Physics-constrained, low-dimensional models for magnetohydrodynamics: First-principles and data-driven approaches

Alan A. Kaptanoglu, Kyle D. Morgan, Chris J. Hansen, and Steven L. Brunton

Physics-constrained, low-dimensional models for MHD: First-principles and data-driven approaches

Alan A. Kaptanoglu
Department of Physics, University of Washington, Seattle, WA, 98195, USA

Kyle D. Morgan
Department of Aeronautics and Astronautics, University of Washington, Seattle, WA, 98195, USA

Chris J. Hansen
Department of Aeronautics and Astronautics, University of Washington, Seattle, WA, 98195, USA and
Department of Applied Physics & Applied Mathematics, Columbia University, New York, NY, 10027, USA

Steven L. Brunton
Department of Mechanical Engineering, University of Washington, Seattle, WA, 98195, USA

Plasmas are highly nonlinear and multi-scale, motivating a hierarchy of models to understand and describe their behavior. However, there is a scarcity of plasma models of lower fidelity than magnetohydrodynamics (MHD), although these reduced models hold promise for understanding key physical mechanisms, efficient computation, and real-time optimization and control. Galerkin models, obtained by projection of the MHD equations onto a truncated modal basis, and data-driven models, obtained by modern machine learning and system identification, can furnish this gap in the lower levels of the model hierarchy. This work develops a reduced-order modeling framework for compressible plasmas, leveraging decades of progress in projection-based and data-driven modeling of fluids. We begin by formalizing projection-based model reduction for nonlinear MHD systems. To avoid separate modal decompositions for the magnetic, velocity, and pressure fields, we introduce an energy inner product to synthesize all of the fields into a dimensionally-consistent, reduced-order basis. Next, we obtain an analytic model by Galerkin projection of the Hall-MHD equations onto these modes. We illustrate how global conservation laws constrain the model parameters, revealing symmetries that can be enforced in data-driven models, directly connecting these models to the underlying physics. We demonstrate the effectiveness of this approach on data from high-fidelity numerical simulations of a 3D spheromak experiment. This manuscript builds a bridge to the extensive Galerkin literature in fluid mechanics, and facilitates future principled development of projection-based and data-driven models for plasmas.

I. INTRODUCTION

Plasmas and plasma-enabled technologies are pervasive in everyday life [1], but their nonlinear, multi-scale behavior poses severe challenges for understanding, modeling, and controlling these systems. There are a tremendous number of known plasma models of varying model complexity, from magnetohydrodynamics (MHD) to the Klimontovich equations, but a large gap exists in the lower levels of this hierarchy between simple circuit models and the many MHD variants. These low-level models are motivated because higher fidelity models typically require computationally intensive and high-dimensional simulations [2–4], obfuscating the dynamics and precluding model-based real-time control. Moreover, many high-dimensional nonlinear systems tend to evolve on low-dimensional attractors [5]; plasmas across a large range of parameter regimes, geometry, and degree of nonlinearity exhibit this feature [6–13]. In these cases, the evolution of only a few coherent structures, obtained from model-reduction techniques [14, 15], can closely approximate the evolution of the high-dimensional physical system. Fortunately, recent progress in theoretical, data-

driven, and machine learning methods are revolutionizing the analysis, modeling, and control of high-dimensional, nonlinear systems, especially in the field of fluid mechanics [16]. Reduced-order modeling is advancing particularly rapidly, enabling the modeling of increasingly complex fluid flows [5, 16–21], but many of these advances have not yet been adopted in the plasma physics community. In this manuscript, we provide a framework for physics-constrained, low-dimensional plasma models which address this important gap in the model hierarchy.

The applications of reduced-order models include understanding reduced physical mechanisms [22, 23], computationally efficient simulations [24], digital twins (virtual time-dependent models for a dynamic system, constantly updating with sensor measurements) [25], and real-time control [26–28]. For example, acceleration of inertial confinement fusion simulations and digital twins can facilitate an exploration of the implosion parameter space [24], surrogate closure models can lead to more accurate and efficient fluid simulations [29], surrogate gyrokinetic transport models can speed up tokamak simulations by orders of magnitude [30, 31], and steady-state tokamak operation will require the active avoidance or

mitigation of disruptions, which can seriously damage components of the device [32]. For these real-time control challenges, there are a wealth of model-based control techniques such as model predictive control [33] that can be leveraged for plasma systems. However, existing models can be either too high-dimensional and computationally expensive to operate in real-time, or too low-fidelity to be useful for control.

In addition to being computational efficient, reduced-order models can help uncover key mechanisms that govern the evolution of the dominant coherent structures. This aspect of reduced-order modeling has a rich history, from the famous Lorenz model in 1963 [22], through the present era, including the low-order mechanistic model of the cylinder wake in 2003 by Noack et al. [17]. Recently, data-driven algorithms, like the algorithms used in this work, have shown potential to uncover similarly interpretable and useful models. Examples include related fluid systems [34, 35] as well as recent work that uncovers a Lorenz-like model of electroconvective chaos by Guan et al. [23]. Moreover, increasingly reduced order models are used to describe key mechanisms in plasma physics, including "predator-prey" dynamics in gyrokinetic simulations [36], direct data-driven discovery of reduced MHD or kinetic equations from a plasma dataset [37], and data-driven fluid models for the L-H mode transition in tokamaks [38]. These models are often critical for providing insight into the physical system, including energy transfers and other nonlinear interactions.

Reduced-order models traditionally fall into two categories: projection-based model reduction and data-driven system identification. Projection-based model reduction is achieved by first computing the evolution of a governing partial differential equation (PDE) model, often by spatially discretizing the domain, resulting in a high-dimensional system of ordinary differential equations (ODEs). In our case, we consider a 3D MHD simulation. Then a low-dimensional orthogonal basis is computed, often via the proper orthogonal decomposition (POD) [5, 39]. Finally, the high-dimensional model is "Galerkin-projected" onto this basis [40, 41], resulting in an efficient reduced system that describes how the amplitudes of the POD modes evolve in time. However, this projection is intrusive since it requires knowledge of the governing physics and a working high-fidelity solver.

In contrast, system identification techniques attempt to identify data-driven models directly from measurement data, often without knowledge of the governing equations. Increasingly, data-driven methods are producing effective bases beyond POD for different experimental or computational tasks [5]; modern methods include balanced POD [40, 42], spectral POD [43], DMD [44–46], the Koopman decomposition [47–49], resolvent analysis [50, 51], and neural-network-based autoencoders [52–54]. Data-driven techniques, including modern machine learning, are also being widely applied to discover dynamical systems models of complex physical systems [55–68], with a particular emphasis on hybrid physics-inspired or physics-informed machine learning [34, 59, 69–73]. In fluid mechanics, sparse model discovery has been used to develop interpretable nonlinear models that enforce known physics by construction [34, 57, 74]. Particular emphasis is put on understanding how these models alter or retain the "direct energy cascade" coming from the interaction of terms in the Navier-Stokes equations; in Hall-MHD there are also inverse cascades and bidirectional cascades [75], complicating the analysis of model stability and generalizability.

In this work we develop theoretical foundations for principled projection-based and data-driven plasma models. In fluid mechanics, careful development of a dimensionalized inner product enabled the extension of POD from incompressible to compressible fluid flows [76]. It is also common in fluid mechanics to obtain nonlinear reduced-order models by Galerkin projection of the Navier-Stokes equations onto POD modes, making it possible to enforce known symmetries and conservation laws, such as conservation of energy [19, 77–79]. These symmetries have recently been utilized to constrain the identification of data-driven fluid models [34]. The present work extends and unifies these three innovations for compressible plasmas, enabling a wealth of advanced modeling and control machinery. The culmination of this work is illustrated by accurately forecasting the evolution of a 3D isothermal Hall-MHD simulation of the HIT-SI experiment [80], described in detail in Appendix A.

In Section II, we propose a formalism for reduced plasma models, and in Section III we derive how global conservation laws in MHD manifest in the subsequent reduced-order models. After establishing a framework for projection-based model reduction in Hall-MHD, in Section IV we utilize physics-constrained machine learning to discover low-dimensional plasma models directly from data. Figure 1 summarizes the steps for using plasma data with projection-based model reduction and data-driven system identification. As described in detail in Sections II and IV, data is collected, projected onto a low-dimensional basis, constrained by any known physical laws, and analyzed to find a descriptive model for its evolution. This last task may be achieved through projection-based, data-driven, or hybrid methods.

II. PROJECTION-BASED REDUCED-ORDER MODELS IN MHD

Despite the many ways to obtain low-dimensional models, projection-based model reduction such as Galerkin methods have seen significant use and remarkable success in fluid mechanics. This success stems from the direct ties with the first-principles physics, lending interpretability to Galerkin models so that physical characteristics such as linear stability can be investigated. When the low-dimensional basis is obtained from the POD, the resulting models are referred to as POD-Galerkin models. Since the 1990's, it has been common in fluid mechan-

$$q(\boldsymbol{x},t) = \begin{bmatrix} \boldsymbol{B}_u \\ \boldsymbol{B} \\ B_T \end{bmatrix}$$

$$q \approx \bar{q}(\boldsymbol{x}) + \sum_j \boldsymbol{\chi}_j(\boldsymbol{x}) a_j(t_k)$$

3D MHD system   injector

$Z = 0$

Dimensional 3D POD

System

B_y

injector   2D visual

Gauss 500 ... −500

Model

B_y

a_1 a_2 a_3

B_x

χ_1 χ_2 χ_3

$\dot{\boldsymbol{a}} = \boldsymbol{f}(\boldsymbol{a})$

Galerkin Projection

$\dot{\boldsymbol{a}} \quad = \quad \boldsymbol{\Theta}(\boldsymbol{a})$

Ξ

$\xi_i^{a_j} = -\xi_j^{a_i}$

t

a_3 a_2 a_1

(a)   (b)   (c)
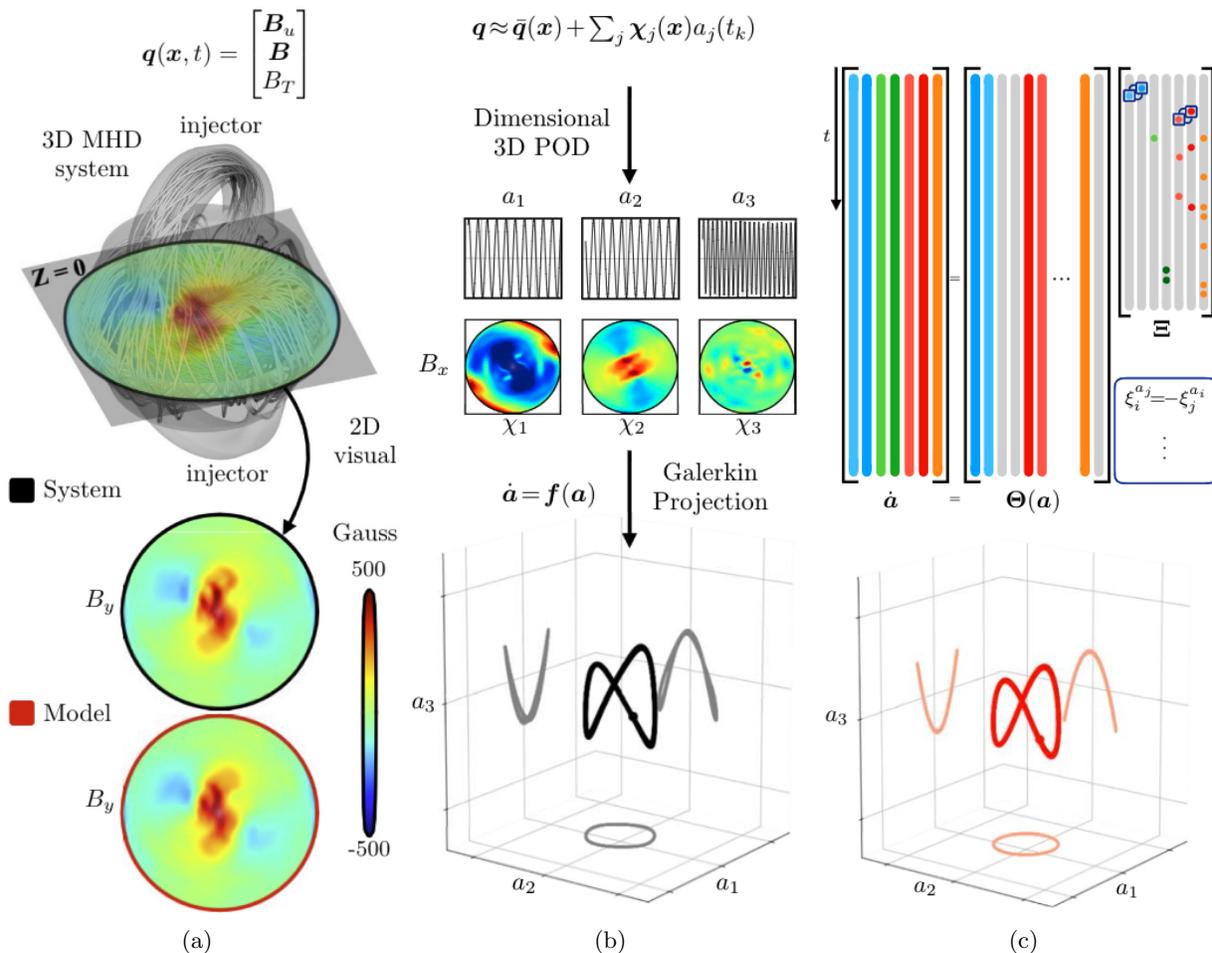
FIG. 1: Proposed approach for filling in lower rungs of the plasma model hierarchy: (a) Collect data, (b) perform projection-based model reduction (c) discover data-driven models using physics-constrained system identification. Throughout this work, see the online version of this manuscript for the color bars and other color-coding.

ics to obtain these nonlinear reduced-order models by Galerkin projection of the Navier-Stokes equations onto POD modes. Further developments increased the utility of these methods, including a dimensionalized inner product which enabled the extension of POD-Galerkin models from incompressible to compressible fluids [76]. Towards the goal of extending these developments into plasma physics, we first need to define a new inner product for MHD plasmas in Section II A and then illustrate how the subsequent proper orthogonal decomposition is performed for plasma datasets in Section II B. This formalism facilitates our the derivation of a POD-Galerkin model for Hall-MHD in Section II C.

