Application of a non-intrusive reduced order modeling approach to magnetohydrodynamics

RESEARCH ARTICLE | OCTOBER 29 2024

Application of a non-intrusive reduced order modeling approach to magnetohydrodynamics

M. Lo Verso ; S. Riva ; C. Introini ; E. Cervi ; F. Giacobbo; L. Savoldi ; M. Di Prinzio ; M. Caramello; L. Barucca; A. Cammi

Physics of Fluids 36, 107167 (2024)
https://doi.org/10.1063/5.0230708

Physics of Fluids ARTICLE pubs.aip.org/aip/pof

Application of a non-intrusive reduced order modeling approach to magnetohydrodynamics

Cite as: Phys. Fluids **36**, 107167 (2024); doi: 10.1063/5.0230708
Submitted: 25 July 2024 · Accepted: 8 October 2024 ·
Published Online: 29 October 2024

M. Lo Verso,[1] S. Riva,[1] C. Introini,[1] E. Cervi,[1] F. Giacobbo,[1] L. Savoldi,[2] M. Di Prinzio,[3] M. Caramello,[3] L. Barucca,[3] and A. Cammi[1,a)]

AFFILIATIONS

[1]Politecnico di Milano, Department of Energy, CeSNEF-Nuclear Engineering Division, Nuclear Reactors Group, via La Masa, 34 20156 Milano, Italy
[2]MATHEP Group, Department of Energy "Galileo Ferraris", Politecnico di Torino, Torino, Italy
[3]Ansaldo Nucleare SpA, Genova, Italy

[a)]**Author to whom correspondence should be addressed:** antonio.cammi@polimi.it

ABSTRACT

Magnetohydrodynamics (MHD) investigates the intricate relationship between electromagnetism and fluid dynamics, offering a complete insight into the behavior of conducting fluids under the influence of magnetic fields. This theory plays a pivotal role in the framework of magnetic confinement fusion, where it can be applied to describe both thermonuclear plasmas confined inside the vacuum vessel and operating fluids, such as liquid metals and molten salts, flowing within the blanket of future tokamaks. Currently, the state-of-the-art numerical modeling of MHD scenarios employs a multi-physics framework to examine the interplay between magnetic fields and thermal hydraulics; however, due to the complexity of the involved physics, detailed models are required, resulting in a significant computational burden. In this regard, reduced order modeling (ROM) techniques may represent a promising solution, as they enable approximating complex systems with lower-dimensional models. Indeed, ROM methodologies can significantly reduce the required computational time while maintaining accuracy in capturing the convoluted physics involved in fusion reactors, especially in the contexts of sensitivity analysis, uncertainty quantification, and control. Despite their potential, ROM methods are relatively under-explored within the MHD framework; this study applies ROM techniques to MHD scenarios, focusing on their capabilities and possible limitations. To this aim, the backward-facing step, which is well suited for exploring the effects of different magnetic fields on turbulent dynamics, is adopted as case study. In particular, this work evaluates the potentialities of the ROM approach in enhancing computational efficiency within the MHD domain. Each of the methods evaluated was effective in precisely reconstructing flow dynamics at any given time and across the full range of magnetic field values tested while significantly reducing computational costs compared to full-order simulations. Practically, this study demonstrates the feasibility to create simplified models that accurately represent the magnetohydrodynamic flows of fluids within the blanket.

I. INTRODUCTION

Magnetohydrodynamics (MHD) studies the dynamics of electrically conducting fluids, i.e., fluids with a non-negligible electrical conductivity, under the influence of a magnetic field.[1] This theory provides mathematical models widely exploited in the nuclear field, including magnetic confinement fusion (MCF), since thermonuclear plasmas can be treated as conducting fluids confined by intense magnetic fields.[2] Indeed, charged particles tend to follow the magnetic field lines, reducing their ability to escape the confined region. This is critical for controlling plasma in tokamaks, where magnetic fields help to maintain the stability of high-temperature plasmas. Moreover, the MHD theory can also describe conducting liquid metals that flow in the blanket of EU-DEMO-like reactors,[3] or molten salts as foreseen in the ARC reactor.[4] In these systems, the magnetic field confines the plasma inside the vacuum vessel. At the same time, residual magnetic field lines could emerge from the plasma chamber and reach the liquid metals or molten salts flowing in the blanket, impacting their fluid dynamics. Then, when designing this component, the presence of a magnetic field varying in amplitude and direction over time and affecting the flow regime of these conducting fluids must be considered. In addition to fusion reactor applications, the MHD theory may be adopted for the analyses of innovative engineering systems that make

use of conducting fluids in the presence of an electromagnetic field, as the magnetic field might enhance the thermal-hydraulics and the overall dynamics of the operating fluid (in terms of pressure drops, velocity, thermal diffusion, flow regime). For these reasons, being crucial to know the effects of the magnetic field on the motion of these fluids and how it influences their dynamics, numerical investigations become essential to explore these phenomena.

The actual effects on fluid dynamics strongly depend on the intensity and direction of the imposed magnetic field, and simulating every possible case is prohibitive from a computational point of view. The presence of a large number of potential cases becomes even more relevant when it comes to real-time applications for control purpose: in general, highly detailed models might be able to predict even unforeseen conditions; however, their response time will likely be too high for any meaningful real-time action. Indeed, MHD models are systems of nonlinear and highly complex equations, where velocity and magnetic field are coupled in a multi-physics framework.[5] Therefore, the necessity arises to find a strategy to reduce the computational complexity in simulating MHD scenarios. In this context, reduced order modeling (ROM) techniques[6,7] can be a promising solution for approximating complex systems with surrogate models, especially in multi-query and real-time contexts.[8] These approaches minimize the computational efforts whilst keeping the accuracy at a desired level; in particular, they are frequently employed to simulate unseen fluid dynamics scenarios[9–12] without performing new calculations on the original high-fidelity model. On one side, for multi-query scenarios that require multiple evaluations of the high-fidelity model, the availability of ROM approaches could significantly reduce the computational costs: for instance, sensitivity analyses may exploit the ROM itself. Conversely, reducing the computational time would allow the application of ROMs in real-time contexts, such as control-oriented applications, where it is essential to rapidly verify the instantaneous effects of residual magnetic fields on the blanket fluids, both for operational and safety reasons.

