representation of the data. The reconstruction quality has significantly improved with the 5-layer autoencoder with $n_z = 10$ compared to what was observed with 10 modes from the single-layer architecture. The 5-layer autoencoder with $n_z = 10$ has reduced the data dimensionality by more than 6000 times.

By employing a 7-layer nonlinear autoencoder, we can further reduce the dimensionality to 3 while still achieving a remarkably good representation of the data. This degree of compression corresponds to more than 20,000 times reduction in data dimensionality. Indeed, comparing against the ground-truth, the reconstructed snapshots of J_{ey} from the 7-layer autoencoder are seen to have retained a high level of detail. This demonstrates the capability of sufficiently deep networks to effectively represent complex plasma states in highly compressed latent spaces.

The ability to represent a complex high-dimensional plasma state using only 3 latent variables is thought provoking. It, in fact, poses the question that whether plasmas are truly high-dimensional or that they only appear as such because the physical space (which is intuitive for human's understanding) is not the intrinsic/natural space for plasmas.

The great efficacy of deep autoencoders in providing a highly reduced nonlinear embedding of the data, as demonstrated here for a sample case of seemingly high-dimensional plasma data, means that these architectures serve as an essential component for developing physics-inspired ML-based ROMs that can be interpretable and readily generalizable. In fact, in Figure 10, we have presented the time evolutions of the three latent space variables in the bottleneck of the 7-layer autoencoder. Being in a sufficiently reduced coordinate comprised by these three latent

variables, the next step toward building a ROM for the plasma evolution in our test case amounts to employing an ML/DD approach that learns the dynamics (time evolutions) of these three variables (essential dominant modes) as illustrated in Figure 10. We will discuss this further in the next subsection.

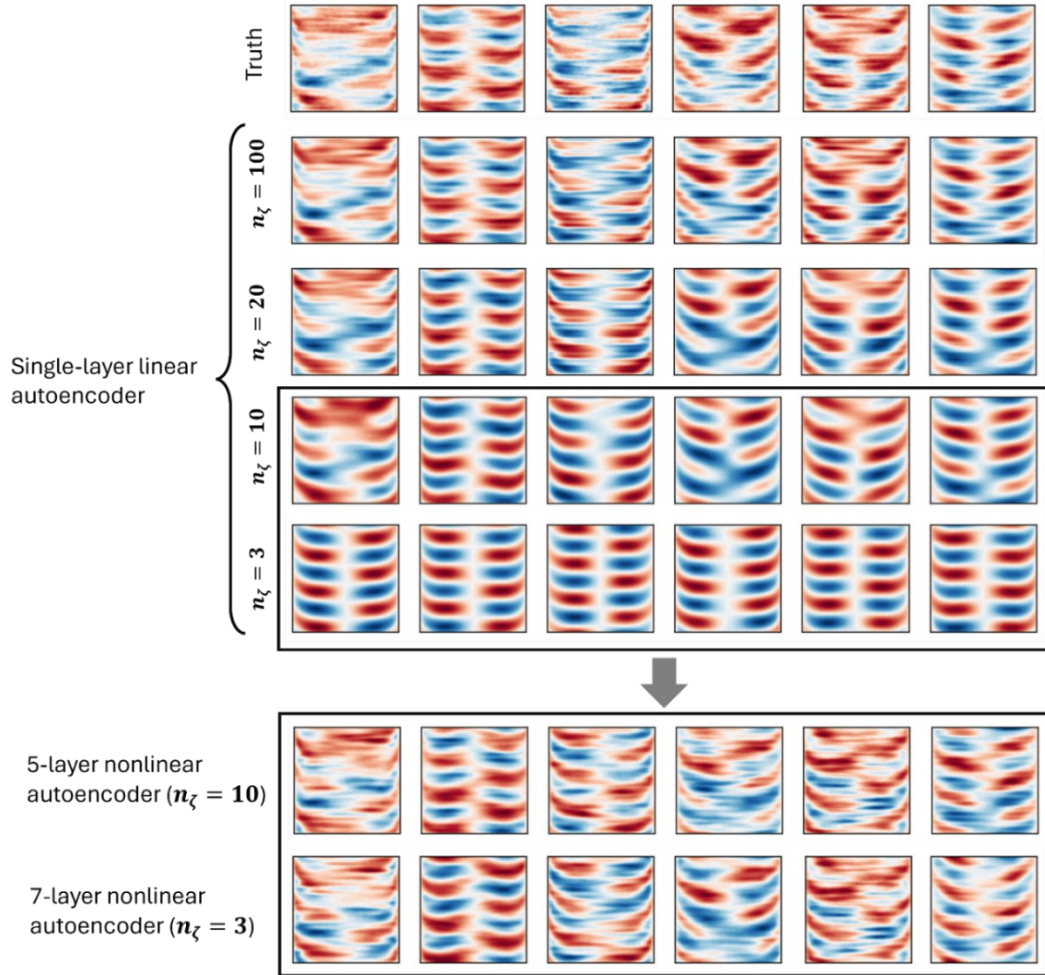


Figure 9. Demonstration of the data reconstruction from compressive representations obtained using the different autoencoder architectures shown in Figure 5 at several random time instants within the test window; (**top subplot**) reconstructions of the J_{ey} data using a single-layer linear autoencoder with varying number of latent variables compared against the ground-truth data snapshots, (**bottom subplot**) J_{ey} reconstructions using nonlinear autoencoders with different hidden layers and latent variables.