The present work adapts and extends these innovations for plasmas, enabling a wealth of advanced modeling and control machinery. For clarity and robust connection with the Galerkin literature in fluid mechanics, we primarily consider models which are quadratic in nonlinearity. This includes ideal MHD, incompressible Hall-MHD, and variants such as compressible Hall-MHD with a slowly time-varying density, which together describe a broad class of space and laboratory plasmas [81–86].

A. An MHD energy inner product

Traditional use of the POD on the MHD fields (velocity, magnetic, and temperature) would either require separate decompositions for $\boldsymbol{u}$, $\boldsymbol{B}$, and T, or an arbitrary choice of dimensionalization. However, separate decompositions of the fields obfuscates the interpretability and increases the complexity of a low-dimensional model, and choosing the units of the combined matrix of measurement data can have a significant impact on the performance and energy spectrum of the resulting POD basis. Inspired by the inner product defined for compressible fluids [76], we define an inner product for MHD through

$$\boldsymbol{q}(\boldsymbol{x},t) = \begin{bmatrix} \boldsymbol{B}_u \\ \boldsymbol{B} \\ B_T \end{bmatrix}, \quad \boldsymbol{B}_u = \sqrt{\rho\mu_0}\,\boldsymbol{u}, \quad B_T = \sqrt{\frac{4\rho\mu_0 k_b T}{m_i(\gamma-1)}}. \quad (1)$$

Here ρ is the mass density, k_b is Boltzmann's constant, μ_0 is the permeability, m_i is the ion mass, γ is the adiabatic index, $p = 2\rho T / m_i$ is the plasma pressure, and the total plasma energy is

$$W = \frac{1}{2\mu_0} \langle \boldsymbol{q}, \boldsymbol{q} \rangle = \int \left(\frac{1}{2} \rho u^2 + \frac{B^2}{2\mu_0} + \frac{p}{\gamma - 1} \right) d^3 \boldsymbol{x}. \qquad (2)$$

Normalizing the MHD fields to magnetic field units produces a natural interpretation of inner products of the vector $\boldsymbol{q}$ as the total plasma energy. This formulation is also useful because reduced order models built for $\boldsymbol{q}$ can be constrained by conservation of energy via Eq. (2), as we illustrate in detail in Section III.

B. Proper orthogonal decomposition for plasma datasets

The POD is already used extensively for interpreting plasma physics data across a range of parameter regimes [28, 87–90], but some formalism is required to effectively use it for modeling and forecasting. For POD, a set of point measurements at time t_k are arranged in a vector $\boldsymbol{q}_k \in \mathbb{R}^D$, called a snapshot, where the dimension D is the product of the number of spatial locations and the number of variables measured at each point. For instance, we could have obtained the magnetic field data from $D/3$ magnetic probes that measure the magnetic field components at a fixed location and sampling rate. Now we assume that the data is sampled at some times t_1, t_2, ..., t_M, arranged in a matrix $\boldsymbol{X} \in \mathbb{R}^{D \times M}$, and the average in time $\bar{\boldsymbol{q}}$ is subtracted off. The singular value decomposition (SVD) provides a low-rank approximation

$$\boldsymbol{X} = \begin{bmatrix} q_1(t_1) & q_1(t_2) & \cdots & q_1(t_M) \\ q_2(t_1) & q_2(t_2) & \cdots & q_2(t_M) \\ \vdots & \vdots & \ddots & \vdots \\ q_D(t_1) & q_D(t_2) & \cdots & q_D(t_M) \end{bmatrix} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^*, \quad (3)$$

where $\boldsymbol{U} \in \mathbb{R}^{D \times D}$ and $\boldsymbol{V} \in \mathbb{R}^{M \times M}$ are unitary matrices, and $\boldsymbol{\Sigma} \in \mathbb{R}^{D \times M}$ is a diagonal matrix containing non-negative and decreasing entries s_{jj} called the singular values of $\boldsymbol{X}$. $\boldsymbol{V}^*$ denotes the complex-conjugate transpose of $\boldsymbol{V}$. The singular values indicate the relative importance of the corresponding columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ for describing the spatio-temporal structure of $\boldsymbol{X}$.

It is often possible to discard small values of $\boldsymbol{\Sigma}$, resulting in a truncated matrix $\boldsymbol{\Sigma}_r \in \mathbb{R}^{r \times r}$. With the first $r \ll \min(D, M)$ columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, denoted $\boldsymbol{U}_r$ and $\boldsymbol{V}_r$, we have

$$\boldsymbol{X} \approx \boldsymbol{U}_r \boldsymbol{\Sigma}_r \boldsymbol{V}_r^*. \qquad (4)$$

The truncation rank r is typically chosen to balance ac-

curacy and complexity [21]. The computational complexity of the SVD is $\mathcal{O}(DM^2 + M^3)$ [91], although there are randomized singular value decompositions [92–94] for very large problems that can be as fast as $\mathcal{O}(DM\log(r))$. Therefore, even for $r \gg 1$, the SVD typically produces significant computational speedup over codes which evolve the full spatio-temporal dynamics. The computational speed [91, 94] of the SVD also enables online computations to update a model for real-time control.

To proceed, a well-defined SVD requires that the measurements in $\boldsymbol{X}$ have the same physical dimensions. With a dimensionalized measurement vector $\boldsymbol{q}$, the matrix $\boldsymbol{X}^* \boldsymbol{X}$ satisfies

$$\boldsymbol{X}^* \boldsymbol{X} \approx \langle \boldsymbol{q}(t_k), \boldsymbol{q}(t_m) \rangle, \qquad k, m \in \{1, 2, ..., M\}. \qquad (5)$$

The equality is not exact because the inner product (an integral) is approximated by the discrete sum from the matrix product of $\boldsymbol{X}^* \boldsymbol{X}$, but it is important that we can relate the matrix $\boldsymbol{X}$ to the total plasma energy through Eq. (2). The temporal SVD modes, or chronos, $\boldsymbol{v}_j$ are the columns of $\boldsymbol{V}_r$. The spatial modes, or topos, $\boldsymbol{\chi}$ form the columns of $\boldsymbol{U}_r$. We scale $a_j(t_k) = v_j(t_k)/\sum_{j=1}^r \max_k |v_j(t_k)|$. Finally,

$$\boldsymbol{q}(\boldsymbol{x}_i, t_k) \approx \bar{\boldsymbol{q}}(\boldsymbol{x}_i) + \sum_{j=1}^r \boldsymbol{\chi}_j(\boldsymbol{x}_i) a_j(t_k). \qquad (6)$$

We have absorbed the normalization of $a_j(t_k)$ and the singular values into the definition of $\boldsymbol{\chi}_j(\boldsymbol{x}_i)$. By construction $\langle \boldsymbol{\chi}_i, \boldsymbol{\chi}_j \rangle \propto \delta_{ij}$. Note that, in principle, we could have expanded $\boldsymbol{q}$ in any set of modes, although orthonormal modes are preferred because this property facilitates the analysis in Section III. Non-orthogonal modes are also suitable, but introduce a complication in the form of a mass matrix [95]. The advantage of the POD basis is that the modes are ordered by energy content; a truncation of the system still captures a majority of the dynamics. A separate POD of each of the MHD fields would lead to three sets of POD modes with independent time dynamics and mixed orthogonality properties. In contrast, our approach captures all the fields simultaneously, resulting in a single set of modes $a_i(t)$ in Eq. (6).

An example of this decomposition is illustrated in Fig. 2 for a 3D isothermal Hall-MHD simulation, described in detail in Appendix A and modeled in Section IV B; the dominant dynamics are harmonics and sub-harmonics of a driving frequency imposed by two actuating injectors on the top and bottom of the device. In general, examining the structure and symmetry in the spatial and temporal POD modes can inform physical understanding. For instance, in Fig. 2, the short-wavelength structures exhibited in the 3D spatial modes derive both from dispersive whistler waves via the Hall term and the small characteristic scale associated with the injectors (actuators). The steep fall-off in the singular values also indicates that models of only the first few modes would be enough to accurately forecast and control the domi-
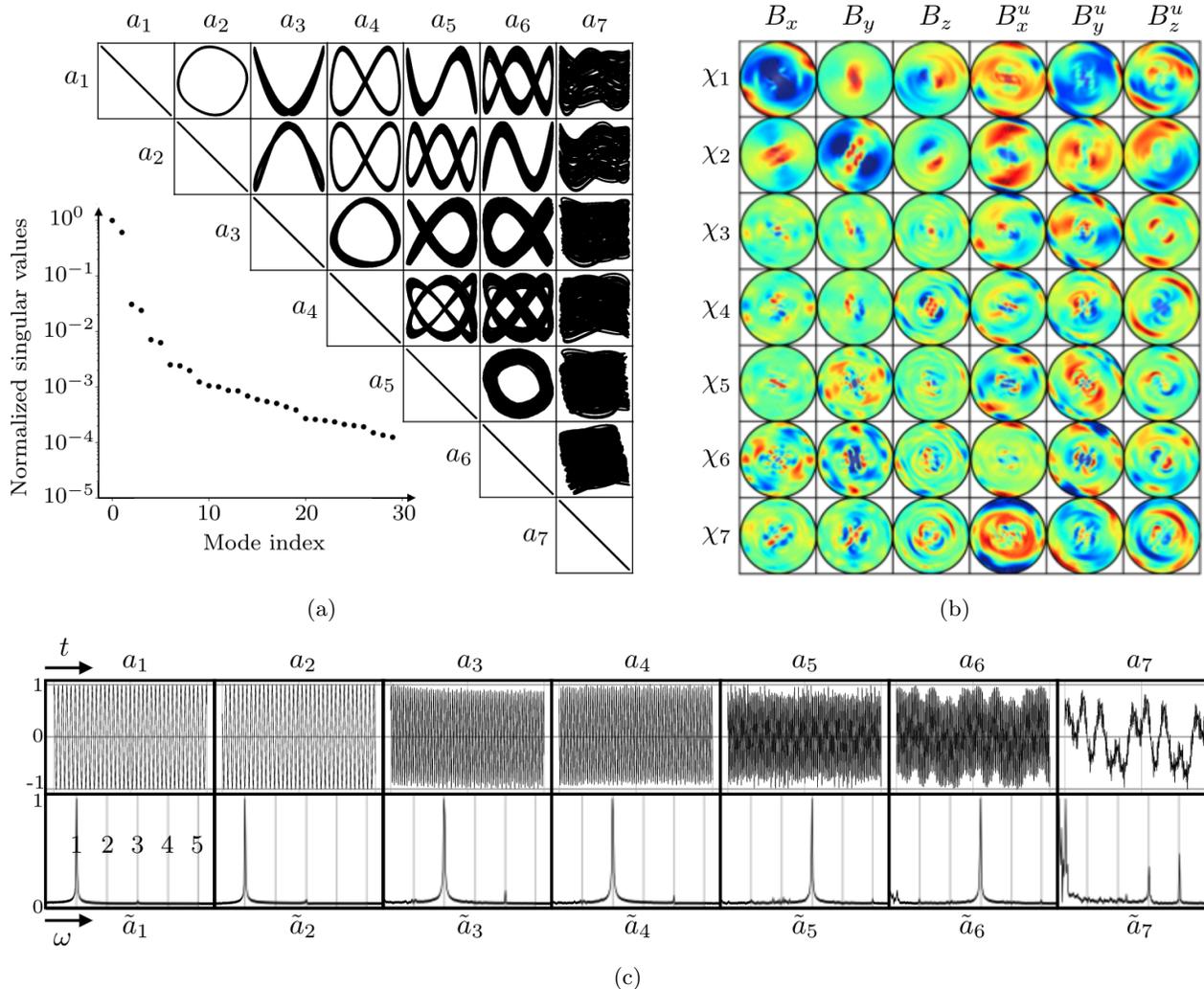
FIG. 2: The first seven POD modes for a 3D isothermal Hall-MHD simulation of the HIT-SI device detailed in Appendix A. The mean-flow-subtracted chronos indicate that the primary dynamics are forcing at the driving injector frequency and its harmonics; (a) Mode pair trajectories evolved in time and the corresponding singular values; (b) 3D spatial modes in the $Z=0$ midplane illustrate a complicated mix of length scales; (c) Normalized temporal modes and corresponding Fourier transforms exhibit harmonics of the driving frequency, labeled 1-5.

nant dynamics. While we show stable and accurate $r=16$ models in Section IV B to illustrate the strengths of our methods for more complicated dynamics, in our test case it is true that more modest models of three or four modes can already achieve reasonable forecasting accuracies.