The application of ROM methods to MHD scenarios, where the magnetic field also comes into play, is still limited in the literature. In particular, all the few studies that adopt ROM techniques for MHD scenarios involve thermonuclear plasma. For example, Taylor *et al.*[13] applied the dynamic mode decomposition (DMD) for the diagnostic analysis of the non-linear dynamics of a magnetized plasma in resistive magnetohydrodynamics; Kaptanoglu *et al.*[14] applied DMD to extract spatiotemporal magnetic coherent structures from the experimental and simulation datasets of the helicity injected torus with steady inductive experiment; Kaptanoglu *et al.*[15] applied a data-driven version of the proper orthogonal decomposition (POD) Galerkin models for compressible plasmas, relying on physically constrained sparse-identification of non-linear dynamics.[16] However, according to the authors' knowledge, there are not many studies using ROM methodologies for MHD scenarios where the conductive fluids are the liquids flowing in the reactor blanket. Therefore, the present work aims to apply ROM techniques, well-known for incompressible flows,[10,17] to magnetohydrodynamic scenarios involving fluids foreseen in the design of the blanket of future tokamaks: in this work, the focus is on the generation of surrogate models, encoding sufficiently high information from the high-fidelity data, for their future use in sensitivity analysis and uncertainty quantification. This paper applies a *non-intrusive* ROM technique, based on the combination of the POD with Gaussian process regression to learn the dynamics[18,19] on an MHD flow in a backward facing step (BFS), with different initial value of the vertical component of the magnetic field.

The structure of the present paper is now reported. Section II provides an overview of the reduced order modeling approach and a description of the specific techniques adopted in this work. Section III describes the MHD model and presents the numerical results obtained with the ROM approach. Finally, Sec. IV resumes the main conclusions of the present work.

II. DATA-DRIVEN REDUCED ORDER MODELING TECHNIQUES

Dimensionality reduction methods are typically used to decrease the computational cost required to numerically solve a partial differential equation (PDE), with the natural application for multi-query and real-time scenarios.[6,7] In particular, one of the most wide-spread approaches is the reduced basis (RB) method in which the solution manifold $\mathscr{U} \subset L^2(\Omega)$, given $\Omega \subset \mathbb{R}^3$ the spatial domain, is approximated by a finite-dimensional space, spanned by basis functions obtained by some training solutions $\{u_i = u(\mathbf{x}; \boldsymbol{\mu}_i)\}_{i=1}^{N_s}$ of the PDEs under study, called *snapshots*, dependent on the spatial coordinate $\mathbf{x} \in \Omega$ and the parameters $\boldsymbol{\mu}$ (including, for instance, time t, Reynolds number Re, initial conditions, and material properties). In this context, the RB approach allows us to approximate any element of the solution manifold with a linear combination of the basis functions, i.e.,

$$u(\mathbf{x}; \boldsymbol{\mu}) \simeq \sum_{n=1}^{N} \alpha_n(\boldsymbol{\mu}) \cdot \varphi_n(\mathbf{x}), \tag{1}$$

in which $\{\varphi_n\}_{n=1}^{N}$ are the basis functions, spanning the reduced space X_N approximating the manifold $\mathscr{U}$, and $\{\alpha_n\}_{n=1}^{N}$ are the modal/reduced coefficients in which the parametric dependencies are embedded.

All reduced order modeling methods are characterized by the so-called offline-online paradigm,[7] which can be resumed as follows: first, the high-fidelity full-order model (FOM) is solved several times to generate the training snapshots needed to compute the basis functions (and hence the reduced space itself); then, the modal coefficients are computed according to the adopted technique. These approaches can belong to two main categories, *intrusive* and *non-intrusive*. The former takes the governing equations and adopts a Galerkin projection[9] through the basis functions previously computed to obtain a reduced order model consisting of a relatively small set of ordinary differential equations (ODEs), which can be solved several times with much lower computational costs. The latter can either use online measurements[20,21] or machine learning methods[18,22,23] to obtain the modal coefficients in a *grey-box* framework, starting from some available data (either full-order *snapshots* or measurements). Even though the intrusive approach is more physically consistent, since a surrogate model is directly derived from the FOM itself, it may be difficult to define a ROM for strongly non-linear problems with several coupled PDEs, such as the MHD case. Moreover, the system of ODEs is typically stiff, for which proper numerical schemes are needed.[24,25] On the other hand, non-intrusive approaches do not require access to the governing equations and they are more suited to be adopted in a fully data-driven framework.

A. Proper orthogonal decomposition

In the context of ROM, proper orthogonal decomposition (POD) is the most widespread approach to obtain low-rank approximations and is often taken as reference ROM technique against which to compare other ROM methods. The POD lays its foundation on the singular value decomposition (SVD),[26] a linear algebra procedure aiming at generalizing the eigen-decomposing of a matrix. For the particular case of continuous functions living in a functional space endowed with an inner product [e.g., $L^2(\Omega)$], the POD algorithm is based on the eigendecomposition of the correlation matrix $\mathbb{C} \in \mathbb{R}^{N_s \times N_s}$, defined as

$$\mathbb{C}_{nm} = (u_n, u_m)_{L^2(\Omega)} = \int_\Omega u_n \cdot u_m \, d\Omega \, n, m = 1, \ldots N_s. \tag{2}$$

given N_s is the amount of training solutions of the PDEs. For this matrix, the associated eigenvalue problem is solved

$$\mathbb{C}\boldsymbol{\eta}_n = \lambda_n \boldsymbol{\eta}_n, \tag{3}$$

to obtain the correspondent eigenvalues $\{\lambda_n\}_{n=1}^{N_s}$ and eigenvectors $\{\boldsymbol{\eta}_n \in \mathbb{R}^{N_s}\}_{n=1}^{N_s}$: these are used to generate a set of basis functions named POD modes

$$\varphi_n(\mathbf{x}) = \frac{1}{\sqrt{\lambda_n}} \sum_{l=1}^{N_s} \eta_{n,l} \cdot u_l(\mathbf{x}) \tag{4}$$

recalling that $u_l(\mathbf{x}) = u(\mathbf{x}, \boldsymbol{\mu}_l)$ and that $\eta_{n,l}$ represents the l-component of the n-th eigenvector. The basis functions $\{\varphi_n\}_{n=1}^{N}$ are ortho-normal to each other in L^2-sense,[8] from which a reduced space $X_N = \text{span}(\varphi_1, \varphi_2, \ldots, \varphi_N) \subset \mathscr{U}$ is defined such that any function $u(\mathbf{x}; \boldsymbol{\mu}) \in \mathscr{U}$ can be approximated as follows:

$$u(\mathbf{x}; \boldsymbol{\mu}) \simeq \sum_{n=1}^{N} \alpha_n(\boldsymbol{\mu}) \cdot \varphi_n(\mathbf{x}) \text{ with } \alpha_n = (u, \varphi_n)_{L^2(\Omega)} = \int_\Omega u \cdot \varphi_n \, d\Omega. \tag{5}$$