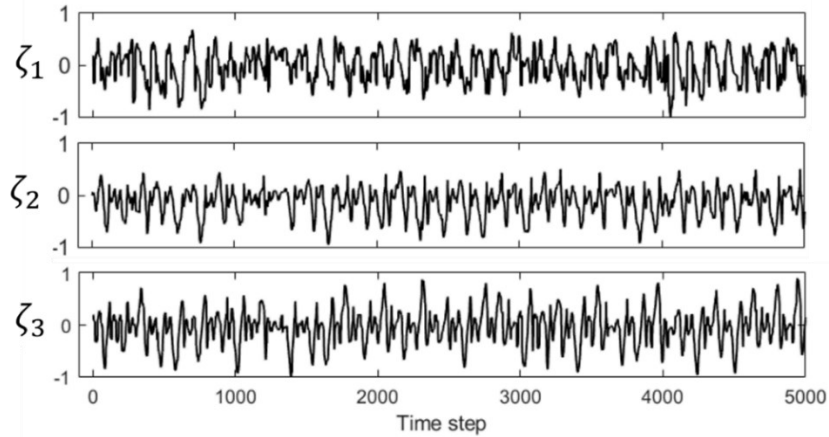


Figure 10. Time evolutions over the test window of the values of the three latent space variables from the nonlinear 7-layer autoencoder (Figure 5(c)).

B. Data-driven discovery of the dynamics

We observed from the above results that ML and, in particular, the autoencoders can derive a much more compressed (low-dimensional) representation of the system beyond what is feasible with linear methods like SVD/POD. The discovery of the dynamics, i.e., finding relations and/or governing equations for how the state of a system evolves over time, can also majorly benefit from ML, especially for complex physical systems such as plasma technologies for which the complete and closed first-principles system of governing equations may be at least partially unknown. The concepts of data-driven and ML-enabled dynamics discovery were described at length in our accompanying paper to this contribution [43].

Here, in the context of reduced-order modeling, what we would emphasize is that, once we have found a suitable reduced coordinate system, there are several ML/DD dynamics discovery algorithms that can be applied to represent the dynamics of the system in that reduced coordinate. This, in fact, corresponds to the second step of ROM development as was discussed in section II. We recall from those discussions that this dynamics discovery in the latent space can be performed consecutively to the determination of the system's low-dimensional representation. We can, for instance, apply data-driven regression techniques such as DMD [19][44], SINDy [26], and Phi Method [45][46] or neural-network-based time-series learning architectures like recurrent neural networks (RNN), long short-term memory (LSTM) [47], and echo-state networks [48] to the time evolutions of the latent variables ζ_1 to ζ_3 in Figure 10 to derive reduced-order models describing the plasma dynamics in the radial-azimuthal test case. We look further at some results and/or perspectives regarding the application of the named data-driven regression techniques in the remainder of this paper.

In any case, before proceeding further, discussing one point is important. We pointed out in section II that our overarching objective toward ROM development for plasmas is to obtain models that are low-dimensional but as well parsimonious (sparse) in the dynamical space, i.e., featuring as few number of terms as possible. In this sense, sparsity serves as another constraint that we would like to impose on our ROMs. However, if we aim for learning the dynamics asynchronously to finding a low-dimensional representation of our data, it may happen that the system would not have a sparse dynamical description in the identified reduced coordinate of the autoencoder for example. This is where the framework of simultaneous discovery of the reduced coordinates and the dynamics becomes relevant and essential. A schematic representation of an architecture suitable for this simultaneous learning is illustrated in Figure 11. As can be seen from this schematic, the coordinate discovery (encoder) part of the architecture reduces the dimensionality of the high-dimensional data at a generic time step k (x^k) to a low-dimensional latent space parameterized by the variables ζ^k . The dynamics discovery algorithm now learns in the bottleneck of the autoencoder the time evolution of the latent variables. In a discrete sense, the dynamic discovery algorithm finds the transformation or mapping that advances the variables ζ one time step to the future (ζ^{k+1}). The decoder part now maps the low-dimensional representation of the predicted future state to the high-dimensional space ($\hat{x}^{k+1}$).

By coupling/pairing the two optimizations corresponding to dimensionality reduction and sparse dynamics identification, we can enforce the network to find a low-dimensional coordinate for the system where the dynamics can be described parsimoniously, i.e., “as simple as possible but no simpler” – a compressed version of a quote accredited to Einstein [49]).

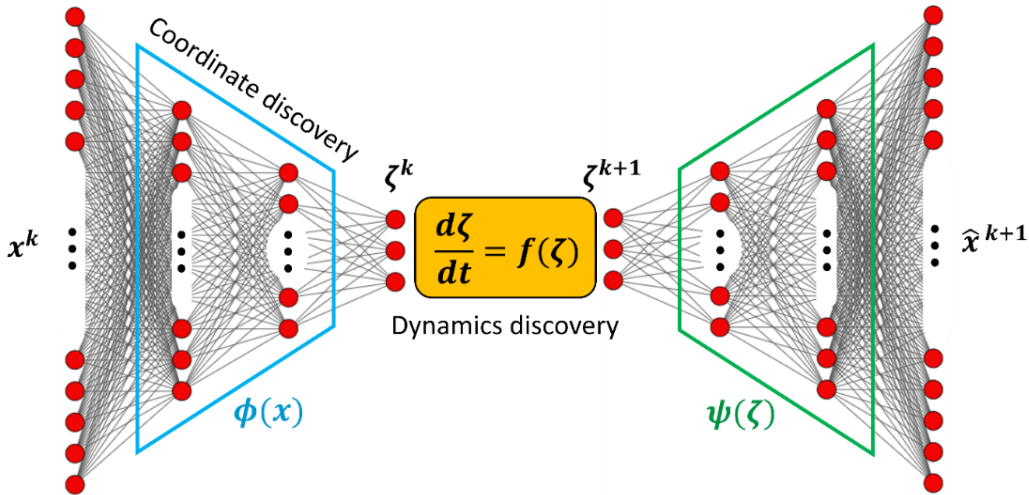


Figure 11. Illustration of a joint dimensionality reduction-dynamics discovery ML architecture for simultaneous coordinate transformation and learning of the dynamics in the reduced coordinates.

1. Example of ROMs based on the optimized dynamic mode decomposition

We can imagine that a rather simple, yet effective and powerful example of a coupled dimensionality reduction-dynamics discovery algorithm is the DMD method. DMD is a technique introduced by Schmid [44] in fluid dynamics to decompose a flow field into series of coherent spatiotemporal modes with exponential time dynamics. In a more general sense, the principal idea behind DMD is to ideally combine a spatial dimensionality-reduction technique, such as POD, with that of the Fourier transform in time. The method provides a best-fit linear-time-dynamics model describing the spatiotemporal data from a system.