C. POD-Galerkin models

Now that we have an expansion of the fields in a low-dimensional basis in Eq. (6), we can project the Hall-MHD equations onto these POD modes in order to obtain a POD-Galerkin model. Hall-MHD, using the definitions of the electromagnetic current $\mu_0 \boldsymbol{J} = \nabla \times \boldsymbol{B}$, electron fluid velocity $\boldsymbol{u}_e = \boldsymbol{u} - \boldsymbol{J}/ne$, electron and ion temperature $T_e = T_i = T$, and the definitions in Eq. (1), can be written:

$$\dot{\rho} = -\nabla \cdot \left(\sqrt{\frac{\rho}{\mu_0}} \, \boldsymbol{B}_u \right), \tag{7}$$

$$\dot{\boldsymbol{B}} = \nabla \times \left[\frac{1}{\sqrt{\rho \mu_0}} \left(\boldsymbol{B}_u \times \boldsymbol{B} - d_i((\nabla \times \boldsymbol{B}) \times \boldsymbol{B}) \right) \right] + \eta_0 \rho^{\frac{3}{2}} B_T^{-3} \nabla^2 \boldsymbol{B} + \frac{d_i}{\sqrt{\rho \mu_0}} (\gamma - 1) B_T \nabla B_T \times \frac{\nabla \rho}{2\rho},$$

$$\dot{\boldsymbol{B}}_u = -\frac{1}{\sqrt{\rho \mu_0}} \left[\frac{1}{2} \boldsymbol{B}_u \nabla \cdot \boldsymbol{B}_u + \boldsymbol{B}_u \cdot \nabla \boldsymbol{B}_u - \frac{1}{4\rho} \boldsymbol{B}_u (\nabla \rho \cdot \boldsymbol{B}_u) - (\nabla \times \boldsymbol{B}) \times \boldsymbol{B} + \frac{(\gamma - 1) B_T^2}{2} \frac{\nabla \rho}{\rho} - (\gamma - 1) B_T \nabla B_T \right]$$

$$+ \nu \left[\nabla^2 \boldsymbol{B}_u - \frac{\nabla^2 \rho}{2\rho} \boldsymbol{B}_u + \frac{3 \boldsymbol{B}_u}{4\rho^2} (\nabla \rho)^2 + \frac{1}{\rho} (\nabla \rho \cdot \nabla) \boldsymbol{B}_u - \frac{1}{6\rho} \nabla (\nabla \rho \cdot \boldsymbol{B}_u) + \frac{1}{4\rho^2} (\nabla \rho \cdot \boldsymbol{B}_u) \nabla \rho + \frac{1}{3} \nabla (\nabla \cdot \boldsymbol{B}_u) - \frac{1}{6\rho} (\nabla \cdot \boldsymbol{B}_u) \nabla \rho \right],$$

$$\dot{B}_T = -\frac{1}{\sqrt{\rho \mu_0}} \left[\boldsymbol{B}_u \cdot \nabla B_T - \gamma B_T (\nabla \cdot \boldsymbol{B}_u - \frac{\nabla \rho}{2\rho} \cdot \boldsymbol{B}_u) \right] - \frac{2}{B_T} [\nabla \cdot \boldsymbol{h} + Q_{\text{visc}}] + 4\eta_0 \rho^{\frac{3}{2}} B_T^{-4} (\nabla \times \boldsymbol{B})^2,$$

where we have used that $\nabla \cdot \boldsymbol{B} = 0$ and the definitions of the heat flux $\boldsymbol{h}$ and viscous heating Q_{visc},

$$\boldsymbol{h} = -\frac{(\gamma - 1) B_T}{4} \left[\chi_\parallel \hat{\boldsymbol{b}} \hat{\boldsymbol{b}} + \chi_\perp \left(\boldsymbol{I} - \hat{\boldsymbol{b}} \hat{\boldsymbol{b}} \right) \right] \cdot \left(\nabla B_T - B_T \frac{\nabla \rho}{\rho} \right), \tag{8}$$

$$Q_{\text{visc}} = -\tilde{\nu} (\nabla \boldsymbol{B}_u - \boldsymbol{B}_u \frac{\nabla \rho}{2\rho})^T : \left[(\nabla \boldsymbol{B}_u - \boldsymbol{B}_u \frac{\nabla \rho}{2\rho}) + (\nabla \boldsymbol{B}_u - \boldsymbol{B}_u \frac{\nabla \rho}{2\rho})^T - \frac{2}{3} \boldsymbol{I} (\nabla \cdot \boldsymbol{B}_u - \boldsymbol{B}_u \cdot \frac{\nabla \rho}{2\rho}) \right].$$

Here $\nu = \tilde{\nu}/\rho$ is the dynamic viscosity, $\eta = \eta_0 T^{-\frac{3}{2}} \propto \rho^{\frac{3}{2}} B_T^{-3}$ is the Spitzer resistivity [96], $d_i = m_i/(e\sqrt{\rho \mu_0})$ is the ion inertial length, and $\chi_\perp$ and $\chi_\parallel$ are the anisotropic Braginskii thermal diffusivities with temperature and magnetic field dependencies [97]. Although many of the nonlinear terms are only quadratic in $\boldsymbol{q}$, we consider the isothermal limit and limit of time-independent density to restrict ourselves to the pure quadratic nonlinear case:

$$\dot{\boldsymbol{q}} = \boldsymbol{C} + \boldsymbol{L}(\boldsymbol{q}) + \boldsymbol{Q}(\boldsymbol{q}, \boldsymbol{q}), \tag{9}$$

$$\boldsymbol{C} = \begin{bmatrix} -\frac{(\gamma - 1) B_T^2}{2} \sqrt{\frac{1}{\mu_0 \rho}} \frac{\nabla \rho}{\rho} \\ 0 \\ 0 \end{bmatrix},$$

$$\boldsymbol{Q}(\boldsymbol{q}, \boldsymbol{q}) = \begin{bmatrix} -\frac{1}{\sqrt{\rho \mu_0}} \left(\boldsymbol{B}_u (\nabla \cdot \boldsymbol{B}_u) + \boldsymbol{B}_u \cdot \nabla \boldsymbol{B}_u - (\nabla \times \boldsymbol{B}) \times \boldsymbol{B} \right) \\ \nabla \times \left(\frac{1}{\sqrt{\rho \mu_0}} \left(\boldsymbol{B}_u \times \boldsymbol{B} - d_i (\nabla \times \boldsymbol{B}) \times \boldsymbol{B} \right) \right) \\ 0 \end{bmatrix},$$

$$\boldsymbol{L}(\boldsymbol{q}) = \begin{bmatrix} \nu \left(\nabla^2 \boldsymbol{B}_u - \cdots - \frac{1}{3\rho} (\nabla \cdot \boldsymbol{B}_u) \nabla \rho \right) \\ \frac{\eta}{\mu_0} \nabla^2 \boldsymbol{B} \\ 0 \end{bmatrix}.$$

Increasingly sophisticated models may be tractable in future work, since the data-driven approach that we adopt in Section IV is not limited to quadratic nonlinearities. Other models, such as those assuming incompressibility and finite temperature evolution, can also be derived straightforwardly from the results here. Substituting Eq. (6) into Eq. (9) and utilizing the orthonormality of the $\boldsymbol{\chi}_j$ produces:

$$\dot{a}_i(t) = C_i^0 + \sum_{j=1}^r L_{ij}^0 a_j + \sum_{j,k=1}^r Q_{ijk}^0 a_j a_k, \tag{10}$$

$$C_i^0 = \langle \boldsymbol{C} + \boldsymbol{L}(\bar{\boldsymbol{q}}) + \boldsymbol{Q}(\bar{\boldsymbol{q}}, \bar{\boldsymbol{q}}), \boldsymbol{\chi}_i \rangle,$$

$$L_{ij}^0 = \langle \boldsymbol{L}(\boldsymbol{\chi}_j) + \boldsymbol{Q}(\bar{\boldsymbol{q}}, \boldsymbol{\chi}_j) + \boldsymbol{Q}(\boldsymbol{\chi}_j, \bar{\boldsymbol{q}}), \boldsymbol{\chi}_i \rangle,$$

$$Q_{ijk}^0 = \langle \boldsymbol{Q}(\boldsymbol{\chi}_j, \boldsymbol{\chi}_k), \boldsymbol{\chi}_i \rangle.$$

The model is quadratic in the temporal POD modes $a_i(t)$. The zero superscript is meant to distinguish the coefficient tensors C_i^0, L_{ij}^0, and Q_{ijk}^0 from the spatio-temporal operators $\boldsymbol{C}$, $\boldsymbol{L}$, and $\boldsymbol{Q}$. If $\bar{\boldsymbol{q}}$ satisfies the steady-state MHD equations, then $C_i^0 = 0$ for all i. This is a reasonable assumption for any approximately steady-state device, such as a tokamak, which can be sustained for many characteristic timescales. In contrast to Eq. (10), a Galerkin model based on separate POD expansions for each field would involve significant mixing and a lack of orthonormality $\langle \boldsymbol{\chi}_i^{\boldsymbol{u}}, \boldsymbol{\chi}_j^{\boldsymbol{B}} \rangle \neq \delta_{ij}$ between the POD modes. Although Eq. (10) contains only quadratic nonlinearities, the influence of truncated low-energy modes can sometimes be modeled with cubic nonlinearities in the Galerkin model [17, 34].

D. Relation to Fourier-Galerkin methods

Similar analytic Fourier-Galerkin models (also called MHD shell models) have been used for modeling incompressible MHD turbulence [98]. Shell models in MHD have primarily been used to describe the statistics of homogeneous and isotropic turbulence in spectral space, rather than as reduced order models [99]. The differ-

ences in application likely stem from shell models preserving the MHD invariants within each triad of wave vectors but POD models providing a dataset-tailored and energy-optimal basis. However, in various homogeneous and symmetric limits the POD reduces to the Fourier basis [39, 100]. In both Fourier-Galerkin and POD-Galerkin models, truncation of the model at some rank r can lead to under-resolving the dissipation rate or approximately breaking the global conservation laws, and a closure scheme may be required to re-introduce the full dissipation. Additionally, if energy is not conserved, as in some dissipative MHD models, the stability of the truncated system is no longer guaranteed. Two advantages of the data-driven approach in Section IV over either POD-Galerkin or Fourier-Galerkin is that 1) we need not laboriously compute the coefficients in Eq. (10) from full state knowledge, and 2) we can enforce global energy or cross-helicity conservation directly into the truncated model (even without energy conservation, we may be able to enforce other generic stability properties [101]). Lastly, preserving the features of the direct energy cascade in truncated Galerkin and data-driven models for incompressible fluid flows is a current field of research. Since even this "simple case" is unsettled, there is much research to be done regarding the preservation of direct, inverse, and even bidirectional cascades [75] in truncated Galerkin models for Hall-MHD beyond highly simplified cases such as isotropic, incompressible, isothermal Hall-MHD turbulence on simple geometries.

III. DERIVING CONSTRAINTS ON PROJECTION-BASED MODELS

We have successfully obtained a POD-Galerkin model for the dynamic fields in Hall-MHD. However, there is substantial additional structure in the coefficients in Eq. (10) because local and global MHD conservation laws are in principle retained in this low-dimensional basis. Vanishing $\nabla \cdot \boldsymbol{B}$ and the linear independence of the temporal POD modes produce

$$\nabla \cdot \boldsymbol{\chi}_i^B = 0, \quad \forall i. \tag{11}$$

In other words, there is a local divergence constraint for each of the $\boldsymbol{\chi}_i^B$, but this does not produce insight into the coefficients defined in Eq. (10) In contrast, global energy conservation produces substantial constraints on the structure of the Galerkin model coefficients.

A. Global conservation of energy

For an examination of the global conservation laws, we consider isothermal Hall-MHD with a very slowly time-varying density. This model reduces to ideal MHD and incompressible resistive or Hall MHD in the appropriate limits, and produces (Galtier [102] Eq. 3.22)

$$\frac{\partial W}{\partial t} = -\int \left[\tilde{\nu}(\nabla \times \boldsymbol{u})^2 + \frac{\eta}{\mu_0}(\nabla \times \boldsymbol{B})^2 + \frac{4}{3}\tilde{\nu}(\nabla \cdot \boldsymbol{u})^2 \right] d^3\boldsymbol{x} \tag{12}$$

$$-\oint \left[\left(\frac{1}{2}\rho u^2 + p \right) \boldsymbol{u} + \boldsymbol{P} - \frac{4}{3}\tilde{\nu}(\nabla \cdot \boldsymbol{u})\boldsymbol{u} - \tilde{\nu}\boldsymbol{u} \times (\nabla \times \boldsymbol{u}) \right] \cdot \hat{\boldsymbol{n}} dS.$$

Here $\hat{\boldsymbol{n}}$ is a unit normal vector to the boundary, and

$$\boldsymbol{P} = \frac{1}{\mu_0}\boldsymbol{E} \times \boldsymbol{B} = \frac{\boldsymbol{u}_e}{\mu_0} \cdot (B^2\boldsymbol{I} - \boldsymbol{B}\boldsymbol{B}) - \frac{\eta}{\mu_0^2}(\nabla \times \boldsymbol{B}) \times \boldsymbol{B}, \tag{13}$$

is the Poynting vector ($\boldsymbol{E}$ is the electric field), which is often an imposed and experimentally-known function of space and time. Omission of the Hall term changes $\boldsymbol{u}_e$ to $\boldsymbol{u}$ in Eq. (13). Even with temperature evolution, the electron diamagnetic term in $\boldsymbol{P}$ does not alter the energy balance if Dirichlet conditions are used for ρ and T. To simplify, we assume that $\boldsymbol{u} \cdot \hat{\boldsymbol{n}} = \boldsymbol{u} \times \hat{\boldsymbol{n}} = 0$, $\boldsymbol{J} \cdot \hat{\boldsymbol{n}} = 0$, and $\boldsymbol{B} \cdot \hat{\boldsymbol{n}} = 0$ at the wall, consistent with the Hall-MHD HIT-SI simulation described in Appendix A and modeled in Section IV B. Now assume steady-state, define $a_0(t) = 1$, and substitute Eq. (6) into Eq. (12),

$$0 \approx \frac{\partial W}{\partial t} = \oint \frac{\eta}{\mu_0^2}((\nabla \times \boldsymbol{B}) \times \boldsymbol{B}) \cdot \hat{\boldsymbol{n}} dS - \int \left[\frac{\nu}{\mu_0}(\nabla \times \boldsymbol{B}_u - \frac{\nabla \rho}{2\rho} \times \boldsymbol{B}_u)^2 + \frac{\eta}{\mu_0^2}(\nabla \times \boldsymbol{B})^2 + \frac{4}{3}\frac{\nu}{\mu_0}(\nabla \cdot \boldsymbol{B}_u - \frac{\nabla \rho}{2\rho} \cdot \boldsymbol{B}_u)^2 \right] d^3\boldsymbol{x}, \tag{14}$$

$$= W^C + \sum_{i=1}^{r} W_i^L a_i + \sum_{i,j=1}^{r} W_{ij}^Q a_i a_j = \sum_{i,j=0}^{r} W_{ij}^Q a_i a_j,$$