The dimension of the reduced space N, sometimes also called rank, is typically selected such that the projection error is smaller than a desired tolerance ε_{POD}: namely, N is the smallest integer satisfying

$$I(r) = \frac{\sum_{n=1}^{r} \lambda_n}{\sum_k \lambda_k} \geq 1 - \varepsilon_{POD}^2, \tag{6}$$

with $I(r)$ representing the energy content retained by the first r POD modes, which is known as relative information content.[8]

These basis functions define the dominant spatial behavior of the snapshots, whereas the parametric dependencies are hidden in the modal coefficients $\{\alpha_n(\boldsymbol{\mu})\}_{n=1}^{N}$. The focus now shifts to the methodology to compute these coefficients in a quick, efficient, and reliable way. In this work, a *non-intrusive* approach is pursued, in which the functional dependence between $\boldsymbol{\mu}$ and $\alpha_n(\boldsymbol{\mu})$ is learned through Gaussian process regression (GPR) supervised learning method,[27,28] adopting the same strategy proposed by Guo and Hesthaven,[22,23] and later applied by Ortali *et al.*[18] and Cicci *et al.*[19]

The generation of the POD modes and the state estimation [Eq. (1)] have been implemented in Python within the *pyforce* package,[29,30] built with the *dolfinx* package (v. 0.6.0), part of the FEniCSx library[31–34] and openly available on Github under the MIT license.

B. Gaussian process regression

A Gaussian process (GP) is any collection of random variables, in which each finite subset obeys a joint Gaussian distribution; the Gaussian process regression (GPR) is a supervised learning algorithm able to generate a model based on the knowledge coming from some training data, to predict continuous quantities of interest.[28]

Let $\mathscr{D} = \{(\mathbf{x}_i, y_i) : i = 1, 2, \ldots, N_s\}$ be the training set of N_s observations, given $\mathbf{x}_i \in \mathbb{R}^d$ as the input with d entries and y_i the corresponding output. Supervised learning algorithms find a model f that can infer the relationship between $\mathbf{x}$ and y. The GPR method[28] infers a probabilistic distribution over functions given the observation data, used to make predictions given new (unknown) inputs $\mathbf{x}$. For GPR, given a mean $\mathbb{E}(\mathbf{x})$ and a covariance function (or kernel) $\kappa_\theta(\mathbf{x}, \mathbf{x}')$, the prior distribution of f is a GP written as

$$f(\mathbf{x}) \sim \text{GP}\big(\mathbb{E}(\mathbf{x}), \kappa_\theta(\mathbf{x}, \mathbf{x}')\big), \tag{7}$$

in which $\mathbb{E}(\mathbf{x})$ is typically assumed to be a constant function based on the training data and the kernel $\kappa_\theta(\mathbf{x}, \mathbf{x}')$ is dependent on some hyper-parameters θ, to be tuned. In the literature, there are plenty of options for κ: one of the most common choices is the radial basis function (RBF) kernel

$$\kappa_\theta(\mathbf{x}, \mathbf{x}') = \theta_1 e^{-\frac{\|\mathbf{x}-\mathbf{x}'\|_2^2}{2\theta_2^2}}. \tag{8}$$

The hyper-parameters $\theta = [\theta_1, \theta_2]$ need to be tuned to augment the predictive performance of the GPR. GPR models embed knowledge from training data $\mathscr{D}$ into the prior distribution to obtain a posterior. Hence, given a set of training input $\mathbb{X} = [\mathbf{x}_1 | \mathbf{x}_2 | \ldots \mathbf{x}_{N_s}] \in \mathbb{R}^{d \times N_s}$ with correspondent output $\mathbf{y} = [y_1, y_2, \ldots y_{N_s}]^T \in \mathbb{R}^{N_s}$ and a set of new test input data $\mathbb{X}^* \in \mathbb{R}^{d \times M}$, predictions $\mathbf{y}^* \in \mathbb{R}^M$ can be made by exploiting the theorem of conditional Gaussian so that the conditional distribution of the predicted GP is given by

$$\mathbf{y}^* = f(\mathbb{X}^* | \mathbf{y}, \mathbb{X}, \mathbb{X}^*) \sim \text{GP}\Big(\hat{\mathbb{E}}(\mathbb{X}^*), \hat{\kappa}(\mathbb{X}, \mathbb{X}^{*\prime})\Big), \tag{9}$$

in which the knowledge of the mean and covariance function updated as follows:[19]

$$\begin{aligned} \hat{\mathbb{E}}(\mathbb{X}^*) &= \mathbb{E}(\mathbb{X}^*) + \mathbb{K}_* \mathbb{K}^{-1} (\mathbf{y} - \mathbb{E}(\mathbb{X})), \\ \hat{\kappa}(\mathbb{X}, \mathbb{X}^{*\prime}) &= \mathbb{K}_{**} - \mathbb{K}_* \mathbb{K}^{-1} \mathbb{K}_*^T, \end{aligned} \tag{10}$$

given $\mathbb{K} = \kappa(\mathbb{X}, \mathbb{X}')$, $\mathbb{K}_* = \kappa(\mathbb{X}^*, \mathbb{X}')$, $\mathbb{K}_*^T = \kappa(\mathbb{X}, \mathbb{X}^{*\prime})$, $\mathbb{K}_{**} = \kappa(\mathbb{X}^*, \mathbb{X}^{*\prime})$. As stated before, the kernel functions depend on hyper-parameters θ, which are required to be tuned by maximizing the loss function log-marginal likelihood as in the works of Cicci *et al.*[19] and Rasmussen and Williams"[28]

$$\begin{aligned} \theta_{\text{opt}} &= \arg\max_\theta \log P(\mathbf{y}|\mathbb{X}), \\ &= \arg\max_\theta \left[-\frac{1}{2} \mathbf{y}^T \mathbb{K}^{-1}(\theta) \mathbf{y} - \frac{1}{2} \log |\mathbb{K}(\theta)| - \frac{N_s}{2} \log (2\pi) \right], \end{aligned} \tag{11}$$

where $P(\mathbf{y}|\mathbb{X})$ is the conditional density function of $\mathbf{y}$ given $\mathbb{X}$, also known as marginal likelihood. The implementation of the GPR and the hyper-parameter tuning are carried out in Python through the open-source package GPy.