Noting these aspects of DMD, we can represent the method as the combination of a single-layer linear autoencoder with a linear operator for dynamics at the bottleneck (Figure 12). In other words, we can think of DMD as an approach that finds a linear regression model for the data in the reduced SVD/POD coordinates.

Of course, to compute a DMD model for the data, we do not use a shallow autoencoder. However, considering the equivalence between DMD and the architecture shown in Figure 12 provides a tangible example for the discussions so far in this section III.B.

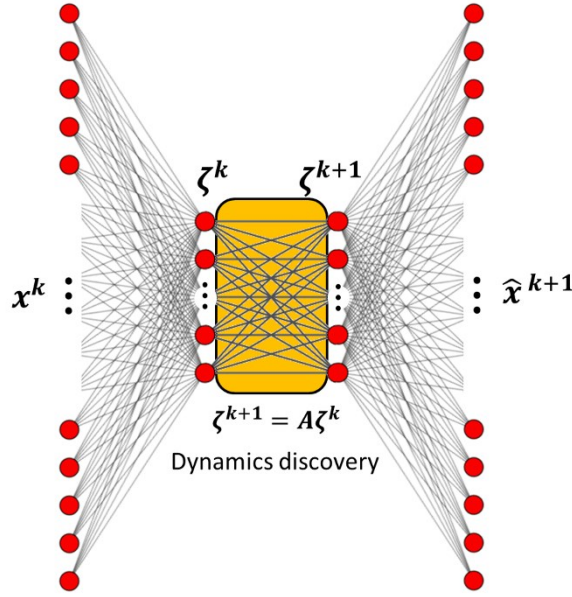


Figure 12. Representation of the dynamic mode decomposition method as the combination of a shallow linear autoencoder network and a linear-in-time transformation in the network's latent space.

The results we would present in the following are from applying a powerful variant of DMD, namely optimized DMD (OPT-DMD), to the data from the radial-azimuthal plasma test case. The dataset here includes an expanded set of outputs from the PIC simulation, comprising several plasma properties – electrons' drift velocity components along the radial, azimuthal, and axial directions ($V_{de,x}, V_{de,z}, V_{de,y}$), radial and azimuthal components of the electron temperature (T_{ex}, T_{ez}), and the azimuthal component of the electric field (E_z).

The OPT-DMD method, originally devised by Askham and Kutz [50], is a major advance over the basic DMD method, particularly with regard to noise robustness. The method has significantly changed the potential of DMD to perform critical modeling tasks. Importantly, OPT-DMD can be used to constrain best-fit linear models to be stable by construction.

Accordingly, the demonstration that we present here involves using OPT-DMD to find a linear ROM that best describes the time evolution of an expanded multi-property state vector of the plasma system simultaneously. This is different from the results we reported in Ref. [25], where we developed and tested linear ROMs using OPT-DMD by applying the method separately to the spatiotemporal signals of individual plasma properties. In that work [25], we observed a promising predictive potential for the OPT-DMD-derived ROMs. Nonetheless, the enhancement of the DMD's state vector by including the data of several plasma properties allows capturing the interdependency between the properties. This is, hence, expected to further improve the predictive performance of the OPT-DMD ROMs [51]. Another significance of the demonstrations here is that they present the performance of OPT-DMD ROMs for long-term forecasting of the plasma dynamics.

The results are presented in Figure 13 and Figure 14. The DMD (SVD) truncation rank of the model for all the following results is 150.

Figure 13 compares the time evolutions of the spatially averaged values of the abovementioned plasma properties from the ground-truth PIC simulation against the signals obtained from the OPT-DMD ROM. The OPT-DMD model simultaneously learned the time dynamics of all the properties and, in turn, provided a simultaneous prediction of their time variations over a rather extended time window of $45 \mu\text{s}$ ($90\text{-}135 \mu\text{s}$). Here, for a close-up assessment, the zoomed-in views over a time interval around the end of the training window, i.e., $t = 90 \mu\text{s}$, are shown. Moreover, an additional comparison is shown in Figure 13 between the ground-truth data and the OPT-DMD ROM predictions in terms of the time variations of the local values of the plasma properties at the mid-point within the radial-azimuthal domain.