We have padded the matrix in the last step so that $W_{0i}^Q=0$, $W_{i0}^Q=W_i^L$ for $i\in\{1,...,r\}$, and $W_{00}^Q=W^C$. Eq. (14) is generally satisfied for anti-symmetric W_{ij}^Q, from which it follows that

$$0=W_{00}^Q=\frac{\eta}{\mu_0}\oint\left[(\nabla\times\bar{\boldsymbol{B}})\times\bar{\boldsymbol{B}}\right]\cdot\hat{\boldsymbol{n}}dS-\int\left[\nu(\nabla\times\bar{\boldsymbol{B}}_u-\frac{\nabla\rho}{2\rho}\times\bar{\boldsymbol{B}}_u)^2+\frac{\eta}{\mu_0}(\nabla\times\bar{\boldsymbol{B}})^2+\frac{4}{3}\nu(\nabla\cdot\bar{\boldsymbol{B}}_u-\frac{\nabla\rho}{2\rho}\cdot\bar{\boldsymbol{B}}_u)^2\right]d^3\boldsymbol{x},\tag{15}$$

$$0=W_{i0}^Q=\frac{\eta}{\mu_0}\oint\left[(\nabla\times\bar{\boldsymbol{B}})\times\boldsymbol{\chi}_i^B+(\nabla\times\boldsymbol{\chi}_i^B)\times\bar{\boldsymbol{B}}\right]\cdot\hat{\boldsymbol{n}}dS$$

$$-2\int\left[\nu(\nabla\times\bar{\boldsymbol{B}}_u-\frac{\nabla\rho}{2\rho}\times\bar{\boldsymbol{B}}_u)\cdot(\nabla\times\boldsymbol{\chi}_i^{B_u}-\frac{\nabla\rho}{2\rho}\times\boldsymbol{\chi}_i^{B_u})+\frac{\eta}{\mu_0}(\nabla\times\bar{\boldsymbol{B}})\cdot(\nabla\times\boldsymbol{\chi}_i^B)+\frac{4}{3}\nu(\nabla\cdot\bar{\boldsymbol{B}}_u-\frac{\nabla\rho}{2\rho}\cdot\bar{\boldsymbol{B}}_u)\cdot(\nabla\cdot\boldsymbol{\chi}_i^{B_u}-\frac{\nabla\rho}{2\rho}\cdot\boldsymbol{\chi}_i^{B_u})\right]d^3\boldsymbol{x},$$

$$W_{ij}^Q=-W_{ji}^Q=\frac{\eta}{\mu_0}\oint\left[(\nabla\times\boldsymbol{\chi}_i^B)\times\boldsymbol{\chi}_j^B\right]\cdot\hat{\boldsymbol{n}}dS$$

$$-\int\left[\nu(\nabla\times\boldsymbol{\chi}_i^{B_u}-\frac{\nabla\rho}{2\rho}\times\boldsymbol{\chi}_i^{B_u})\cdot(\nabla\times\boldsymbol{\chi}_j^{B_u}-\frac{\nabla\rho}{2\rho}\times\boldsymbol{\chi}_j^{B_u})+\frac{\eta}{\mu_0}(\nabla\times\boldsymbol{\chi}_i^B)\cdot(\nabla\times\boldsymbol{\chi}_j^B)+\frac{4}{3}\nu(\nabla\cdot\boldsymbol{\chi}_i^{B_u}-\frac{\nabla\rho}{2\rho}\cdot\boldsymbol{\chi}_i^{B_u})\cdot(\nabla\cdot\boldsymbol{\chi}_j^{B_u}-\frac{\nabla\rho}{2\rho}\cdot\boldsymbol{\chi}_j^{B_u})\right]d^3\boldsymbol{x}.$$

Evaluating Eq. (15) and the Galerkin coefficients in Eq. (10) relies on the existence of all of the $\nabla\times\boldsymbol{\chi}_i$. These spatial POD modes are evaluated on a discrete set of spatial locations, but in practice we can always choose an interpolation such that the curl operator is well-defined. In such a case, $\nabla\times\boldsymbol{\chi}_i^B$ and $\nabla\times\boldsymbol{\chi}_i^{B_u}$ have natural interpretations as the spatial POD modes of the electromagnetic current and vorticity fields. However, in the present work these computations only serve as formal manipulations so we need not evaluate these curls; our data-driven method in Section IV uses sparse regression to determine these coefficients from data. Continuing on with our analysis, we can compute $a_i\dot{a}_i$ for $i\in\{1,...,r\}$,

$$a_i\dot{a}_i=\sum_{i,j=1}^r a_i\frac{\partial a_j}{\partial t}\int\chi_i\chi_j d^3\boldsymbol{x}=\int\frac{1}{2}\frac{\partial q^2}{\partial t}d^3\boldsymbol{x}=\frac{\partial W}{\partial t},\tag{16}$$

$$a_i\dot{a}_i=a_iC_i^0+a_iL_{ij}^0a_j+a_iQ_{ijk}^0a_ja_k,\ i,j,k\in\{1,...,r\}.\tag{17}$$

First, note that $W_{i0}^Q=0$ produces $C_i^0=0$ for all $i\in\{1,...,r\}$. There are no constant terms in the Galerkin model. This is a physical consequence of our assumption that $\bar{q}$ is steady-state; nonzero constant terms would imply the possibility of unbounded growth in the energy norm. The anti-symmetry of W_{ij}^Q for $i,j\in\{1,...,r\}$ constrains the quadratic structure of the energy $\boldsymbol{a}^T\boldsymbol{a}$,

$$\boldsymbol{a}^T\boldsymbol{L}^0\boldsymbol{a}\approx0.\tag{18}$$

This physical interpretation is also clear; if the plasma is steady-state but has finite dissipation, the input power,

here manifested through a purely quadratic Poynting flux $\boldsymbol{P}\propto\eta\boldsymbol{J}\times\boldsymbol{B}$, must be balancing these losses. Finally, there are no cubic terms in the energy, implying

$$\boldsymbol{a}^T\boldsymbol{Q}^0\boldsymbol{a}\boldsymbol{a}=0,\tag{19}$$

or equivalently,

$$Q_{ijk}^0+Q_{jik}^0+Q_{kij}^0=0.\tag{20}$$

In other words, the quadratic nonlinearities in the Galerkin model of Eq. (10) are energy-preserving; this conclusion did not rely on any assumption of steady-state and energy-preserving structure in other quadratic nonlinearities is well-studied in fluid mechanics [74, 79, 101]. The lack of nonlinear energy losses is a physical consequence coming from the boundary conditions $\boldsymbol{B}\cdot\hat{\boldsymbol{n}}=0$, $\boldsymbol{J}\cdot\hat{\boldsymbol{n}}=0$, $\boldsymbol{u}\cdot\hat{\boldsymbol{n}}=\boldsymbol{u}\times\hat{\boldsymbol{n}}=0$ (and constant temperature).

B. Global conservation of cross-helicity

An analogous derivation can be done to further constrain the model-building for systems which conserve cross-helicity, although this is inappropriate for the Hall-MHD HIT-SI simulation in Section IV B. Consider the local form of cross-helicity $H_c=\boldsymbol{u}\cdot\boldsymbol{B}$. Using Galtier [102] Eq. (3.36),

$$\frac{\partial H_c}{\partial t}=-\nabla\cdot\left[\left(\frac{u^2}{2}+\frac{\gamma p}{(\gamma-1)\rho}\right)\boldsymbol{B}+\boldsymbol{u}\times(\boldsymbol{u}\times\boldsymbol{B})-\frac{d_i}{\sqrt{\rho\mu_0}}\boldsymbol{u}\times\left((\nabla\times\boldsymbol{B})\times\boldsymbol{B}\right)-\eta\boldsymbol{u}\times(\nabla\times\boldsymbol{B})\right]\tag{21}$$

$$+\nu\nabla\cdot\left(\boldsymbol{B}\times(\nabla\times\boldsymbol{u})+\frac{4}{3}(\nabla\cdot\boldsymbol{u})\boldsymbol{B}\right)-\frac{d_i}{\sqrt{\rho\mu_0}}(\nabla\times\boldsymbol{u})\cdot\left((\nabla\times\boldsymbol{B})\times\boldsymbol{B}\right)-(\eta+\nu)(\nabla\times\boldsymbol{B})\cdot(\nabla\times\boldsymbol{u}).$$

Consider again the simplifying case $\boldsymbol{J}\cdot\hat{\boldsymbol{n}}=0$, $\boldsymbol{B}\cdot\hat{\boldsymbol{n}}=0$, and $\boldsymbol{u}\cdot\hat{\boldsymbol{n}}=\boldsymbol{u}\times\hat{\boldsymbol{n}}=0$. If global cross-helicity is conserved, the integral form is

$$0\approx\int\frac{\partial H_c}{\partial t}d^3\boldsymbol{x}=\int\left[\nu\frac{\nabla\rho}{\rho}\cdot\left(\boldsymbol{B}\times(\nabla\times\boldsymbol{u})+\frac{4}{3}(\nabla\cdot\boldsymbol{u})\boldsymbol{B}\right)-\frac{d_i}{\sqrt{\rho\mu_0}}(\nabla\times\boldsymbol{u})\cdot\left((\nabla\times\boldsymbol{B})\times\boldsymbol{B}\right)-(\eta+\nu)(\nabla\times\boldsymbol{B})\cdot(\nabla\times\boldsymbol{u})\right]d^3\boldsymbol{x}. \quad (22)$$

Substituting in Eq. (6) produces terms up to cubic in the temporal POD modes,

$$0\approx\int\frac{\partial H_c}{\partial t}d^3\boldsymbol{x}=\frac{\partial}{\partial t}(a_i a_j)\int\frac{1}{\sqrt{\rho\mu_0}}\boldsymbol{\chi}_i^{B_u}\cdot\boldsymbol{\chi}_j^B d^3\boldsymbol{x} \quad (23)$$

$$=A_{ij}\frac{\partial}{\partial t}(a_i a_j)\rightarrow\begin{bmatrix}A_{ij}C_j^0 a_i\\A_{ij}L_{jk}^0 a_i a_k\\A_{ij}Q_{jkl}^0 a_i a_k a_l\end{bmatrix}\approx\begin{bmatrix}0\\0\\0\end{bmatrix}$$

Note that if the system is energy-preserving, $C_j^0=0$ for all j, so the first equality is already satisfied. The second equality determines that $A_{ij}L_{jk}^0$ is anti-symmetric under swapping i and k, and energy-preservation in Eq. (18) produces anti-symmetry under swapping j and k. The most straightforward solution is $L_{jk}^0=0$ for all j,k; this solution is precisely the ideal limit corresponding to $\eta=\nu=0$. Since A_{ij} is not symmetric, this constraint can also apply to systems which conserve cross-helicity despite finite dissipation.

Lastly, $A_{ij}Q_{jkl}^0$, containing only the contribution from the Hall-term, exhibits the same structure as (and is compatible with) our constraint on the energy-preserving nonlinearities in Eq. (19). The simplest solution is $A_{ij}Q_{jkl}^0=0$ for all i,k,l, since this corresponds to standard MHD without the Hall term. Like the analysis of the linear terms, this constraint indicates that it is possible that there are indices for which $A_{ij}Q_{jkl}^0\neq0$ but overall satisfy $A_{ij}Q_{jkl}^0 a_i a_k a_l=0$, so that nonzero Hall contributions can still conserve cross-helicity. Lastly, although inviscid Hall-MHD has two other time-invariants, enforcing the remaining invariants may require alternative formulations to the one presented here, since derived fields like the vector potential are involved.