1. Tensor-decomposition

The standard version of the GPR, which we will henceforth refer to as monolithic GPR, can consider multiple inputs providing a scalar output quantity; thus, a surrogate model for each reduced coefficient $\alpha_n(\boldsymbol{\mu})$ can be built, i.e.,

$$f_n^{GP} : \boldsymbol{\mu} \rightarrow \alpha_n, \tag{12}$$

and each of them can be trained using data obtained from the training snapshots. For transient problems with large datasets, it is convenient to separate the parameters $\boldsymbol{\mu}$ into two main contributions: the pseudo-parameter t, i.e., the time, and all other parameters, labeled with $\boldsymbol{\mu}$ from now on with abuse of notation, e.g., the initial magnetic field or the Reynolds number. This distinction is usually performed since the time t rigorously is an independent parameter. Supposing the training set of snapshots is composed of N_s parametric solutions of a transient sampling N_t time instants, the total dimension of the snapshot matrix is $\mathcal{N}_h \times (N_t \cdot N_s)$, given $\mathcal{N}_h$ the dimension of the spatial grid. Therefore, the GPR model requires training for a $N_t \cdot N_s$-dimensional dataset per each coefficient α_n.

The literature shows an alternative approach for parametric-transient datasets, as Guo[23] proposed a tensor-decomposition algorithm to separate the parametric dependency from the time one; this can be useful to reduce the time required for the training of the ML regression models. Let $\mathbb{A}_n \in \mathbb{R}^{N_t \times N_s}$ the corresponding matrix for the nth coefficient $\alpha_n(t, \boldsymbol{\mu})$, defined as

$$\mathbb{A}_n = \begin{bmatrix} \alpha_n\left(t^1, \boldsymbol{\mu}_1\right) & \dots & \alpha_n\left(t^1, \boldsymbol{\mu}_{N_s}\right) \\ \vdots & & \vdots \\ \alpha_n\left(t^{N_t}, \boldsymbol{\mu}_1\right) & \dots & \alpha_n\left(t^{N_t}, \boldsymbol{\mu}_{N_s}\right) \end{bmatrix}. \tag{13}$$

The truncated SVD can be performed on $\mathbb{A}_n$ as

$$\mathbb{A}_n \simeq \boldsymbol{\psi}^{(n)} \boldsymbol{\Lambda}^{(n)} \left(\boldsymbol{\phi}^{(n)}\right)^T, \tag{14}$$

given $\boldsymbol{\psi}^{(n)} \in \mathbb{R}^{N_t \times l}$ and $\boldsymbol{\phi}^{(n)} \in \mathbb{R}^{N_s \times l}$ are the time and parametric modes, respectively, $\boldsymbol{\Lambda}^{(n)} \in \mathbb{R}^{l \times l}$ is a diagonal matrix with the singular values of the SVD, and l is the truncation rank obtained using the SVD tolerance as in Eq. (6). For $n = 1, \dots N$, the following maps can be learned:

$$t \rightarrow \{\psi_l^{(n)}(t)\}_{k=1}^{l} \quad \text{and} \quad \boldsymbol{\mu} \rightarrow \{\phi_l^{(n)}(t)\}_{k=1}^{l} \tag{15}$$

for each component of the truncated SVD, hence two Gaussian processes (one for time t and one for the parameters $\boldsymbol{\mu}$) are trained with two smaller datasets (N_t and N_s). This algorithm is known as tensor-decomposition (TD): it allows a strong reduction in the computational times of the hyper-parameter tuning phase,[19,23] which becomes significant when large datasets are considered.

C. POD with GPR as a data-driven technique

In this section, the POD is integrated with the GPR,[18] referred to as POD-GPR, for the specific application in MHD: the basic ideas consist of generating a low-rank representation of some training snapshots through the POD, then infer the reduced coefficients in the time-parameters space to approximate the solution manifold using GPR, in a fully data-driven framework: in fact, no information about the physical model and the correspondent governing equations is employed for the encoding with POD and the inferring with GPR.

The input dataset is represented by the snapshot matrix $\mathbb{S} = [\mathbf{u}_1, \dots, \mathbf{u}_{Nt \cdot N_s}] \in \mathbb{R}^{\mathcal{N}_h \times (N_t \cdot N_s)}$, given $\mathbf{u}_i \in \mathbb{R}^{\mathcal{N}_h}$ the discrete representation of the i th snapshot, e.g., the evaluation of the continuous function at the degrees of freedom of the spatial mesh of dimension $\mathcal{N}_h$. Each snapshot $\mathbf{u}_i$ depends on the training time and the parameters $\boldsymbol{\mu}$, belonging to the training set Ξ_{train}. The POD is used to generate a reduced representation of the training snapshots through the POD modes $\{\varphi_n\}_{n=1}^{N}$ with the truncation rank N obtained from Eq. (6).

Hence, the POD modes can be used to find a reduced representation of the snapshots through the coefficients $\{\alpha_n(t, \boldsymbol{\mu})\}_{n=1}^{N}$ dependent on training time t and parameters $\boldsymbol{\mu}$. Each component can be considered as a scalar output onto which a probabilistic response model can be built using Gaussian process regression. This allows learning the map between the input parameters (t and $\boldsymbol{\mu}$) and each reduced coefficient α_n. These steps are all executed during the offline phase, which is computationally demanding even though it is performed only once. The steps of the offline phase are reported in Algorithm I.

During the online phase, the aim is to obtain an approximation of the full-order solution for a case not used for the training with $(t^\star, \boldsymbol{\mu}^\star)$. At first, the GPR models, generated during the offline phase, are evaluated for the unseen parameters to find the correspondent reduced coefficients $\{\alpha_n(t^\star, \boldsymbol{\mu}^\star)\}_{n=1}^{N}$. In the end, the decoding step is performed: the full state is reconstructed using a linear combination of the POD modes weighted against the aforementioned reduced coefficients.

III. NUMERICAL RESULTS

A. Test case

As a test case, this work considers a flow in the bi-dimensional backward-facing step (BFS). Although this geometry is not representative of any realistic configuration involved in fusion reactors, it is chosen for its simplicity in investigating the effects of different magnetic

ALGORITHM 1: POD-GPR (Offline Phase).