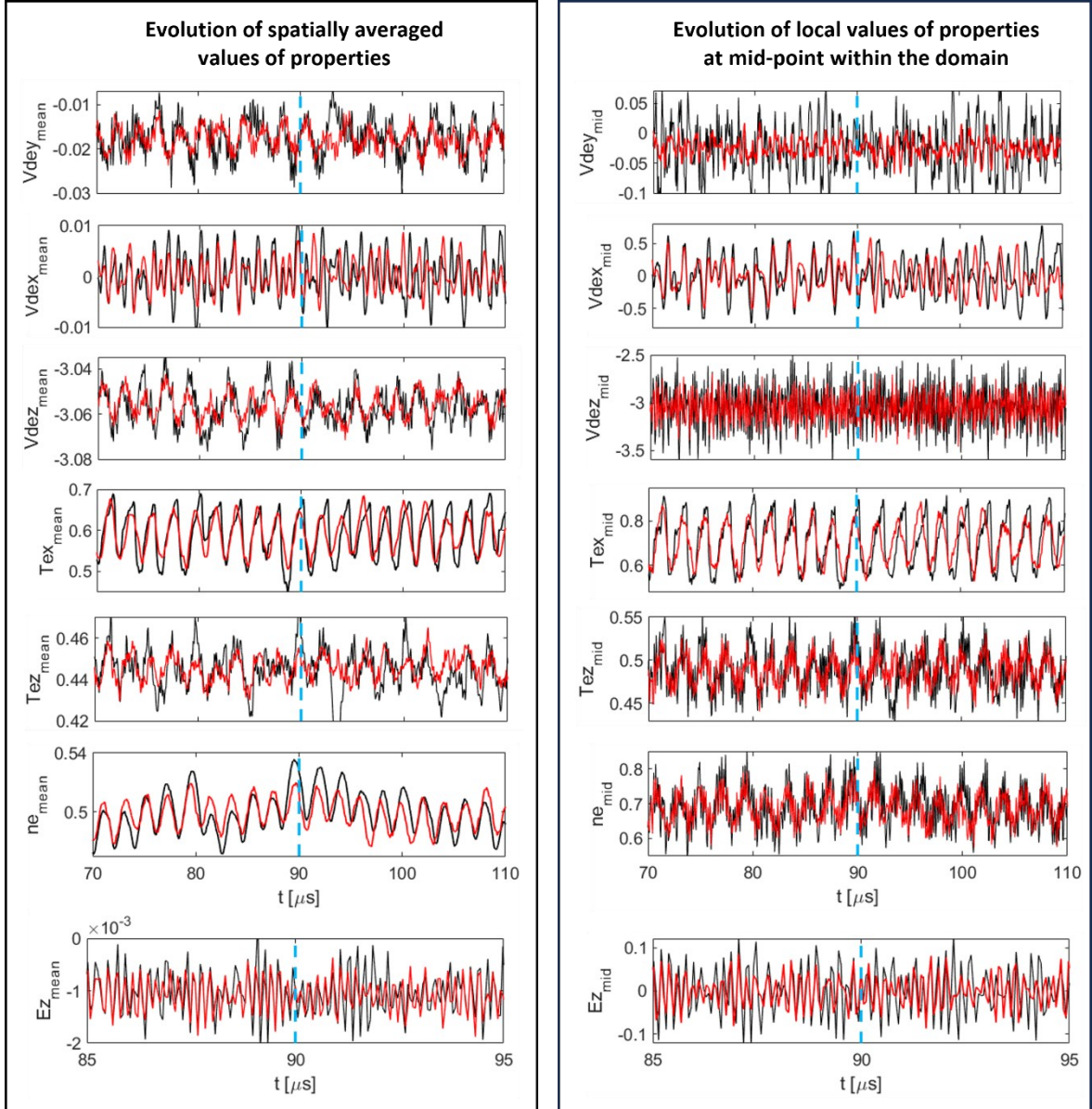


Figure 13. Time evolutions of the plasma properties that comprised the constructed multi-property plasma state vector; **(left column)** temporal variation plots of the spatially averaged plasma properties over the time interval of $70\text{-}110 \mu\text{s}$ ($20 \mu\text{s}$ before and after the training end time) for all the plasma properties, except for the E_z . For this property, the timeframe is $85\text{-}95 \mu\text{s}$. This is done to aid better distinguishing the ground-truth and the predicted evolutions. **(right column)** time variations of the local values of the plasma properties. The signals' duration for all the properties are consistent with those of the plots on the left column.