C.   Conservation laws with velocity units

The previous sections have illustrated that our choice of magnetic field units in Eq. (1) allowed us to relate global MHD conservation laws to the structure of the coefficients in the POD-Galerkin model. It is worth exploring any alterations in velocity units (in closer analogy to fluid dynamics) $\boldsymbol{q}=[\boldsymbol{u},\boldsymbol{u}_A,u_s]$,

$$u_s^2=\frac{4T}{m_i(\gamma-1)},\quad\boldsymbol{u}_A=\frac{\boldsymbol{B}}{\sqrt{\mu_0\rho}}, \quad (24)$$

$$\frac{1}{2}\langle\boldsymbol{q},\boldsymbol{q}\rangle=\frac{1}{2}\int\left(u^2+u_A^2+u_s^2\right)d^3\boldsymbol{x}. \quad (25)$$

We have defined a scaled plasma sound speed, u_s. If ρ is uniform $\rho\langle\boldsymbol{q},\boldsymbol{q}\rangle/2=W$. The isothermal and time-independent density assumptions allow us to derive another quadratic model in $\boldsymbol{q}$, for which a POD-Galerkin model is readily available (the form is identical to Eq. (10) but the POD modes and coefficients have changed). Once again, assume $\boldsymbol{u}\cdot\hat{\boldsymbol{n}}=\boldsymbol{u}\times\hat{\boldsymbol{n}}=0$, $\boldsymbol{J}\cdot\hat{\boldsymbol{n}}=0$, and $\boldsymbol{B}\cdot\hat{\boldsymbol{n}}=0$ on the boundary, so that

$$\int\frac{\rho}{2}\frac{dq^2}{dt}d^3\boldsymbol{x}=\frac{\partial W}{\partial t}. \quad (26)$$

This is equivalent to Eq. (16) in the particular case of time-independent density. Without this assumption, an extra term appears, proportional to $\int\boldsymbol{u}\cdot\nabla(u^2+u_A^2)d^3\boldsymbol{x}$. Although from dimensional analysis this term is potentially very large, this may not be the case for many laboratory devices with strong anisotropy introduced by a large external magnetic field. For instance, steady-state toroidal plasmas with large closed flux surfaces would expect $\boldsymbol{u}\cdot\nabla u_A^2$ and $\boldsymbol{u}\cdot\nabla u^2$ to be small, as the fluid velocity is primarily along field lines and gradients in both the magnetic and velocity fields are primarily across field lines. For this reason, in certain devices the use of $\boldsymbol{q}=[\boldsymbol{u},\boldsymbol{u}_A,u_s]$ could be a useful alternative to the formulation used in the main body of this work. It is possible that, in these units, the structure of the nonlinearities in the associated POD-Galerkin model may prove more amenable to analysis or computation.

D.   Hyper-reduction techniques

Now that we have illustrated how global conservation laws manifest as structure in Galerkin models, we could compute the coefficients in Eq. (10) and evolve the subsequent model. However, in order to calculate the model coefficients, spatial derivatives for ρ, $\boldsymbol{B}_u$, and $\boldsymbol{B}$ (and B_T if temperature is evolved) must be well-approximated in the region of experimental interest. In some cases, high-resolution diagnostics can resolve these quantities in a particular plasma region. Even if the high-quality data is available, for instance through simulations, computing these inner products and evaluating the nonlinear terms is expensive, because the fields have the original spatial dimension D. This somewhat reduces the usefulness of projection-based model reduction. Fortunately, there are hyper-reduction techniques from fluid dynamics [41], such as the discrete empirical interpolation method (DEIM) [103], QDEIM [104], missing-

point estimation (MPE) [105] and gappy POD [106, 107], which can enable efficient computations. Instead of using hyper-reduction, we will turn to emerging and increasingly sophisticated machine learning methods in Section IV to discover Galerkin models from data. There are two primary reasons we have derived the POD-Galerkin model structure here anyways: 1) it indicates that we can search plasma datasets for systems of ODEs consisting only up to quadratic polynomials, 2) it provides a theoretical basis for projection-based model reduction and hyper-reduction techniques in future MHD work.

IV. CONSTRAINED IDENTIFICATION OF DATA-DRIVEN MODELS

In projection-based modeling, the expansion in the POD basis is data-driven, but the projection step is intrusive, requiring access to a numerical solver of the known governing equations. Purely data-driven techniques are useful because they are non-intrusive and in principle do not require high-resolution simulations or knowledge of the governing equations. From the projection-based analysis in Sections II−III, we are now able to define physical constraints for improved data-driven models.

This is an opportune time to discover data-driven models; throughout the scientific community, emerging techniques in system identification and optimization are increasingly facilitating the discovery of physical models directly from data [21, 56]. We use the sparse identification of nonlinear dynamics (SINDy) algorithm [57] to identify nonlinear reduced-order models for plasmas because SINDy models are *parsimonious*, having as few terms as are required to explain the dynamics. This feature of the SINDy algorithm promotes models that are interpretable and generalizable.

A. The constrained SINDy method

In our case, we compute a set of POD modes from a plasma dataset, and then use SINDy to search for low-dimensional models for $\boldsymbol{a}(t)$ as a sparse linear combination of elements from a library of candidate terms $\boldsymbol{\Theta}$,

$$\boldsymbol{\Theta}(\boldsymbol{a}) = \left[\begin{array}{cccc} | & | & | & \\ \mathbf{1} & \boldsymbol{a} & \boldsymbol{a} \otimes \boldsymbol{a} & \cdots \\ | & | & | & \end{array} \right]. \tag{27}$$

Here $\boldsymbol{a} \otimes \boldsymbol{a}$ is all combinations of $a_i a_j$ without duplicates, and similarly for the other candidate terms. We now assume that the evolution of $\boldsymbol{a}$ can be approximated as

$$\dot{\boldsymbol{a}} = \boldsymbol{f}(\boldsymbol{a}) \approx \boldsymbol{\Theta}(\boldsymbol{a})\boldsymbol{\Xi}. \tag{28}$$

The optimization problem solves for a sparse matrix of coefficients $\boldsymbol{\Xi}$, which represents the coefficients (strengths) of the candidate terms in $\boldsymbol{\Theta}$. To address

this combinatorially hard problem, it leverages sparse regression techniques, optimizing for the sparsest set of equations that produces an accurate fit of the data. To incorporate known physical laws, a constrained SINDy formulation was first introduced to conserve energy in incompressible fluids [34]. The constrained SINDy optimization problem can be written

$$\min_{\boldsymbol{\Xi}} ||\dot{\boldsymbol{a}} - \boldsymbol{\Theta}(\boldsymbol{a})\boldsymbol{\Xi}||_2^2 + \lambda R(\boldsymbol{\Xi}), \tag{29}$$
$$\text{subject to } \boldsymbol{D}\boldsymbol{\Xi}[:] = \boldsymbol{d},$$

where $R(\boldsymbol{\Xi})$ is a regularizer such as the L^0 or L^1 norm, which promotes sparsity in the coefficients $\boldsymbol{\Xi}$. $\boldsymbol{D}$ is a constraint matrix that allows us to impose that affine combinations of the coefficients in $\boldsymbol{\Xi}$ have the fixed values in $\boldsymbol{d}$. The original unconstrained SINDy algorithm solves Eq. (29) without using the constraint $\boldsymbol{D}\boldsymbol{\Xi}[:] = \boldsymbol{d}$. Here $\boldsymbol{a}, \dot{\boldsymbol{a}} \in \mathbb{R}^{M \times r}$, $\boldsymbol{\Theta}(\boldsymbol{a}) \in \mathbb{R}^{M \times N}$, $\boldsymbol{\Xi} \in \mathbb{R}^{N \times r}$, $\boldsymbol{D} \in \mathbb{R}^{N_c \times rN}$, $\boldsymbol{\Xi}[:] \in \mathbb{R}^{rN}$, $\boldsymbol{d} \in \mathbb{R}^{N_c}$, where N is the number of candidate terms, N_c is the number of constraints, and $\boldsymbol{\Xi}[:] = \left[\xi_1^{a_1} \cdots \xi_1^{a_r} \cdots \xi_N^{a_1} \cdots \xi_N^{a_r} \right]$ denotes the flattened or "vectorized" set of model coefficients. Motivated by the Galerkin model we have derived, we restrict the library of candidate terms $\boldsymbol{\Theta}$ to first and second order polynomials in $\boldsymbol{a}(t)$, although this is not a requirement of the SINDy algorithm; more complicated nonlinear terms may also be included for modeling the effect of truncated POD modes [34] or capturing POD-Galerkin models which exhibit higher order nonlinearities. $\boldsymbol{\Xi}$ is typically identified via sparse regression, for example by sequentially thresholded least-squares [57, 108], LASSO [109], or sparse regularized relaxed regression (SR3) [110]. In Appendix B, we explicitly derive the SINDy constraints required for the identified models to satisfy the global conservation laws discussed in Section III.

To summarize, we use a physics-informed sparse regression method that requires only $\dot{\boldsymbol{a}}$ to discover data-driven models for the evolution of $\boldsymbol{a}$. In the next section, we compute the POD for an example 3D MHD simulation and feed the chronos into the SINDy algorithm to identify data-driven models that we can use for forecasting future data.

B. Initial results

The theoretical structure of this reduced-order modeling framework is appealing, but its value to the community ultimately depends on the quality of the analysis when applied to plasma systems. Guided by the theory, we construct a nonlinear, physics-constrained SINDy model for an isothermal Hall-MHD simulation of this device, described in detail in Appendix A. The density, velocity, and magnetic field are sampled at a set of equally-spaced 3D points in the volume and sampling intervals $\Delta\phi = \pi/16$, $\Delta R \approx \Delta Z \approx 2$ cm. The result is that each component of $\boldsymbol{u}$ and $\boldsymbol{B}$ has 47712 samples. This high-resolution is ideal for visualization but substantial

(a)

(b)
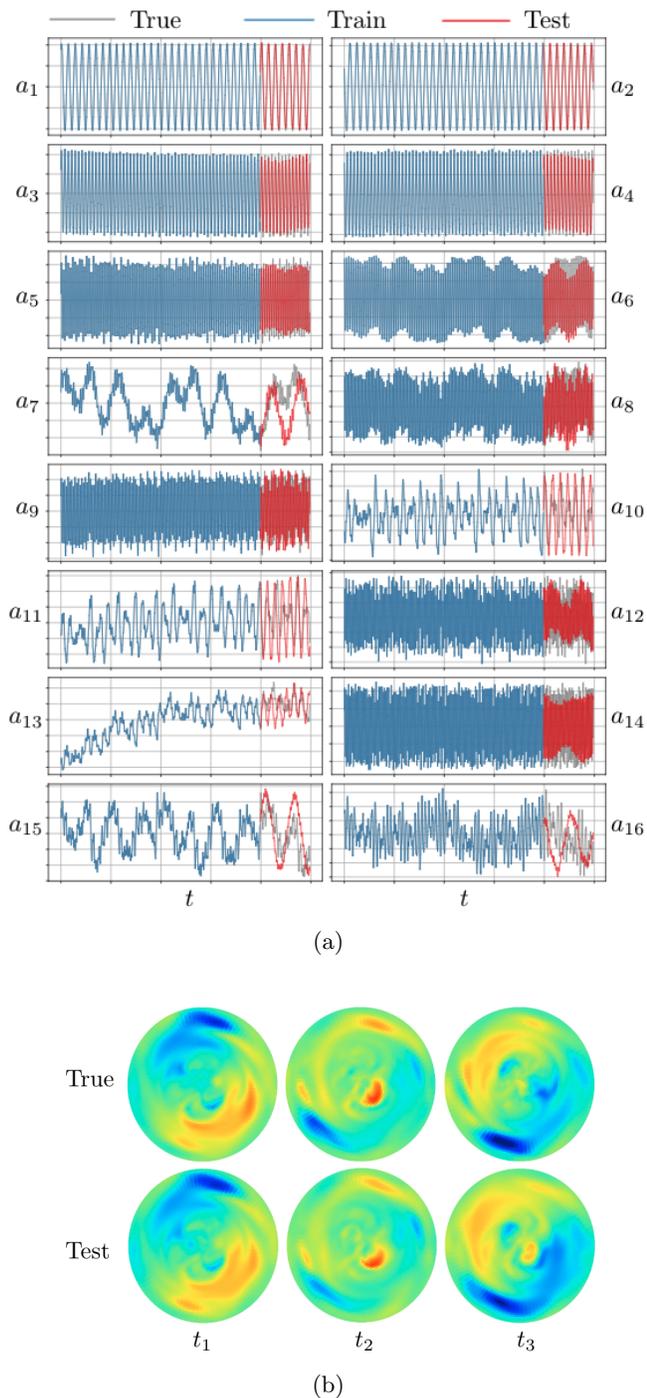
FIG. 3: Summary of the constrained SINDy performance on a 3D Hall-MHD simulation of the HIT-SI device described in Appendix A. (a) Constrained SINDy prediction of $a_1,...,a_{16}$. The true evolution is in gray, the training data used for the model-finding is in blue, and the SINDy prediction is in red (see online version for color). (b) Constrained SINDy predictions of u_z (Test) in the $Z=0$ midplane are compared with the true u_z evolution at three snapshots in time, indicating strong algorithm performance.

size reduction can be done with little or no change to the spatial or temporal POD modes. For instance, in Figures 2 and 3a, the $Z=0$ visualizations of the 3D spatial POD modes are constructed from the 1440 sample locations at $Z=0$; with a non-uniform set of 50 points in the midplane, the only change to the visualization is a smoothing out of the shortest wavelengths. The temporal resolution of the measurements is $\Delta t_k=1$ μs. The analysis is essentially unchanged for time steps as large as 10 μs, but smaller time steps are required in HIT-SI to resolve harmonics of the injector frequency that appear in the temporal POD modes. For instance, at $\Delta t_k=10$ μs, the fourth injector harmonic is sampled, on average, less than twice per period.

From these measurements of the density, velocity, and magnetic field, we compute the topos and chronos via the SVD in Eq. (4), obtaining a Galerkin expansion for the velocity and magnetic fields in magnetic field units, as in Eq. (6). Now a constrained SINDy model is identified for the first 16 chronos $a_i(t)$ and the forecasting is illustrated in Fig. 3a. The SINDy model accurately captures most of the $a_j(t)$ dynamics, with larger errors for the less energetic modes. Some of the low-frequency content in the $a_j(t)$ is not captured by this particular constrained SINDy model, but this is largely because the low frequencies are not well-resolved in the time range used for training. Despite this deficiency in the data, the SINDy model illustrates strong prediction performance in the $Z=0$ midplane reconstructions of the simulation data in Fig. 3b and forecasts much of the time evolution for a high-dimensional simulation that used 589,824 grid points, a tremendous efficiency gain of $\mathcal{O}(10^5)$. Furthermore, this model was obtained by training on a dataset representing a single discharge. Further improvements are likely accessible by training on a dataset of many discharges of varying trajectories.