Input
Training Parameters $(t, \boldsymbol{\mu}) \in \Xi_{\text{train}}$;
Snapshot matrix $\mathbb{S} = [\mathbf{u}_1, \dots, \mathbf{u}_{Nt \cdot N_s}] \in \mathbb{R}^{\mathcal{N}_h \times (N_t \cdot N_s)}$;
Output
POD modes $\{\varphi_1(\mathbf{x}), \dots, \varphi_N(\mathbf{x})\}$;
GPR models for the reduced coefficients $\{\alpha_n\}_{n=1}^{N}$;
Proper Orthogonal Decomposition
Computation of correlation matrix $\mathbb{C}_{nm} = (u_n, u_m)_{L^2(\Omega)} = \int_\Omega u_n \cdot u_m \, d\Omega$;
Eigen-decomposition $\mathbb{C}\boldsymbol{\eta}_n = \lambda_n \boldsymbol{\eta}_n$;
Selection of the rank with Eq. (6);
Definition of the POD modes $\{\varphi_n(\mathbf{x}) = \sqrt{\lambda_n^{-1}} \sum_{l=1}^{N_s} \eta_{n,l} \cdot u_l(\mathbf{x})\}_{n=1}^{N}$;
Gaussian Process Regression
Computation of reduced coefficients $\{\alpha_n\}_{n=1}^{N}$ by projection;
Training of a GPR model per each α_n;

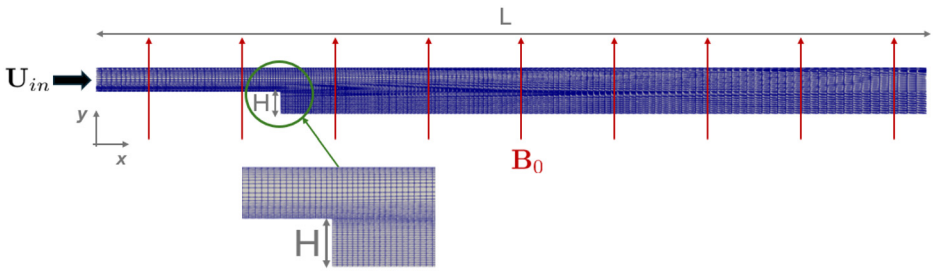

FIG. 1. Computational domain and grid for the BFS case study.

fields on recirculation dynamics. The assumption of 2D flow is accepted since no significant variation in span-wise direction is expected in the separation and reattachment regions. Indeed, the goal of this work is to investigate the potential applications of ROM for MHD problems, focusing on the application itself rather than on the complexity of the test case: further works will focus on more realistic geometries.

The considered fluid is the lead–lithium eutectic, which is the typical operating liquid metal foreseen in several concepts of breeding blankets, such as the water-cooled lead–lithium candidate system for the European DEMO.[35] The flow is described with the following incompressible, viscoresistive magnetohydrodynamic full-order model:[36]

$$\rho\frac{\partial \mathbf{u}}{\partial t}+\rho(\mathbf{u}\cdot\nabla)\mathbf{u}=-\nabla p+\mu\Delta\mathbf{u}+\left(\frac{1}{\mu_0}\nabla\times\mathbf{B}\right)\times\mathbf{B}, \tag{16a}$$

$$\frac{\partial \mathbf{B}}{\partial t}=\nabla\times(\mathbf{u}\times\mathbf{B})+\frac{\eta}{\mu_0}\Delta\mathbf{B}, \tag{16b}$$

$$\nabla\cdot\mathbf{B}=0, \tag{16c}$$

where $\mathbf{u}$ is the velocity of the fluid; $\mathbf{B}$ is the magnetic field; p is the pressure; and ρ, μ, and η represent the density, the dynamic viscosity, and the electrical resistivity of the fluid, while μ_0 denotes the magnetic permeability. This model combines the Navier–Stokes and Maxwell equations with a bidirectional coupling between velocity and magnetic field. The energy equation is neglected in this phase, as the emphasis is on the flow field rather than on the temperature field.

The fluid dynamics is affected by the magnetic field through the Lorentz force, represented by the term $(\frac{1}{\mu_0}\nabla\times\mathbf{B})\times\mathbf{B}$ in Eq. (16a). Then, the magnetic field reacts to the change in the fluid velocity, due to the term $\nabla\times(\mathbf{u}\times\mathbf{B})$ in Eq. (16b), and an induced magnetic field arises and superimposes itself to the external one. To complete the set-up, initial and boundary conditions are required. The fluid is considered initially still; moreover, for the sake of simplicity (without affecting the applicability of the method), the imposed magnetic field will be assumed to be oriented in the vertical direction

$$\begin{cases}\mathbf{u}|_{t=0}=\mathbf{0},\\ \mathbf{B}|_{t=0}=\mathbf{B}_0=B_{0,y}\mathbf{e}_y,\end{cases} \tag{17}$$

given $\mathbf{e}_y=[0,1,0]^T$ the unit vector of the y-axis. Moreover, a uniform velocity is supposed to enter the inlet while no-slip conditions and perfect electrical conductance are assumed for the top and bottom walls

$$\begin{cases}\mathbf{u}|_{inlet}=\mathbf{u}_{in},\\ \left.\dfrac{\partial\mathbf{u}}{\partial\mathbf{n}}\right|_{outlet}=0,\\ \mathbf{u}|_{wall}=\mathbf{0},\\ \left.\dfrac{\partial\mathbf{B}}{\partial\mathbf{n}}\right|_{\partial\Omega}=0,\\ p|_{outlet}=10^5\,Pa,\\ \left.\dfrac{\partial p}{\partial\mathbf{n}}\right|_{\partial\Omega\setminus outlet}=0,\end{cases} \tag{18}$$

where u_{in} is taken equal to 0.32 m/s in order to have $Re\simeq 8000$, thus ensuring a turbulent regime for the case without a magnetic field. For the BFS case, the Reynolds number, which represents the ratio between inertial and viscous forces, is defined as

$$Re=\frac{\rho u_{in}H}{\mu}, \tag{19}$$

where H (m) represents the depth of the step. No turbulence model has been implemented since the magnetic field tends to suppress turbulence, favoring a completely laminar flow. Indeed, the Lorentz force induced by the presence of a vertical magnetic field acts in the opposite direction to the flow with a direct proportionality to the local velocity, resulting in a global slowdown of the flow and a damping of turbulence. The laminarisation induced by the magnetic field is observed even at the lowest values imposed for the intensity of $\mathbf{B}$. Figure 1 reports a sketch of the simulated benchmark, along with the adopted numerical mesh. All the details on the physical and numerical parameters are summarized in Table I (the temperature reported is the one at which the thermophysical properties are evaluated).