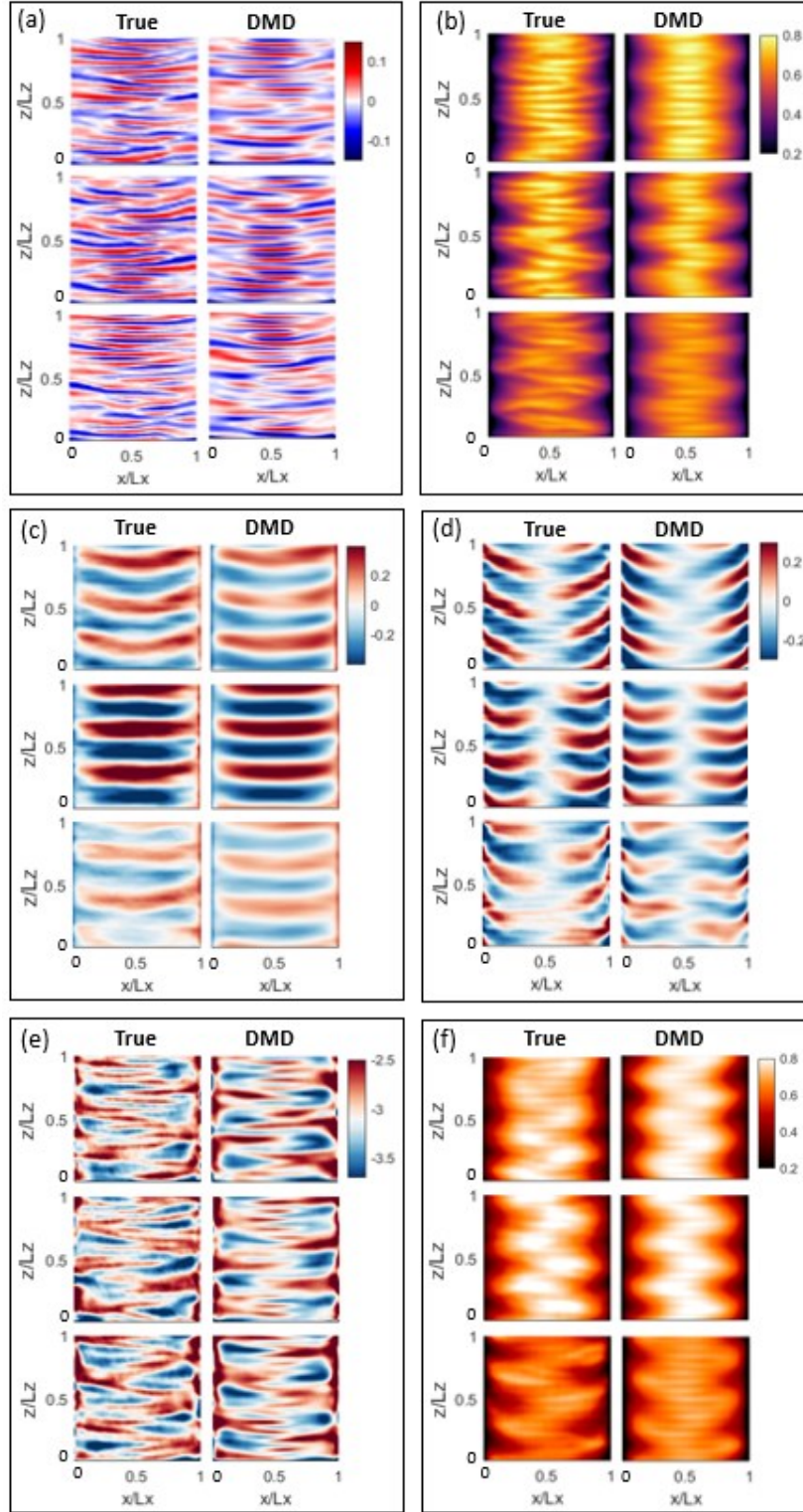


Figure 14. Comparison between the 2D snapshots of several plasma properties from the ground-truth radial-azimuthal Q2D simulation and the OPT-DMD model at three randomly selected time instants within the test window (90-130 μ s); (a) azimuthal electric field (E_z), (b) electron number density (n_e), (c) electrons' radial drift velocity ($V_{de,x}$), (d) electrons' axial drift velocity ($V_{de,y}$), (e) electrons' azimuthal drift velocity ($V_{de,z}$), (f) radial electron temperature (T_{ex}). Figure adopted from Ref. [51].

From Figure 13, it is remarkable to observe that the predicted signals from the OPT-DMD model remain fully stable over time and are closely following the ground-truth signals across all the properties. In fact, the plots on the left column of Figure 13 show that the OPT-DMD-predicted signals faithfully follow the ground-truth traces for all the plasma properties. In particular, the predicted time dynamics (oscillations' frequency) remains consistent with the ground-truth from the training into the testing interval.

The time evolution plots of the local values of the plasma properties (Figure 13(right-column)) also show an overall good agreement between the predictions and the ground-truth, with the main oscillation patterns of the ground-truth signals being captured by the OPT-DMD-predicted signals. It is emphasized that the ability of a data-driven model to provide rather accurate predictions of the local behaviors of complex systems is most often a major challenge. However, the OPT-DMD model trained on a multi-property state vector is seen here to overcome this challenge in an effective manner.

Now, for a visual qualitative assessment of the performance of the OPT-DMD ROM for simultaneous forecasting of a multi-property plasma state vector, we refer to Figure 14. This figure compares the predicted 2D snapshots of several plasma properties from the OPT-DMD ROM against the ground-truth snapshots at three randomly selected time instants within the test window (beyond the 90- μ s mark in Figure 13). Across all the time instants and plasma properties, we observe a remarkable degree of similarity between the OPT-DMD-predicted and the ground-truth snapshots.

To conclude the discussions, we underline that the performance of any ML/DD ROM, including the one from OPT-DMD here, shall be always assessed against the reduction in model complexity and the computational cost advantage. In this regard, the accuracy of the predictions from OPT-DMD ROM can be deemed notably good considering that it is a linear-time-dynamics model for a nonlinear plasma system and has a negligible computational cost compared to the reference full-order PIC model.

C. Future directions for ROM development for plasmas

The DMD method provides linear ROMs to describe the time dynamics of systems. Nevertheless, many physical systems of applied interest, including plasma systems, generally exhibit nonlinear and/or transient dynamics. In this regard, there are several ways by which we can incorporate nonlinearity in a data-driven dynamics discovery model.

One possible way is to use library-based data-driven regression methods [26][45][46][52][53]. In such a ROM development framework, we augment the potential model terms to include nonlinear observables or functions of the latent space variables in addition to the variables themselves. Essentially, we build a library of candidate terms that represent various nonlinear interactions and dynamics within the system. The objective is to identify the fewest terms from this library that best describe the system's dynamics and to determine their corresponding coefficients.

This process involves sparse regression and sparsity-promoting optimization techniques, such as sequential thresholding [54], and L1 regularization or LASSO regression. These techniques help in isolating the most relevant terms by penalizing the inclusion of unnecessary terms, thereby promoting a sparse solution. It is important to note that, despite the nonlinearity within the library terms themselves on which regression is performed, the regression framework remains linear.

By employing the above approach, we can capture the underlying nonlinear dynamics of the system in a computationally efficient manner while maintaining the interpretability of the model. This methodology allows for a robust and flexible way to model complex physical phenomena, making it highly suitable for applications in plasma dynamics.

Two prominent examples of the library-based data-driven regression methods that can be used within the described ROM framework are SINDy and Phi Method. Since the dynamics discovery here happens in the latent space of an autoencoder, these two approaches become rather equivalent. As such, because we have discussed Phi Method at length in the accompanying paper on dynamics discovery [43], we elaborate a bit further in the following on SINDy.

Figure 15 shows a schematic of an autoencoder architecture with SINDy at its bottleneck. This autoencoder-SINDy architecture has been already applied and demonstrated by Champion et al. in several applications [54]. Further to this, Mars Gao and Kutz [55] have combined SINDy with Bayesian autoencoder networks in order to facilitate uncertainty quantification through a probabilistic joint framework for simultaneous dimensionality reduction and dynamics discovery.

In the architecture of Figure 15, SINDy uses time series data of the latent variables (ζ) to extract a dynamical system model best describing the time evolutions (e.g., the signals in Figure 10). To do this, SINDy performs a sparse regression on a library (Θ) of candidate linear and nonlinear terms such that the resulting model best approximates the time derivatives ($\dot{\zeta}$) of the latent variables (ζ). The sparse regression finds the coefficient values (a_i s) of the minimum number of "active" terms in matrix Θ .