We have found a quality forecasting model from the SINDy system identification method, but it is interesting to see how the model quality varies with the algorithm hyperparameters like the model sparsity λ and model rank r. In Figure 4, we illustrate how the normalized reconstruction errors of $\boldsymbol{X}$ and $\dot{\boldsymbol{X}}$ vary in the "Pareto-space" of (r,λ) for both the unconstrained and constrained SINDy algorithms, with the goal to explore the space of possible models obtained from this system identification technique. Although the exact reconstruction error values are unique to the simulation examined here, there are some interesting qualitative features that we expect to be quite general. The unconstrained SINDy algorithm indicates a significant region of (r,λ) where numerically unstable models are found. For $r\gtrsim10$, the models are typically either unstable or too sparse to be effective for forecasting. In contrast, by construction the constrained SINDy algorithm conserves the energy and therefore exhibits no unstable models. This is promising for discovering models on historically challenging systems for machine learning methods − multi-scale or turbulent systems that require $r\gg1$ to properly capture the

dynamics. At first glance, it may appear that the constrained SINDy errors in $\dot{\mathbf{X}}$ are worse than the unconstrained errors, but the low-error values in the unconstrained case are precisely the unstable models. These nonsparse models are overfitting, leading to instability in the numerical integration. Finally, we can see that at $\lambda \approx 0.091$, all the SINDy models are rendered ineffective. This value is precisely at the driving frequency of the HIT-SI injectors in this simulation; if λ is larger than this frequency, SINDy thresholds off the primary dynamics in the system.

V. CONCLUSIONS

A hierarchy of models with varying fidelity is essential for understanding and controlling plasmas, and our work provides a principled lower level on this hierarchy − low-dimensional and interpretable plasma models which can be used for physical discovery, forecasting, stability analysis, and real-time control. We have discussed how these models are obtained from either projection-based or data-driven methods. Furthermore, we illustrated how Galerkin plasma models retain the global conservation laws of MHD, and machine learning or system identification methods like SINDy can use these constraints directly in an optimization procedure for discovering such models from data. We demonstrated the effectiveness of this approach for a 3D isothermal Hall-MHD simulation of a self-organized plasma. This framework may be used more broadly for discovering low-dimensional models, forecasting, or real-time control of complex plasmas. This principled enforcement of global conservation laws is critical for the stability and success of future low-dimensional plasma models.

There are a number of potential numerical limitations to the methodology presented here, including stability issues, the curse of dimensionality, addressing turbulent or stochastic systems, and extrapolation beyond the training dataset. Fortunately, all of these potential caveats are currently the subjects of intense ongoing research efforts. Generally, the systems of nonlinear ODEs identified by the unconstrained SINDy algorithm tend to have depreciating stability properties as the number of modes increases; unless steps to constrain the model structure are taken, as in the present work and other recent work in provably stable data-driven models [101, 118, 119], this may prove a difficult obstacle. Fortunately, there are also several alternative formulations of the SINDy algorithm that may be more robust in some circumstances and are worth exploring on plasma systems [111–117, 120].

In regards to the curse of dimensionality, the SINDy library grows combinatorially with the number of state variables and the optimization problem can become very ill-conditioned and computationally intensive. Even for the relatively modest candidate library used in the present work, limited to quadratic polynomials, the size scales as $\mathcal{O}(r^3)$ because Q_{ijk}^0 is a three index tensor. How-

ever, there is recent work utilizing low-rank tensor decompositions to significantly reduce memory usage and computational latency [111]. There has also been considerable recent progress in the modeling of turbulent systems that exhibit broadband turbulence, which generally require a prohibitive number of modes to faithfully reconstruct the field. New approaches bypass this requirement by using stochastic techniques [112, 116, 120] or finding new data-driven closures for the Navier-Stokes equations [117, 121]. Finally, extrapolation beyond the training set is a central challenge for all machine learning techniques and this issue is primarily addressed in system identification methods by additional steps to mitigate overfitting, such as the additional of a sparsity-promoting regularizer in the optimization problem.

Lastly, to promote reproducible research, the python code used for this analysis can be found at `https://github.com/akaptano/POD-Galerkin_MHD`. The results presented below have also been incorporated into an advanced example of the PySINDy software package [108].

A   The HIT-SI experiment and simulations

HIT-SI was a laboratory plasma device that formed and sustained spheromak plasmas for the study of plasma self-organization and steady inductive helicity injection (SIHI) [80]. It consisted of an axisymmetric flux conserver and two inductive injectors (actuators) mounted on each end as illustrated in the top left panel of Fig. 1. Magnetic coils on each injector, generating helical fields linking through the flux conserver, were oscillated in phase at a frequency with values between $10-70$ kHz. The magnetic fields generated by the two injectors were spatially and temporally $90°$ out of phase, resulting in approximately constant power and helicity injection. The fields from these injectors provided the power and magnetic helicity to both form and sustain a spheromak during experimental discharges, with a quasi-steady-state period of roughly constant spheromak amplitude lasting <1 ms. Additional details of the experiment and its operation can be found in references [80, 122, 123].
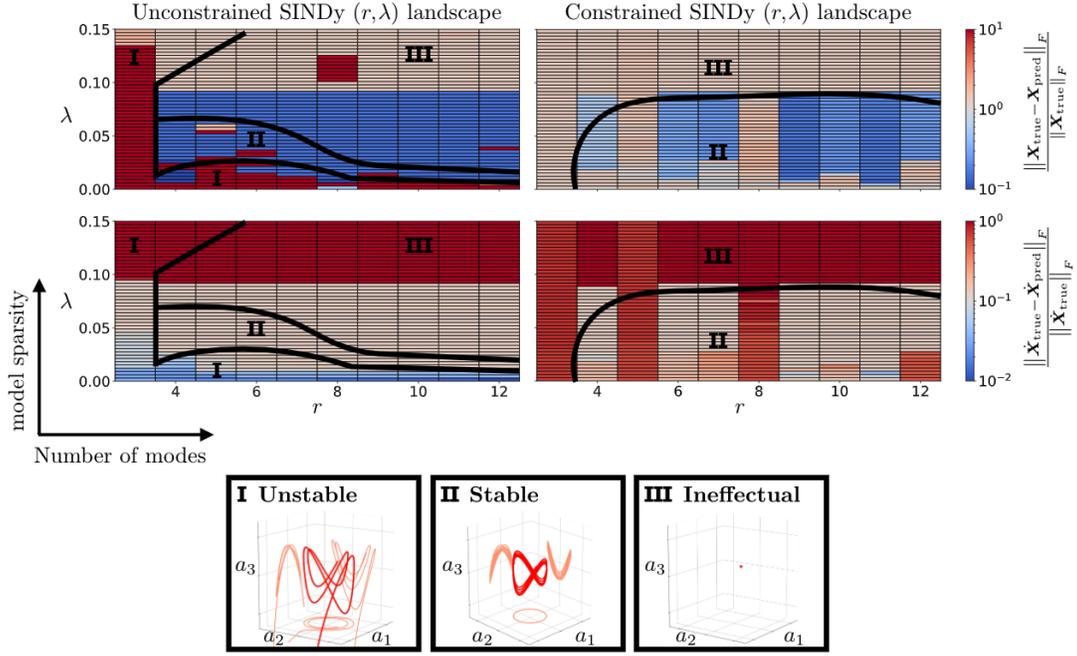
FIG. 4: Summary of the (r,λ) space of unconstrained and constrained SINDy models from the HIT-SI simulation. The unconstrained models approximately separate into three distinct classes. Class I illustrates nonsparse and typically unstable models. Class II consists of sparse and accurate solutions. Class III denotes solutions which are too sparse to accurately capture the dynamics. Computed errors are for the testing part of the dataset; the colorbar (see online version for color) is limited to 10^1 as unstable model errors grow arbitrarily large. Constrained SINDy guarantees the energy norm is preserved and thus class I vanishes. Algorithmic advances [101, 111–117] may help further expand the size of class II.

Simulations of HIT-SI were performed using the Hall-MHD equations and solved by the NIMROD code [124]. NIMROD discretizes equations in cylindrical coordinates (R,Z,ϕ); the R-Z plane is composed of finite elements and the ϕ component is expanded in a finite Fourier series. Mesh convergence was obtained previously by Akcay [125] on a grid with 28×28 finite elements of polynomial degree 4 and 22 Fourier components, so we use the same grid for the simulation here. Due to the 2D grid, the HIT-SI injectors cannot be directly modeled in the simulation. Rather, the injectors are implemented as $B_\perp$ and $E_\parallel$ boundary conditions at the top and bottom device surfaces to match the experimental waveforms. A detailed description of the implementation of these boundary conditions can be found in Akcay [126]. Dirichlet boundary conditions are used for all other variables; the plasma density satisfies $n_e = 2 \times 10^{19}$ m^{-3} and the temperatures satisfy $T_i = T_e = 14$ eV. Isotropic viscosity $\nu = 550$ m^2/s and Spitzer resistivity [96] is used. The remaining boundary conditions are $\boldsymbol{u} \times \hat{\boldsymbol{n}} = \boldsymbol{u} \cdot \hat{\boldsymbol{n}} = 0$, $\boldsymbol{J} \cdot \hat{\boldsymbol{n}} = 0$, and $\boldsymbol{B} \cdot \hat{\boldsymbol{n}} = 0$. For more information on the numerical model used in this simulation, see Morgan et al. [127]. The data for training and testing are obtained during the approximately steady-state phase of the simulation so that the energy constraints derived in Section III A are applicable.

B   Derivation of the SINDy constraints

In Sec. III, we derived constraints for the POD-Galerkin model coefficients from global conservation laws; our goal here is to rewrite these constraints to be compatible with the formulation of the SINDy system identification method. The conclusions for the global conservation of energy were: 1) no constant terms, 2) an anti-symmetry constraint on the linear part of the coefficient matrix $\boldsymbol{\Xi}$, and 3) a more complicated energy-preserving structure in the quadratic coefficients. Consider a quadratic library in a set of r modes, ordered as $\boldsymbol{\Theta}(\boldsymbol{a}) = [a_1,...,a_r,a_1 a_2,...,a_{r-1}a_r,a_1^2,...,a_r^2]$. Note that this arrangement of the polynomials in $\boldsymbol{\Theta}$ differs from Loiseau et al. [74], so the indexing and subscripts are also different here. First we will consider the constraint on the linear part of the Galerkin model in Eq. (10), $\boldsymbol{a}^T \boldsymbol{L}_0 \boldsymbol{a} \approx 0$. We can rewrite this in the SINDy notation as

$$0 = \begin{bmatrix} a_1 & \cdots & a_r \end{bmatrix} \begin{bmatrix} \xi_1^{a_1} & \cdots & \xi_r^{a_1} \\ \vdots & \ddots & \vdots \\ \xi_r^{a_r} & \cdots & \xi_r^{a_r} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_r \end{bmatrix}. \qquad (30)$$

We conclude $\xi_i^{a_j} = -\xi_j^{a_i}$ for $i,j \in \{1,...,r\}$ and identify $\xi_i^{a_j}$ by accessing the $(i-1)r+j$ index in the vector of model coefficients $\boldsymbol{\Xi}[:]$. Note we are only accessing the first r^2

elements of $\boldsymbol{\Xi}[:]$. For models of linear and quadratic polynomials, $N=(r^2+3r)/2$ and the number of constraints from anti-symmetry of the linear coefficients is $N_L=(r^2+r)/2$. Thus there are now only $rN-N_L=r(r^2+2r-1)/2$ free parameters. Since the constrained SINDy algorithm solves linear equality constraints of the form $\boldsymbol{D\Xi}[:]=\boldsymbol{d}$, we can write this out explicitly for $r=3$,

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & \cdots
\end{bmatrix}
\begin{bmatrix}
\xi_1^{a_1} \\
\xi_1^{a_2} \\
\xi_1^{a_3} \\
\xi_1^{a_1} \\
\xi_2^{a_1} \\
\vdots
\end{bmatrix}
=
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{bmatrix}. \quad (31)
$$

The boundary conditions $\boldsymbol{u}\cdot\hat{\boldsymbol{n}}=0$, $\boldsymbol{J}\cdot\hat{\boldsymbol{n}}=0$, $\boldsymbol{B}\cdot\hat{\boldsymbol{n}}=0$ guaranteed that the quadratic nonlinearities were energy-preserving, and thus that cubic terms in Eq. (12) vanish,

$$
\sum_{i,j,k=0}^{r} Q_{ijk}^{0} a_i a_j a_k \approx 0. \quad (32)
$$

This constraint is significantly more involved to reformat. Written in SINDy notation, this is equivalent to

$$
0=\begin{bmatrix} a_1 & \cdots & a_r \end{bmatrix}
\begin{bmatrix}
\xi_{r+1}^{a_1} & \xi_{r+2}^{a_1} & \cdots & \xi_N^{a_1} \\
\vdots & \vdots & \vdots & \vdots \\
\xi_{r+1}^{a_r} & \xi_{r+2}^{a_r} & \cdots & \xi_N^{a_r}
\end{bmatrix}
\begin{bmatrix}
a_1 a_2 \\
\vdots \\
a_{r-1} a_r \\
a_1^2 \\
\vdots \\
a_r^2
\end{bmatrix}. \quad (33)
$$