In the following, different values for $\mathbf{B}_0\in\mathbb{R}^3$ are considered, to explore different effects of various magnetic intensities on the flow regime: specifically, the initial magnetic field is only directed in the y-component as $\mathbf{B}_0=[0,B_{y,0},0]^T$ [Eq. (17)]. For each value of the imposed magnetic field, the flow reaches a steady state before 4 s, which is considered the final time of each simulation; the discretization

TABLE I. Physical and numerical parameters for the FOM.

ρ	9806 kg/m^3	u_{in}	0.321336 m/s	$\mathcal{N}_h$	7500
μ	1.93e-3 Pa · s	B_{in}	[0.25, 1] T	L	0.18 m
μ_0	$1.26e-6$ H/m	p_{out}	10^5 Pa	H	0.0049 m
η	$1.28e-6\Omega$·m	T	600 K	z	$1e-5$ m (1 cell)

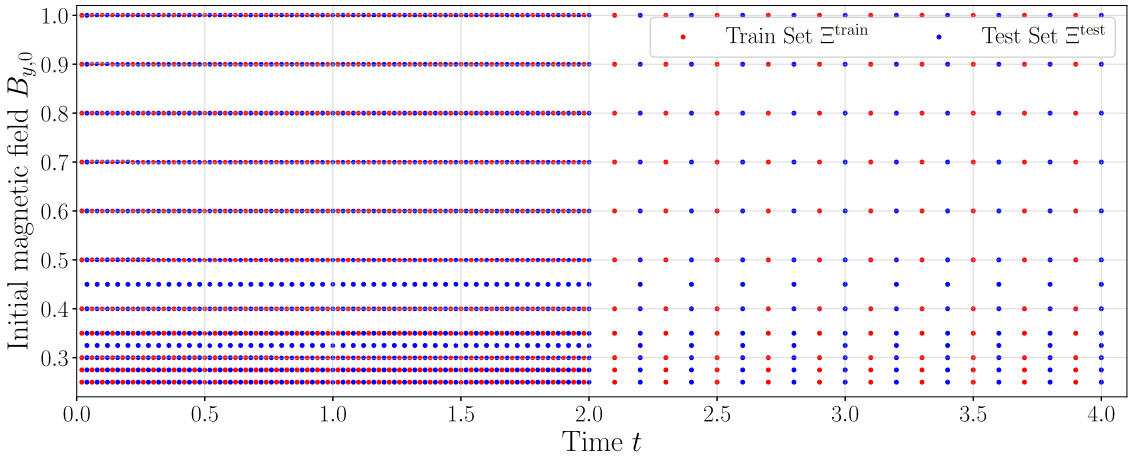

FIG. 2. Separation of the train and test set in the parametric domain D.

time varies at any time step, according to the *CFL* condition, to ensure numerical stability. The simulations have been performed adopting the finite volume discretization in OpenFOAM-org (v.7)[37] using the library *magnetoHDFoam* developed in Ref. 38 and available on Github under the MIT license.

B. Offline phase

The full-order model (FOM), represented by the OpenFOAM model described in the previous section, has been solved for different parametric configurations: the parameters $\boldsymbol{\mu} \in \mathscr{D} \subset \mathbb{R}^2$ are given by the time t and the y-component of the initial magnetic field, i.e., $B_{y,0}$. The parametric domain is then defined as follows:

$$t \in \mathscr{T} = [0, 4]\,\text{s} \quad B_{y,0} \in \mathscr{B} = [0.25, 1]\,\mathscr{T} \text{ such that } \mathscr{D} = \mathscr{T} \times \mathscr{B}, \tag{20}$$

from which a train Ξ^{train} and a test Ξ^{test} set are derived, as displayed in Fig. 2. The non-uniform sampling for the time is driven by physical observation, as most of the dynamics are concentrated in the first time instants (before 2 s), whereas, as the magnetic field increases, the flow is completely laminarised, thus showing less turbulent behavior. Overall, the training dataset is composed of 660 snapshots related to 11 values of the initial magnetic field and 60-time instances. Two different fields are going to be considered: the pressure p and the velocity $\mathbf{u}$; in particular, the former has been rescaled with respect to the outlet boundary condition, i.e., $p|_{outlet} = 10^5$ Pa, thus the reduced pressure $\tilde{p} = p - p|_{outlet}$.

First, the training snapshot matrix is used to perform the POD procedure to obtain the eigenvalues and the associated modes. Both are ordered from the most energetic to the lowest according to the eigenvalues; specifically, low-rank modes contain more information and represent the dominant spatial physics in the system.[6,8] Figure 3 shows on the left the decay of the singular values $\sqrt{\lambda_r}$ as a function of the rank r, presenting good exponential behavior for both the quantities of interest: this means that both solution manifolds can be

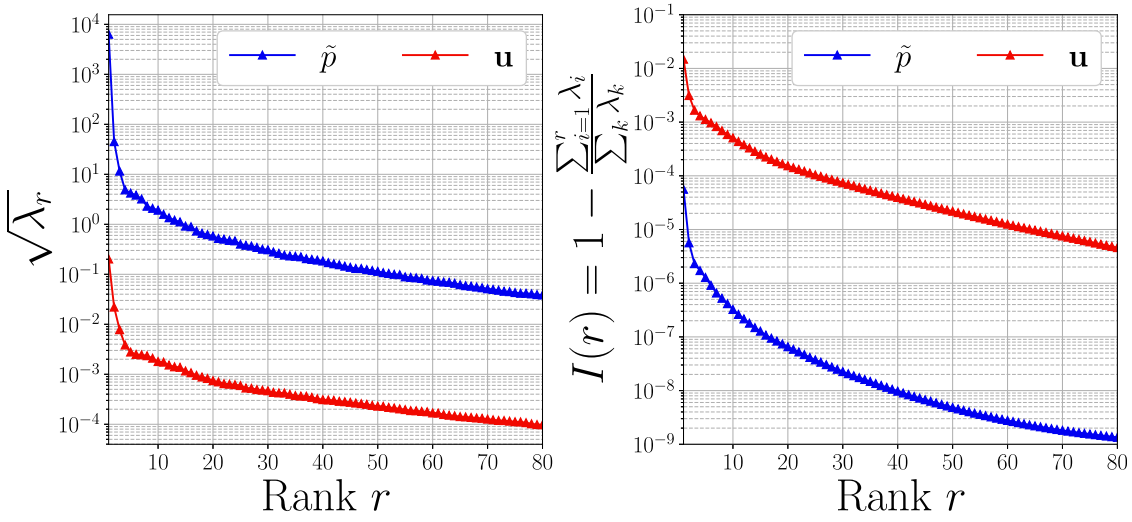

FIG. 3. POD singular values (left) and information/energy content (right) as a function of the rank for the velocity and the reduced pressure field.