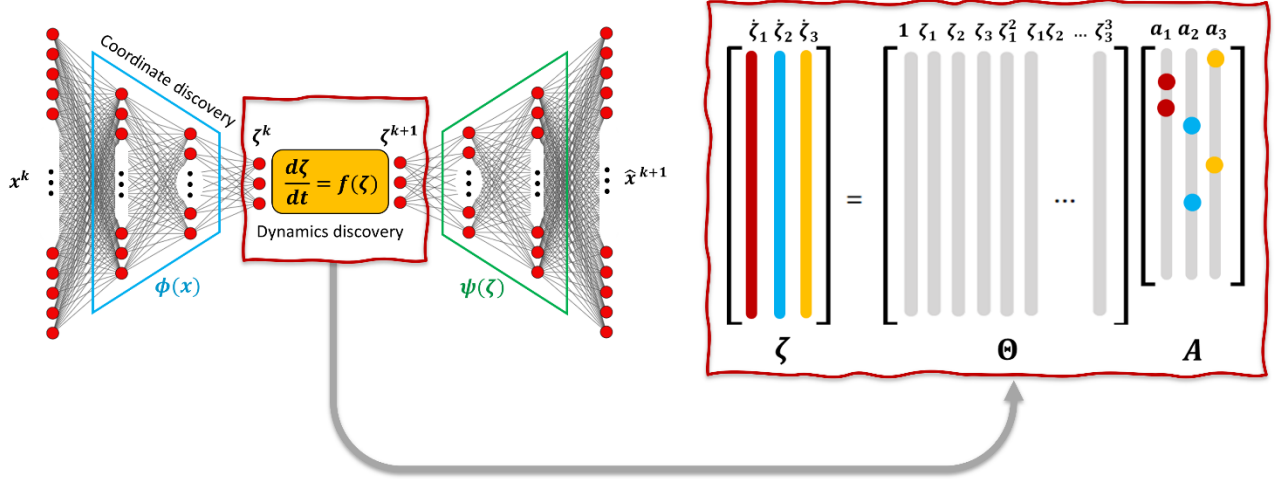


Figure 15. Schematic visualization of the SINDy algorithm as the dynamics discovery method in the latent space of the autoencoder.

For the library-based regression methods like SINDy, some expert knowledge or prior understanding of the physics of the system can be greatly helpful in creating the library of candidate terms. This helps to facilitate the training of the network and the involved regression by avoiding a very large number of terms in the library. However, as we are applying the regression in the reduced latent space of an autoencoder, it is highly likely that suitable observables of the latent variables are not known a-priori to establish the candidate library. As a result, it would be advantageous if the network can itself learn the required dynamical terms from the data.

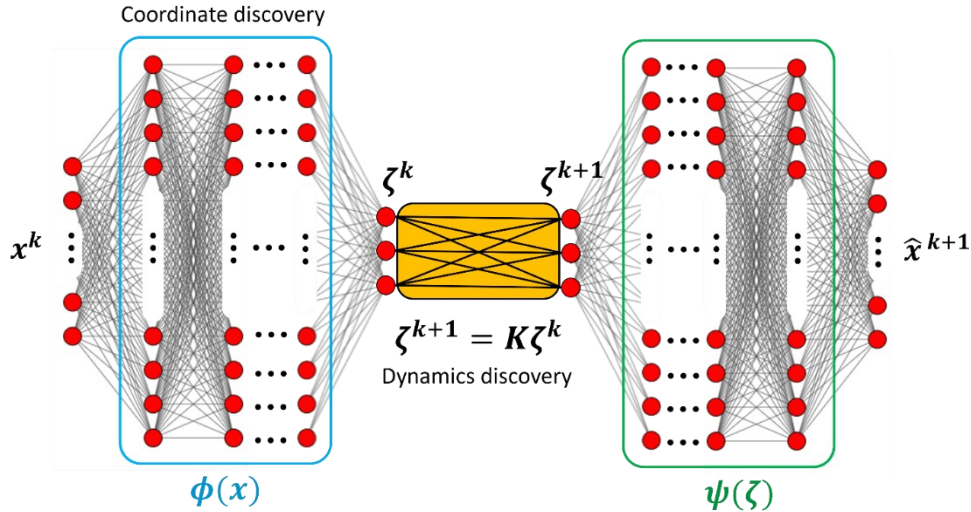


Figure 16. Schematic representation of deep Koopman networks for the simultaneous discovery of the linear embedding of nonlinear dynamics (Koopman transformation) and the associated time-stepping operator K .

A powerful approach to do this is by leveraging the Koopman analysis framework, based on the Koopman theory [56], which aims to find a coordinate transformation (embedding) in which the nonlinear dynamics can be linearly represented.

The Koopman operator, a central concept in this framework, involves defining an infinite-dimensional operator that acts on observables (functions) of the state vector from a finite-dimensional nonlinear dynamical system. By applying this operator to suitable observables, the dynamics can be described linearly in the observables space. The Koopman operator lifts the dynamics from a finite-dimensional state space to an infinite-dimensional function space, where the evolution of the original system becomes linear.

Despite being an abstraction, the Koopman theory suggests that such a mapping exists, although identifying the appropriate observables to achieve this linear representation is nontrivial and a subject of ongoing research. In this regard, several researchers, e.g. in Refs. [57]-[59], have attempted using deep-learning architectures and autoencoders,

like the one illustrated schematically in Figure 16, to identify nonlinear observables and/or coordinate embeddings that may realize Koopman linear representation of nonlinear dynamics.

As a concluding remark here, we would emphasize that the DMD method is closely related to the Koopman analysis framework, and efforts have been made to extend DMD with nonlinear observables [60][61] in order to find an approximation for the Koopman operator. Accordingly, the Koopman analysis and the related DMD implementations serve as powerful tools for ROM development of nonlinear systems. This means that future research efforts on this front, particularly in our context of reduced-order plasma modeling, are indeed warranted.