Expand this all out and group the like terms, i.e. terms which look like a_i^3, $a_i a_j^2$ or $a_i a_j a_k$, $i,j,k\in\{1,...,r\}$, $i\neq j\neq k$. All of the like terms can be straightforwardly

shown to be linearly independent, so we can consider three constraints separately for the three types of terms. The number of each of these respective terms is $\binom{r}{1}=r$, $2\binom{r}{2}=r(r-1)$, and $\binom{r}{3}=r(r-1)(r-2)/6$, for a total of $r(r+1)(r+2)/6=N_Q$ constraints. With both constraints, we have $rN-N_L-N_Q=r(r-1)(2r+5)/6$ free parameters, and $N_c=N_L+N_Q$ constraints. Further considering the quadratic case, we find that coefficients which adorn a_i^3 must vanish, $\xi_{N-r+i}^{a_i}=0$. Now define

$$
\tilde{\xi}_{ijk}=\xi_{r+\frac{j}{2}(2r-j-3)+k-1}^{a_i}. \quad (34)
$$

The second type of constraint, with $i\neq j$, produces

$$
\xi_{N-r+j}^{a_i}=\begin{cases} \tilde{\xi}_{jij} & i<j \\ \tilde{\xi}_{jji} & i>j, \end{cases} \quad (35)
$$

while the third type of constraint produces

$$
\tilde{\xi}_{ijk}+\tilde{\xi}_{jik}+\tilde{\xi}_{kij}=0. \quad (36)
$$

This relation is equivalent to the energy-preserving conditions in Schlegel et al. [79], but the indexing is not straightforward, even after fully expanding Eq. (33). This equation is an arbitrary r generalization to the $r=3$ constraint used in Loiseau et al. [34]. For the specific case where the plasma system is Hamiltonian (for instance in ideal [128], Hall [129], and extended [130] MHD without dissipation) and the measurements are assumed to be sufficient to represent the Hamiltonian, one could alternatively use formulations of SINDy to directly discover the Hamiltonian [131] and subsequently derive the equations of motion. Lastly, if the global energy conservation constraint on the quadratic terms in the SINDy coefficient matrix $\boldsymbol{\Xi}$ is written $D_{jk}\Xi_k=0$, then the quadratic cross-helicity constraint can be written $D_{jk}A_{kl}\Xi_l=0$.

[1] J. R. Roth, *Industrial plasma engineering: Volume 2: Applications to nonthermal plasma processing*, Vol. 2 (CRC press, 2001).

[2] J. Candy and R. E. Waltz, Anomalous transport scaling in the DIII-D tokamak matched by supercomputer simulation, Phys. Rev. Lett. **91**, 045001 (2003).

[3] O. Ohia, J. Egedal, V. S. Lukin, W. Daughton, and A. Le, Demonstration of anisotropic fluid closure capturing the kinetic structure of magnetic reconnection, Phys. Rev. Lett. **109**, 115004 (2012).

[4] D. Grošelj, A. Mallet, N. F. Loureiro, and F. Jenko, Fully kinetic simulation of 3D kinetic Alfvén turbulence, Phys. Rev. Lett. **120**, 105101 (2018).

[5] K. Taira, S. L. Brunton, S. Dawson, C. W. Rowley, T. Colonius, B. J. McKeon, O. T. Schmidt, S. Gordeyev, V. Theofilis, and L. S. Ukeiley, Modal analysis of fluid flows: An overview, AIAA Journal **55**, 4013 (2017).

[6] R. Jiménez-Gómez, E. Ascasíbar, T. Estrada, I. García-Cortés, B. Van Milligen, A. López-Fraguas, I. Pastor, and D. López-Bruna, Analysis of magnetohydro-dynamic instabilities in TJ-II plasmas, Fusion science and technology **51**, 20 (2007).

[7] B. P. van Milligen, E. Sánchez, A. Alonso, M. A. Pedrosa, C. Hidalgo, A. M. de Aguilera, and A. L. Fraguas, The use of the biorthogonal decomposition for the identification of zonal flows at TJ-II, Plasma Physics and Controlled Fusion **57**, 025005 (2014).

[8] M. Pandya, *Low edge safety factor disruptions in the Compact Toroidal Hybrid: Operation in the low-q regime, passive disruption avoidance and the nature of MHD precursors*, Ph.D. thesis, Auburn University (2016).

[9] B. Victor, C. Akcay, C. Hansen, T. Jarboe, B. Nelson, and K. Morgan, Development of validation metrics using biorthogonal decomposition for the comparison of magnetic field measurements, Plasma Physics and Controlled Fusion **57**, 045010 (2015).

[10] E. Strait, J. King, J. Hanson, and N. Logan, Spatial and temporal analysis of DIII-D 3D magnetic diagnostic data, Review of Scientific Instruments **87**, 11D423

(2016).

[11] P. J. Byrne, *Study of External Kink Modes in Shaped HBT-EP Plasmas*, Ph.D. thesis, Columbia University (2017).

[12] S. Gu, B. Wan, Y. Sun, N. Chu, Y. Liu, T. Shi, H. Wang, M. Jia, and K. He, A new criterion for controlling edge localized modes based on a multi-mode plasma response, Nuclear Fusion **59**, 126042 (2019).

[13] A. A. Kaptanoglu, K. D. Morgan, C. J. Hansen, and S. L. Brunton, Characterizing magnetized plasmas with dynamic mode decomposition, Physics of Plasmas **27**, 032108 (2020).

[14] P. Benner, V. Mehrmann, and D. C. Sorensen, *Dimension reduction of large-scale systems*, Vol. 45 (Springer, 2005).

[15] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, *Model reduction and approximation: theory and algorithms* (SIAM, 2017).

[16] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, Machine learning for fluid mechanics, Annual Review of Fluid Mechanics **52**, 477 (2020).

[17] B. R. Noack, K. Afanasiev, M. Morzynski, G. Tadmor, and F. Thiele, A hierarchy of low-dimensional models for the transient and post-transient cylinder wake, Journal of Fluid Mechanics **497**, 335 (2003).

[18] B. R. Noack, W. Stankiewicz, M. Morzynski, and P. J. Schmid, Recursive dynamic mode decomposition of a transient cylinder wake, Journal of Fluid Mechanics **809**, 843 (2016).

[19] K. Carlberg, M. Barone, and H. Antil, Galerkin v. least-squares Petrov–Galerkin projection in nonlinear model reduction, Journal of Computational Physics **330**, 693 (2017).

[20] C. W. Rowley and S. T. Dawson, Model reduction for flow analysis and control, Annual Review of Fluid Mechanics **49**, 387 (2017).

[21] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control* (Cambridge University Press, 2019).

[22] E. N. Lorenz, Deterministic nonperiodic flow, Journal of the atmospheric sciences **20**, 130 (1963).

[23] Y. Guan, S. L. Brunton, and I. Novosselov, Sparse nonlinear models of chaotic electroconvection, arXiv preprint arXiv:2009.11862 (2020).

[24] K. D. Humbird, J. L. Peterson, B. Spears, and R. McClarren, Transfer learning to model inertial confinement fusion experiments, IEEE Transactions on Plasma Science **48**, 61 (2019).

[25] M. G. Kapteyn, D. J. Knezevic, and K. Willcox, Toward predictive digital twins via component-based reduced-order models and interpretable machine learning, in *AIAA Scitech 2020 Forum* (2020) p. 0418.

[26] S. Ravindran, Real-time computational algorithm for optimal control of an MHD flow system, SIAM Journal on Scientific Computing **26**, 1369 (2005).

[27] C. Galperti, S. Coda, B. Duval, X. Llobet, P. Milne, O. Sauter, J. Moret, and D. Testa, Integration of a real-time node for magnetic perturbations signal analysis in the distributed digital control system of the TCV tokamak, IEEE Transactions on Nuclear Science **64**, 1446 (2017).

[28] J. Levesque, N. Rath, D. Shiraki, S. Angelini, J. Bialek, P. Byrne, B. DeBono, P. Hughes, M. Mauel, G. Navratil, *et al.*, Multimode observations and 3D magnetic control of the boundary of a tokamak plasma, Nuclear Fusion **53**, 073037 (2013).

[29] L. Wang, X. Xu, B. Zhu, C. Ma, and Y.-a. Lei, Deep learning surrogate model for kinetic Landau-fluid closure with collision, AIP Advances **10**, 075108 (2020).

[30] J. Citrin, S. Breton, F. Felici, F. Imbeaux, T. Aniel, J. Artaud, B. Baiocchi, C. Bourdelle, Y. Camenen, and J. Garcia, Real-time capable first principle based modelling of tokamak turbulent transport, Nuclear Fusion **55**, 092001 (2015).

[31] K. L. van de Plassche, J. Citrin, C. Bourdelle, Y. Camenen, F. J. Casson, V. I. Dagnelie, F. Felici, A. Ho, S. Van Mulders, and J. Contributors, Fast modeling of turbulent transport in fusion plasmas using neural networks, Physics of Plasmas **27**, 022310 (2020).

[32] A. Loarte, B. Lipschultz, A. Kukushkin, G. Matthews, P. Stangeby, N. Asakura, G. Counsell, G. Federici, A. Kallenbach, K. Krieger, *et al.*, Power and particle control, Nuclear Fusion **47**, S203 (2007).

[33] F. Allgöwer, T. A. Badgwell, J. S. Qin, J. B. Rawlings, and S. J. Wright, Nonlinear predictive control and moving horizon estimationan introductory overview, in *Advances in control* (Springer, 1999) pp. 391–449.

[34] J.-C. Loiseau and S. L. Brunton, Constrained sparse Galerkin regression, Journal of Fluid Mechanics **838**, 42 (2018).

[35] J.-C. Loiseau, Data-driven modeling of the chaotic thermal convection in an annular thermosyphon, Theoretical and Computational Fluid Dynamics **34**, 339 (2020).

[36] S. Kobayashi, Ö. D. Gürcan, and P. H. Diamond, Direct identification of predator-prey dynamics in gyrokinetic simulations, Physics of Plasmas **22**, 090702 (2015).

[37] E. P. Alves and F. Fiuza, Data-driven discovery of reduced plasma physics models from fully-kinetic simulations, arXiv preprint arXiv:2011.01927 (2020).

[38] M. Dam, M. Brøns, J. Juul Rasmussen, V. Naulin, and J. S. Hesthaven, Sparse identification of a predator-prey system from simulation data of a convection model, Physics of Plasmas **24**, 022310 (2017).

[39] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley, *Turbulence, coherent structures, dynamical systems and symmetry* (Cambridge university press, 2012).

[40] K. Willcox and J. Peraire, Balanced model reduction via the proper orthogonal decomposition, AIAA journal **40**, 2323 (2002).

[41] P. Benner, S. Gugercin, and K. Willcox, A survey of projection-based model reduction methods for parametric dynamical systems, SIAM review **57**, 483 (2015).

[42] C. W. Rowley, Model reduction for fluids, using balanced proper orthogonal decomposition, International Journal of Bifurcation and Chaos **15**, 997 (2005).

[43] A. Towne, O. T. Schmidt, and T. Colonius, Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis, Journal of Fluid Mechanics **847**, 821 (2018).

[44] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, Journal of Fluid Mechanics **656**, 5 (2010).

[45] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. Henningson, Spectral analysis of nonlinear flows, J. Fluid Mech. **645**, 115 (2009).

[46] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, On dynamic mode decomposition:

theory and applications, Journal of Computational Dynamics **1**, 391 (2014).

[47] B. O. Koopman, Hamiltonian systems and transformation in Hilbert space, Proceedings of the National Academy of Sciences **17**, 315 (1931).

[48] I. Mezic, Analysis of fluid flows via spectral properties of the Koopman operator, Annual Review of Fluid Mechanics **45**, 357 (2013).

[49] S. L. Brunton, M. Budišić, E. Kaiser, and J. N. Kutz, Modern Koopman theory for dynamical systems, arXiv preprint arXiv:2102.12086 (2021).

[50] B. McKeon and A. Sharma, A critical-layer framework for turbulent pipe flow, Journal of Fluid Mechanics **658**, 336 (2010).

[51] M. Luhar, A. S. Sharma, and B. J. McKeon, Opposition control within the resolvent analysis framework, Journal of Fluid Mechanics **749**, 597 (2014).

[52] B. Lusch, J. N. Kutz, and S. L. Brunton, Deep learning for universal linear embeddings of nonlinear dynamics, Nature communications **9**, 4950 (2018).

[53] K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, Data-driven discovery of coordinates and governing equations, Proceedings of the National Academy of Sciences **116**, 22445 (2019).

[54] K. Lee and K. T. Carlberg, Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders, Journal of Computational Physics **404**, 108973 (2020).

[55] J. Bongard and H. Lipson, Automated reverse engineering of nonlinear dynamical systems, Proceedings of the National Academy of Sciences **104**, 9943 (2007).

[56] M. Schmidt and H. Lipson, Distilling free-form natural laws from experimental data, Science **324**, 81 (2009).

[57] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proceedings of the National Academy of Sciences **113**, 3932 (2016).

[58] M. Raissi and G. E. Karniadakis, Hidden physics models: Machine learning of nonlinear partial differential equations, arXiv preprint arXiv:1708.00588 (2017).

[59] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations, arXiv preprint arXiv:1711.10561 (2017).

[60] O. Yair, R. Talmon, R. R. Coifman, and I. G. Kevrekidis, Reconstruction of normal forms by learning informed observation geometries from data, Proceedings of the National Academy of Sciences , 201620045 (2017).

[61] S. Klus, F. Nüske, P. Koltai, H. Wu, I. Kevrekidis, C. Schütte, and F. Noé, Data-driven model reduction and transfer operator approximation, Journal of Nonlinear Science (2018).

[62] C. Wehmeyer and F. Noé, Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics, The Journal of Chemical Physics **148**, 1 (2018).