approximated by a linear subspace spanned by the POD modes;[8] the right subplot reports the decay of the information/energy content $I(r)$, defined in Eq. (6), depicting how the error of low-rank approximation decreases as more basis functions are added to the subspace. In particular, for the online phase, the maximum number of available modes has been selected to be 20 for both fields in order to have information about the most dominant physics and the smaller scales.

Afterwards, the GP models have been trained using both the monolithic approach and the tensor decomposition (TD), as explained in Sec. II B 1: in both cases, the input data $\mathbf{x}$ of the GP have been rescaled using the *min–max* approach

$$\tilde{\mathbf{x}} = \frac{\mathbf{x} - \mathbf{x}_{min}}{\mathbf{x}_{max} - \mathbf{x}_{min}}, \tag{21}$$

given $\mathbf{x}_{min}$ the minimum of each feature in the training sample and $\mathbf{x}_{max}$ the maximum; whereas, the i th output data y_i using the standardization

$$\tilde{y}_i = \frac{y_i - \langle y_i \rangle}{\sigma_i}, \tag{22}$$

with $\langle y_i \rangle$ as the sample mean on the training dataset and σ_i^2 its sample variance. This is performed to keep the different features in the same range and thus improve the training capabilities of the GPR.[18,26] As already mentioned in Sec. II B, the kernel of the GP is chosen to be a radial basis function (RBF),[28] which is the most widespread choice: the prediction of the surrogate models was satisfying; thus, no hyperparameter tuning on the kernel has been done.

C. Online phase

Once the POD modes and the GP models (both monolithic and tensor-decomposed) have been generated in the offline phase, in this section, the results of the online phase will be presented, focusing on the reconstruction error and the prediction of the reattachment length x_R. Before diving into the discussion, the average absolute error E_N^ϕ and relative error ε_N^ϕ for the generic field ϕ have been defined as follows:

$$E_N^\phi = \frac{1}{\dim(\Xi^{\text{test}})} \sum_{\boldsymbol{\mu} \in \Xi^{\text{test}}} ||r_N^\phi(\mathbf{x}; \boldsymbol{\mu})||_{L^2(\Omega)},$$
$$\varepsilon_N^\phi = \frac{1}{\dim(\Xi^{\text{test}})} \sum_{\boldsymbol{\mu} \in \Xi^{\text{test}}} \frac{||r_N^\phi(\mathbf{x}; \boldsymbol{\mu})||_{L^2(\Omega)}}{||\phi(\mathbf{x}; \boldsymbol{\mu})||_{L^2(\Omega)}}, \tag{23}$$

given $\Xi^{\text{test}} \subset \mathscr{D}$ a subset of the parameter space with test data (not used for the training) and $r_N^\phi(\mathbf{x}; \boldsymbol{\mu})$ the residual field defined as the absolute difference between the full-order solution ϕ and the reconstruction with N modes $\mathscr{P}_N[\phi]$, i.e., $r_N^\phi(\mathbf{x}; \boldsymbol{\mu}) = |\phi(\mathbf{x}; \boldsymbol{\mu}) - \mathscr{P}_N[\phi](\mathbf{x}; \boldsymbol{\mu})|$.

The results of the non-intrusive approach proposed in this work, i.e., the integration of Gaussian process regression with the proper orthogonal decomposition to estimate the modal coefficients, both with the monolithic GPR (POD-GPR) and the GPR with Tensor Decomposition (POD-GPR-TD) have been compared with the best-fit case in which the coefficients are computed through projection (POD-Project) as in Eq. (5). Figure 4 shows the decay of test errors for the reduced pressure $\tilde{p}$ and the velocity $\mathbf{u}$ as the dimension of the reduced space increases: both the cases with GPR interpolation present a "saturation" on the error itself, as found by Cammi *et al.*[29] and Introini *et al.*;[39] nevertheless, the average error on the test set is around 3% for the velocity and 0.3% for the reduced pressure, both values are evaluated by the authors as acceptable since most of the information is encoded correctly and the test output is sufficiently close to the high-fidelity data. The POD-Project case shows a convergence trend; as more modes are added, the reconstruction tends to become more and more similar to the ground-truth given by the full-order solution itself.

A significant quantity of interest for the adopted case study is the reattachment length x_R,[40–42] defined as the maximum distance from

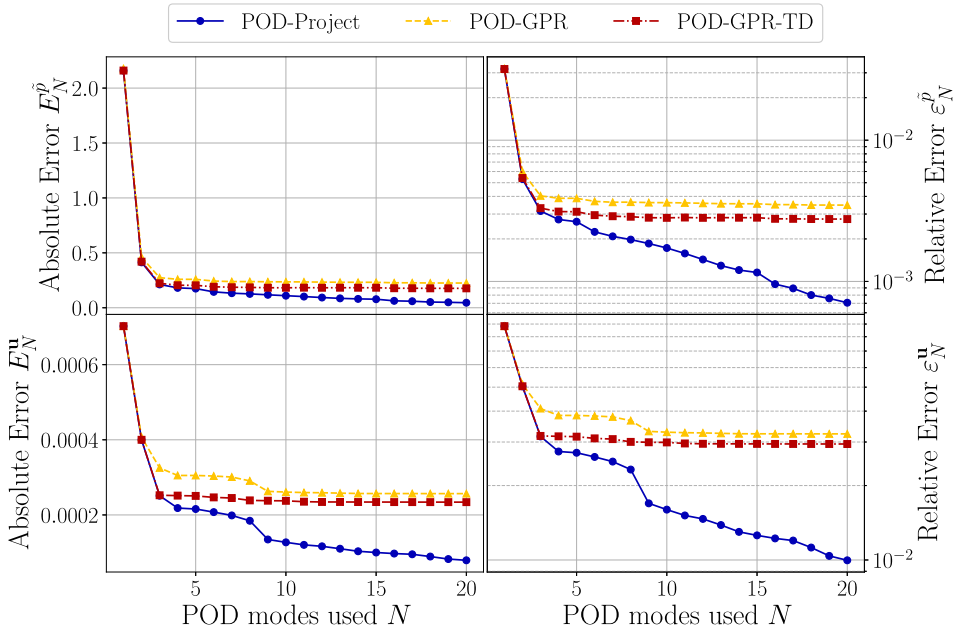

FIG. 4. Online absolute (first column) and relative (second column) test error for the POD-GPR for the pressure (first row) and the velocity (second row), compared with the projection error (best-fit).