References

- [1] Kaganovich ID, Smolyakov A, Raitses Y, Ahedo E, Mikkellides IG, Jorns B, Taccogna F, et al., “Physics of $E \times B$ discharges relevant to plasma propulsion and similar technologies,” *Phys. Plasmas* **27**, 120601 (2020)
- [2] Boeuf JP, Smolyakov A, “Physics and instabilities of low-temperature $E \times B$ plasmas for spacecraft propulsion and other applications,” *Physics of Plasmas* **30**, 050901 (2023)
- [3] Goldstein ML, “Major Unsolved Problems in Space Plasma Physics,” In: Meyer-Vernet N, Moncuquet M, Pantellini F (eds) “Physics of Space: Growth Points and Problems,” *Springer*, Dordrecht (2001)
- [4] Borovsky JE, Delzanno GL, Valdivia JA, Moya PS, Stepanova M, Birn J, Blum LW, Lotko W, Hesse M, “Outstanding questions in magnetospheric plasma physics: The pollenzo view,” *Journal of Atmospheric and Solar-Terrestrial Physics*, **208**:105377 (2020)
- [5] Balescu R, “Aspects of Anomalous Transport in Plasmas,” 1st Edition, CRC Press, 2019
- [6] Villafana W, Cuenot B, Vermorel O, “3D particle-in-cell study of the electron drift instability in a Hall Thruster using unstructured grids,” *Phys. Plasmas* **30**, 033503 (2023)
- [7] Birdsall CK, Langdon AB, “Plasma Physics via Computer Simulation,” CRC Press, 1991
- [8] Taccogna F, Cichocki F, Eremin D, Fubiani G, Garrigues L, “Plasma propulsion modeling with particle-based algorithms,” *J. Appl. Phys.* **134**, 150901 (2023)
- [9] URL to a description of digital engineering and digital tools: <https://insights.sei.cmu.edu/blog/what-digital-engineering-and-how-it-related-devsecops/>, last accessed in May 2024.
- [10] Faraji F, Reza M, Knoll A, “Digital twins for Electric Propulsion: concept, game-changing potentials, and building blocks,” IEPC-2024-256, to be presented at the 38th International Electric Propulsion Conference, Toulouse, France (2024)
- [11] Smucker B, Krzywinski M, Altman N, “Optimal experimental design,” *Nat Methods* **15**, 559–560 (2018)
- [12] Faraji F, Reza M, Knoll A, “Digital twins for 21st-century transformation of the Electric Propulsion sector: Vision 2030,” SP2024_046, In *Proc. of Space Propulsion Conference* (2024)
- [13] Kuldinov DA, Yamashita Y, Mansour AR, Hara K, “Ten-moment fluid model with heat flux closure for gasdynamic flows,” *Journal of Computational Physics*, Vol. **508**, 113030 (2024)
- [14] Bittencourt JA, “Fundamentals of Plasma Physics,” 3rd ed., Springer New York, NY, 2004
- [15] Vinuesa R, Brunton SL, “Enhancing computational fluid dynamics with machine learning,” *Nat Comput Sci* **2**, 358–366 (2022)
- [16] Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P et al., “Physics-informed machine learning,” *Nat Rev Phys* **3**, 422–440 (2021)
- [17] Brunton SL, Noack BR, Koumoutsakos P, “Machine Learning for Fluid Mechanics,” *Annu. Rev. Fluid Mech.* **52**:477–508 (2020)
- [18] Pollard A, Castillo L, Danaila L, Glauser M (eds.), “Whither Turbulence and Big Data in the 21st Century?” Springer Cham, 2017.
- [19] Kutz JN, Brunton SL, Brunton BW, Proctor JL, “Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems,” Philadelphia, PA: Society for Industrial and Applied Mathematics; Series: Other titles in applied mathematics; 149 (2016)
- [20] Sasaki M, Kawachi Y, Dendy RO, Arakawa H, Kasuya N et al., “Using dynamical mode decomposition to extract the limit cycle dynamics of modulated turbulence in a plasma simulation,” *Plasma Physics and Controlled Fusion*, **61** 112001 (2019)
- [21] Shivamoggi BK, “Kelvin-Helmholtz instability of a plasma in a magnetic field: finite-resistivity effects,” *Physica Scripta*. **24** 49 (1981)
- [22] Taylor R, Kutz JN, Morgan K, Nelson BA, “Dynamic mode decomposition for plasma diagnostics and validation,” *Rev Sci Instrum* **89**, 053501 (2018)
- [23] Kaptanoglu AA, Morgan KD, Hansen CJ, Brunton SL, “Characterizing Magnetized Plasmas with Dynamic Mode Decomposition,” *Physics of Plasmas* **27**, 032108 (2020)
- [24] Faraji F, Reza M, Knoll A, Kutz JN, “Dynamic Mode Decomposition for data-driven analysis and reduced-order modelling of $E \times B$ plasmas: I. Extraction of spatiotemporally coherent patterns,” *J. Phys. D: Appl. Phys.* **57** 065201 (2024)
- [25] Faraji F, Reza M, Knoll A, Kutz JN, “Dynamic Mode Decomposition for data-driven analysis and reduced-order modelling of $E \times B$ plasmas: II. Dynamics forecasting,” *J. Phys. D: Appl. Phys.* **57** 065202 (2024)
- [26] Brunton SL, Proctor JL, Kutz JN, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *PNAS* **113** (15) 3932–3937 (2016)