[63] A. Mardt, L. Pasquali, H. Wu, and F. Noé, VAMPnets: Deep learning of molecular kinetics, Nature Communications **9** (2018).

[64] K. Duraisamy, G. Iaccarino, and H. Xiao, Turbulence modeling in the age of data, Annual Reviews of Fluid Mechanics **51**, 357 (2019).

[65] J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach, Physical review letters **120**, 024102 (2018).

[66] F. Noé, S. Olsson, J. Köhler, and H. Wu, Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning, Science **365**, eaaw1147 (2019).

[67] Y. Bar-Sinai, S. Hoyer, J. Hickey, and M. P. Brenner, Learning data-driven discretizations for partial differential equations, Proceedings of the National Academy of Sciences **116**, 15344 (2019).

[68] D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer, Machine learning–accelerated computational fluid dynamics, Proceedings of the National Academy of Sciences **118** (2021).

[69] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics informed deep learning (part ii): Data-driven discovery of nonlinear partial differential equations, arXiv preprint arXiv:1711.10566 (2017).

[70] P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. Zambaldi, M. Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, *et al.*, Relational inductive biases, deep learning, and graph networks, arXiv preprint arXiv:1806.01261 (2018).

[71] M. D. Cranmer, R. Xu, P. Battaglia, and S. Ho, Learning symbolic physics with graph networks, arXiv preprint arXiv:1909.05862 (2019).

[72] M. Mohebujjaman, L. G. Rebholz, and T. Iliescu, Physically constrained data-driven correction for reduced-order modeling of fluid flows, International Journal for Numerical Methods in Fluids **89**, 103 (2019).

[73] M. Cranmer, S. Greydanus, S. Hoyer, P. Battaglia, D. Spergel, and S. Ho, Lagrangian neural networks, arXiv preprint arXiv:2003.04630 (2020).

[74] J.-C. Loiseau, B. R. Noack, and S. L. Brunton, Sparse reduced-order modeling: sensor-based dynamics to full-state estimation, Journal of Fluid Mechanics **844**, 459 (2018).

[75] A. Pouquet, D. Rosenberg, J. E. Stawarz, and R. Marino, Helicity dynamics, inverse, and bidirectional cascades in fluid and magnetohydrodynamic turbulence: a brief review, Earth and Space Science **6**, 351 (2019).

[76] C. W. Rowley, T. Colonius, and R. M. Murray, Model reduction for compressible flows using POD and Galerkin projection, Physica D **189**, 115 (2004).

[77] B. R. Noack, M. Schlegel, M. Morzynski, and G. Tadmor, *Galerkin method for nonlinear dynamics* (Springer, 2011).

[78] M. J. Balajewicz, E. H. Dowell, and B. R. Noack, Low-dimensional modelling of high-Reynolds-number shear flows incorporating constraints from the Navier–Stokes equation, Journal of Fluid Mechanics **729**, 285 (2013).

[79] M. Schlegel and B. R. Noack, On long-term boundedness of Galerkin models, Journal of Fluid Mechanics **765**, 325 (2015).

[80] T. Jarboe, W. Hamp, G. Marklin, B. Nelson, R. ONeill, A. Redd, P. Sieck, R. Smith, and J. Wrobel, Spheromak formation by steady inductive helicity injection, Physical review letters **97**, 115003 (2006).

[81] D. D. Schnack, D. C. Barnes, D. P. Brennan, C. C. Hegna, E. Held, C. C. Kim, S. E. Kruger, A. Y. Pankin, and C. R. Sovinec, Computational modeling of fully ionized magnetized plasmas using the fluid approximation, Physics of Plasmas **13**, 058103 (2006).

[82] Z. Ma and A. Bhattacharjee, Hall magnetohydrodynamic reconnection: The geospace environment model-

ing challenge, Journal of Geophysical Research: Space Physics **106**, 3773 (2001).

[83] V. Krishan and S. Mahajan, Magnetic fluctuations and Hall magnetohydrodynamic turbulence in the solar wind, Journal of Geophysical Research: Space Physics **109** (2004).

[84] F. Ebrahimi, B. Lefebvre, C. B. Forest, and A. Bhattacharjee, Global Hall-MHD simulations of magnetorotational instability in a plasma Couette flow experiment, Physics of Plasmas **18**, 062904 (2011).

[85] N. M. Ferraro, Calculations of two-fluid linear response to non-axisymmetric fields in tokamaks, Physics of Plasmas **19**, 056105 (2012).

[86] A. A. Kaptanoglu, T. E. Benedett, K. D. Morgan, C. J. Hansen, and T. R. Jarboe, Two-temperature effects in Hall-MHD simulations of the HIT-SI experiment, Physics of Plasmas **27**, 072505 (2020).

[87] T. Dudok de Wit, A.-L. Pecquet, J.-C. Vallet, and R. Lima, The biorthogonal decomposition as a tool for investigating fluctuations in plasmas, Physics of Plasmas **1**, 3288 (1994).

[88] C. Galperti, C. Marchetto, E. Alessi, D. Minelli, M. Mosconi, F. Belli, L. Boncagni, A. Botrugno, P. Buratti, B. Esposito, *et al.*, Development of real-time MHD markers based on biorthogonal decomposition of signals from Mirnov coils, Plasma Physics and Controlled Fusion **56**, 114012 (2014).

[89] B. P. Van Milligen, E. Sánchez, A. Alonso, M. Pedrosa, C. Hidalgo, A. M. De Aguilera, and A. L. Fraguas, The use of the biorthogonal decomposition for the identification of zonal flows at TJ-II, Plasma Physics and Controlled Fusion **57**, 025005 (2014).

[90] C. Hansen, B. Victor, K. Morgan, T. Jarboe, A. Hossack, G. Marklin, B. Nelson, and D. Sutherland, Numerical studies and metric development for validation of magnetohydrodynamic models on the HIT-SI experiment, Physics of Plasmas **22**, 056105 (2015).

[91] G. H. Golub *et al.*, Matrix computations, The Johns Hopkins (1996).

[92] A. Frieze, R. Kannan, and S. Vempala, Fast Monte-Carlo algorithms for finding low-rank approximations, Journal of the ACM (JACM) **51**, 1025 (2004).

[93] E. Liberty, F. Woolfe, P.-G. Martinsson, V. Rokhlin, and M. Tygert, Randomized algorithms for the low-rank approximation of matrices, Proceedings of the National Academy of Sciences **104**, 20167 (2007).

[94] F. Woolfe, E. Liberty, V. Rokhlin, and M. Tygert, A fast randomized algorithm for the approximation of matrices, Applied and Computational Harmonic Analysis **25**, 335 (2008).

[95] D. Rempfer and H. F. Fasel, Dynamics of three-dimensional coherent structures in a flat-plate boundary layer, Journal of Fluid Mechanics **275**, 257 (1994).

[96] L. Spitzer, *Physics of fully ionized gases* (Courier Corporation, 2006).

[97] S. Braginskii and M. Leontovich, Reviews of plasma physics (1965).

[98] F. Plunian, R. Stepanov, and P. Frick, Shell models of magnetohydrodynamic turbulence, Physics Reports **523**, 1 (2013).

[99] D. Biskamp, Cascade models for magnetohydrodynamic turbulence, Physical Review E **50**, 2702 (1994).

[100] M. Couplet, P. Sagaut, and C. Basdevant, Intermodal energy transfers in a proper orthogonal decomposition-Galerkin representation of a turbulent separated flow, Journal of Fluid Mechanics **491**, 275 (2003).

[101] A. A. Kaptanoglu, J. L. Callaham, C. J. Hansen, A. Aravkin, and S. L. Brunton, Promoting global stability in data-driven models of quadratic nonlinear dynamics, arXiv preprint arXiv:2105.01843 (2021).

[102] S. Galtier, *Introduction to modern magnetohydrodynamics* (Cambridge University Press, 2016).

[103] S. Chaturantabut and D. C. Sorensen, Discrete empirical interpolation for nonlinear model reduction, in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference* (IEEE, 2009) pp. 4316–4321.

[104] Z. Drmac and S. Gugercin, A new selection operator for the discrete empirical interpolation method—improved a priori error bound and extensions, SIAM Journal on Scientific Computing **38**, A631 (2016).

[105] P. Astrid, S. Weiland, K. Willcox, and T. Backx, Missing point estimation in models described by proper orthogonal decomposition, IEEE Transactions on Automatic Control **53**, 2237 (2008).

[106] K. Willcox, Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition, Computers & Fluids **35**, 208 (2006).

[107] K. Carlberg, C. Farhat, J. Cortial, and D. Amsallem, The GNAT method for nonlinear model reduction: effective implementation and application to computational fluid dynamics and turbulent flows, Journal of Computational Physics **242**, 623 (2013).

[108] B. M. de Silva, K. Champion, M. Quade, J.-C. Loiseau, J. N. Kutz, and S. L. Brunton, PySINDy: A python package for the sparse identification of nonlinear dynamics from data (2020), arXiv:2004.08424 [math.DS].

[109] R. Tibshirani, Regression shrinkage and selection via the lasso, Journal of the Royal Statistical Society. Series B (Methodological) , 267 (1996).

[110] P. Zheng, T. Askham, S. L. Brunton, J. N. Kutz, and A. Y. Aravkin, A unified framework for sparse relaxed regularized regression: SR3, IEEE Access **7**, 1404 (2019).

[111] P. Gelß, S. Klus, J. Eisert, and C. Schütte, Multidimensional approximation of nonlinear dynamical systems, Journal of Computational and Nonlinear Dynamics **14** (2019).

[112] J. L. Callaham, J.-C. Loiseau, G. Rigas, and S. L. Brunton, Nonlinear stochastic modeling with Langevin regression, arXiv preprint arXiv:2009.01006 (2020).

[113] J. J. Bramburger, D. Dylewsky, and J. N. Kutz, Sparse identification of slow timescale dynamics, arXiv preprint arXiv:2006.00940 (2020).

[114] K. Kaheman, J. N. Kutz, and S. L. Brunton, SINDy-PI: a robust algorithm for parallel implicit sparse identification of nonlinear dynamics, Proceedings of the Royal Society A **476**, 20200279 (2020).

[115] A. Cortiella, K.-C. Park, and A. Doostan, Sparse identification of nonlinear dynamical systems via reweighted l_1-regularized least squares, Computer Methods in Applied Mechanics and Engineering **376**, 113620 (2021).

[116] D. B. Brückner, P. Ronceray, and C. P. Broedersz, Inferring the dynamics of underdamped stochastic systems, Physical review letters **125**, 058103 (2020).

[117] S. Beetham and J. Capecelatro, Formulating turbulence closures using sparse regression with embedded form invariance, Physical Review Fluids **5**, 084611 (2020).

[118] J. Z. Kolter and G. Manek, Learning stable deep dynamics models, in *Advances in Neural Information Processing Systems*, Vol. 32 (2019) pp. 11128–11136.

[119] S. Pan and K. Duraisamy, Physics-informed probabilistic learning of linear embeddings of nonlinear dynamics with guaranteed stability, SIAM Journal on Applied Dynamical Systems **19**, 480 (2020).

[120] L. Boninsegna, F. Nüske, and C. Clementi, Sparse learning of stochastic dynamical equations, The Journal of Chemical Physics **148**, 241723 (2018).

[121] V. Zucatti and W. Wolf, Data-driven closure of projection-based reduced order models for unsteady compressible flows, arXiv preprint arXiv:2103.12727 (2021).

[122] J. S. Wrobel, *A study of HIT-SI plasma dynamics using surface magnetic field measurements* (University of Washington, 2011).

[123] B. Victor, T. Jarboe, C. Hansen, C. Akcay, K. Morgan, A. Hossack, and B. Nelson, Sustained spheromaks with ideal n=1 kink stability and pressure confinement, Physics of Plasmas **21**, 082504 (2014).

[124] C. Sovinec, A. Glasser, T. Gianakon, D. Barnes, R. Nebel, S. Kruger, D. Schnack, S. Plimpton, A. Tarditi, M.-S. Chu, *et al.*, Nonlinear magnetohydrodynamics simulation using high-order finite elements, Journal of Computational Physics **195**, 355 (2004).

[125] C. Akcay, *Extended magnetohydrodynamic simulations of the helicity injected torus (HIT-SI) spheromak exper-iment with the NIMROD code*, Ph.D. thesis, University of Washington, Seattle (2013).

[126] C. Akcay, C. C. Kim, B. S. Victor, and T. R. Jarboe, Validation of single-fluid and two-fluid magnetohydrodynamic models of the helicity injected torus spheromak experiment with the NIMROD code, Physics of Plasmas **20**, 082512 (2013).

[127] K. Morgan, T. Jarboe, A. Hossack, R. Chandra, and C. Everson, Validation of extended magnetohydrodynamic simulations of the HIT-SI3 experiment using the NIMROD code, Physics of Plasmas **24**, 122510 (2017).

[128] P. J. Morrison and J. M. Greene, Noncanonical Hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics, Physical Review Letters **45**, 790 (1980).

[129] Z. Yoshida and E. Hameiri, Canonical Hamiltonian mechanics of Hall magnetohydrodynamics and its limit to ideal magnetohydrodynamics, Journal of Physics A: Mathematical and Theoretical **46**, 335502 (2013).

[130] H. M. Abdelhamid, Y. Kawazura, and Z. Yoshida, Hamiltonian formalism of extended magnetohydrodynamics, Journal of Physics A: Mathematical and Theoretical **48**, 235502 (2015).

[131] H. K. Chu and M. Hayashibe, Discovering interpretable dynamics by sparsity promotion on energy and the Lagrangian, IEEE Robotics and Automation Letters **5**, 2154 (2020).