- [27] Kaptanoglu AA, Morgan KD, Hansen CJ, Brunton SL, “Physics-constrained, low-dimensional models for magnetohydrodynamics: First-principles and data-driven approaches,” *Physical Review E*, **104** 015206 (2021)
- [28] Kaptanoglu AA, Hansen C, Lore JD, Landreman M, Brunton SL, “Sparse regression for plasma physics,” *Physics of Plasmas*, **30** 033906 (2023)
- [29] Reza M, Faraji F, Knoll A, Kutz JN, “Shallow Recurrent Decoder architecture with application to super-resolution enhancement of the reduced-order particle-in-cell scheme,” *IEPC-2024-503, to be presented at the 38th International Electric Propulsion Conference, Toulouse, France* (2024)
- [30] Alves EP, Fiuza F, “Data-driven discovery of reduced plasma physics models from fully kinetic simulations,” *Physical Review Research*, **4**, 033192 (2022)
- [31] Abramovic I, Alves EP, Greenwald M, “Data-driven model discovery for plasma turbulence modelling,” *Journal of Plasma Physics*, **88**(6):895880604 (2022)
- [32] Donaghy J, Germaschewski K, “In search of a data-driven symbolic multi-fluid ten-moment model closure,” *Journal of Plasma Physics*, **89**(1):895890105 (2023)
- [33] Jorns B, “Predictive, data-driven model for the anomalous electron collision frequency in a Hall effect thruster,” *Plasma Sources Sci. Technol.* **27** 104007 (2018)
- [34] Villafana W, Petronio F, Denig AC, Jimenez MJ, Eremin D et al., “2D radial-azimuthal particle-in-cell benchmark for E×B discharges,” *Plasma Sources Sci. Technol.* **30** 075002 (2021)
- [35] Reza M, Faraji F, Knoll A, “Parametric investigation of azimuthal instabilities and electron transport in a radial-azimuthal E×B plasma configuration,” *Journal of Applied Physics* **133**, 123301 (2023)
- [36] Reza M, Faraji F, Knoll A, “Influence of the magnetic field curvature on the radial-azimuthal dynamics of a Hall thruster plasma discharge with different propellants,” *J. Appl. Phys.* **134**, 233303 (2023)
- [37] Reza M, Faraji F, Knoll A, “Effects of the applied fields’ strength on the plasma behavior and processes in E×B plasma discharges of various propellants: I. Electric field”, *Phys. Plasmas*. **31**, 032120 (2024)
- [38] Reza M, Faraji F, Knoll A, “Effects of the applied fields’ strength on the plasma behavior and processes in E×B plasma discharges of various propellants: II. Magnetic field”, *Phys. Plasmas*. **31**, 032121 (2024)
- [39] Reza M, Faraji F, Knoll A, “Concept of the generalized reduced-order particle-in-cell scheme and verification in an axial-azimuthal Hall thruster configuration,” *J. Phys. D: Appl. Phys.* **56** 175201 (2023)
- [40] Reza M, Faraji F, Knoll A, “Generalized reduced-order particle-in-cell scheme for Hall thruster modeling: concept and in-depth verification in the axial-azimuthal configuration,” *arXiv preprint*, arXiv:2208.13106 (2022)
- [41] Faraji F, Reza M, Knoll A, “Verification of the generalized reduced-order particle-in-cell scheme in a radial-azimuthal E×B plasma configuration,” *AIP Advances* **13**, 025315 (2023)
- [42] Reza M, Faraji F, Knoll A, “Latest verifications of the reduced-order particle-in-cell scheme: Penning discharge and axial-radial Hall thruster case,” AIAA 2024-2712, *In Proc. of SciTech Forum conference*, Orlando, Florida (2024)
- [43] Faraji F, Reza M, Knoll A, “Data-driven dynamics discovery and discrepancy modeling for advancing plasma science and technology,” *IEPC-2024-502, to be presented at the 38th International Electric Propulsion Conference, Toulouse, France* (2024)
- [44] Schmid PJ, “Dynamic mode decomposition of numerical and experimental data”, *J. Fluid Mech.* **656**.1 (2010)
- [45] Faraji F, Reza M, Knoll A, Kutz JN, “Data-driven local operator finding for reduced-order modelling of plasma systems: I. Concept and verifications,” *ArXiv preprint*, arXiv:2403.01523 (2024)
- [46] Faraji F, Reza M, Knoll A, Kutz JN, “Data-driven local operator finding for reduced-order modelling of plasma systems: II. Application to parametric dynamics,” *ArXiv preprint*, arXiv: 2403.01532 (2024)
- [47] Hochreiter S, Schmidhuber J, “Long Short-Term Memory,” *Neural Comput* **9** (8): 1735–1780 (1997)
- [48] Pathak J, Hunt B, Girvan M, Lu Z, Ott E, “Model-Free Prediction of Large Spatiotemporally Chaotic Systems from Data: A Reservoir Computing Approach,” *Phys. Rev. Lett.* **120**, 024102 (2018)
- [49] Robinson A, “Einstein said that — didn’t he?” *Nature* **557**, 30 (2018)
- [50] Askham T, Kutz JN, “Variable projection methods for an optimized dynamic mode decomposition”, *SIAM Journal on Applied Dynamical Systems* **17**:1, 380-416 (2018)
- [51] Faraji F, Reza M, Knoll A, “Machine-learning-enabled plasma modelling and prediction,” AIAA-2024-2708, *In Proceedings of 2024 SciTech Forum conference*, Orlando, Florida (2024)
- [52] Small M, Judd K, Mees A, “Modeling continuous processes from data”, *Physical Review E*, **65**:046704 (2002)
- [53] Yao C, Bollt EM, “Modeling and nonlinear parameter estimation with Kronecker product representation for coupled oscillators and spatiotemporal systems”, *Physica D: Nonlinear Phenomena*, **227**(1):78–99 (2007)
- [54] Champion K, Zheng P, Aravkin AY, Brunton SL, Kutz JN, “A Unified Sparse Optimization Framework to Learn Parsimonious Physics-Informed Models From Data”, in *IEEE Access*, vol. **8**, pp. 169259-169271 (2020)
- [55] Mars Gao L, Kutz JN, “Bayesian autoencoders for data-driven discovery of coordinates, governing equations and fundamental constants,” *Proceedings of the Royal Society A*, **480**: 20230506 (2024)
- [56] Koopman BO, “Hamiltonian systems and transformation in Hilbert space”, *Proceedings of the National Academy of Sciences*, **17**(5):315–318 (1931)
- [57] Yeung E, Kundu S, Hodas N, “Learning deep neural network representations for Koopman operators of nonlinear dynamical systems,” In 2019 American Control Conference (ACC), *IEEE*, pp. 4832-4839 (2019)
- [58] Takeishi N, Kawahara Y, Yairi T, “Learning Koopman invariant subspaces for dynamic mode decomposition,” *Advances in neural information processing systems* **30** (2017)

- [59] Lusch B, Kutz JN, Brunton SL, “Deep learning for universal linear embeddings of nonlinear dynamics,” *Nature Communications*, vol. **9**, no. 1, p. 4950 (2018)
- [60] Williams MO, Kevrekidis IG, Rowley CW, “A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition,” *Journal of Nonlinear Science* **25** 1307-1346 (2015)
- [61] Li Q, Dietrich F, Boltt EM, Kevrekidis IG, “Extended dynamic mode decomposition with dictionary learning: A data-driven adaptive spectral decomposition of the Koopman operator,” *Journal of Nonlinear Science*, **27**(10):103111 (2017)