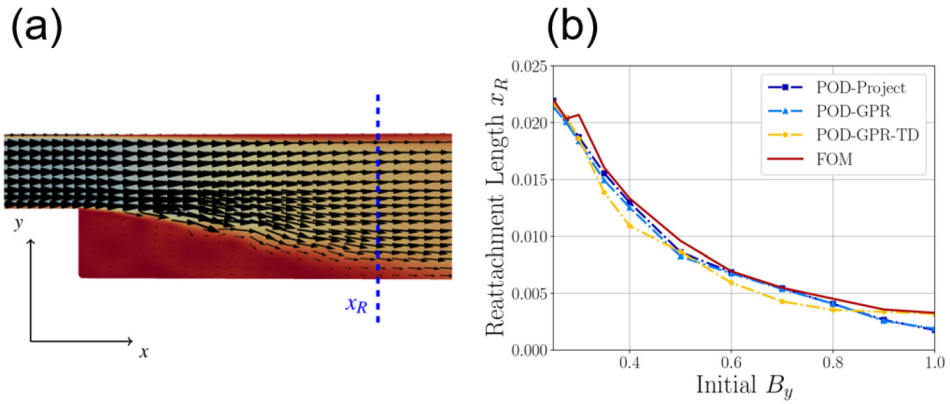

FIG. 5. (a) Definition of the reattachment length x_R. (b) Prediction of x_R (m) by FOM and ROM using ten modes at different magnetic field intensities.

the step where the x-component of the velocity close to the wall ($\Delta y = 0.0004$ m) changes sign, meaning a left-to-right flow along the positive direction of the x-axis. Thus, x_R is a function of time t, and the value is averaged starting from $t = 2$ s, that is, when the flow for each value of the initial magnetic field is laminarised.[43] The results have been plotted as a function of the initial magnetic field B_y in Fig. 5: each ROM method shows quite good agreement with the full-order solution, being able to reconstruct the average flow field and its characteristic quantities.

The algorithms have been executed on a tower computer with an Intel Core i7-9800X processor, whereas the FOM simulations have been performed on an HPC cluster. One of the important advantages of the ROM methods consists in the significant reduction of the computational costs in multi-query and real-time applications compared to a direct numerical solution: Fig. 6 shows the average CPU time needed to estimate a modal coefficient for the pressure and the velocity comparing the three different methods; overall, the time required is quite low even though the POD-GPR-TD case is more CPU consuming because more GPR models have to be evaluated during the online phase. The main advantage of the GPR-TD approach lies in the offline phase due to the reduction of the computational cost of the training phase:[19,23] once the reduced model is built, several cheap simulations can be run very quickly also for parameter values not included in the training range, which is especially useful for uncertainty quantification and parameter estimation. Indeed, for the wall-clock time, the test transient for a given value of the magnetic field can be completed in less than 1 s. In contrast, the computation of the full-order solution requires significantly more time, ranging from 4 to 5 min when high values of the magnetic field are imposed to 50–60 min for low values. The variation in time based on the magnetic field intensity is due to the fact that an elevated magnetic field brings the flow to a steady state much faster, after which all the computed values remain constant.

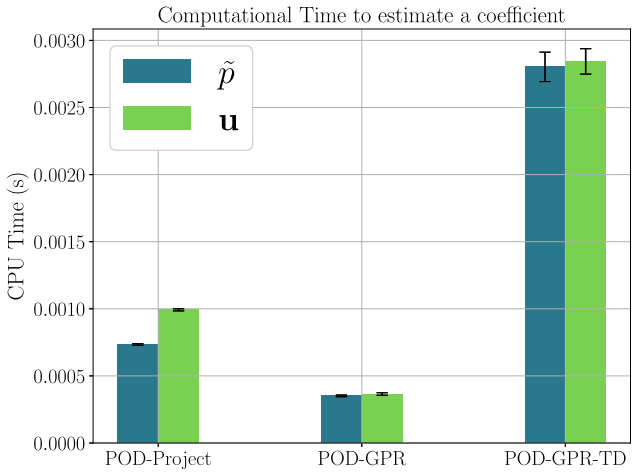

FIG. 6. CPU times of the ROM during the online phase.

In the end, all the considered techniques accurately reconstructed the velocity and pressure fields at any time instant and for any value of a magnetic field in the considered range. Figures 7 and 8 report the velocity field from the FOM and the ROM reconstruction for two-time instants ($t = 2$ s and $t = 4$ s) for the case with $B_y = 0.325$ T, using ten modes: it can be observed that the average flow field is accurately described by any method adopted in this work and that the smaller scales[44] have been suppressed by the POD truncation. This behavior is expected and coherent with the POD algorithm itself, whose first modes correspond to the most energetic ones,[8] being characterized by eddies of larger scales.[45,46] So, the residual is primarily concentrated near the step, where smaller scales evolve with faster dynamics compared to the temporal sampling used. However, by increasing the temporal sampling resolution, it would be possible to capture these higher-frequency dynamics, thereby improving the accuracy of the reconstruction achieved by the POD. Nevertheless, even with the current temporal sampling, the residual remains sufficiently small and acceptable for both pressure and velocity. Moreover, Figs. 7 and 8 put in evidence the effect exerted by the magnetic field on the flow behavior:[43] indeed, since the value of the Reynolds number is very high ($Re = 8000$), the flow would be characterized by extremely turbulent dynamics,[42] if no magnetic fields were applied. Instead, the perpendicular magnetic field laminarises the regime, causing turbulent dynamics to be dampened in favor of a fully laminar flow.

IV. CONCLUSIONS

The present work applies a non-intrusive reduced-order modeling approach to a magnetohydrodynamic scenario consisting of lead–lithium flowing in a BFS geometry. According to the authors' knowledge, this study represents one of the few works focusing on ROM for MHD flows involving fluids representing liquid metals

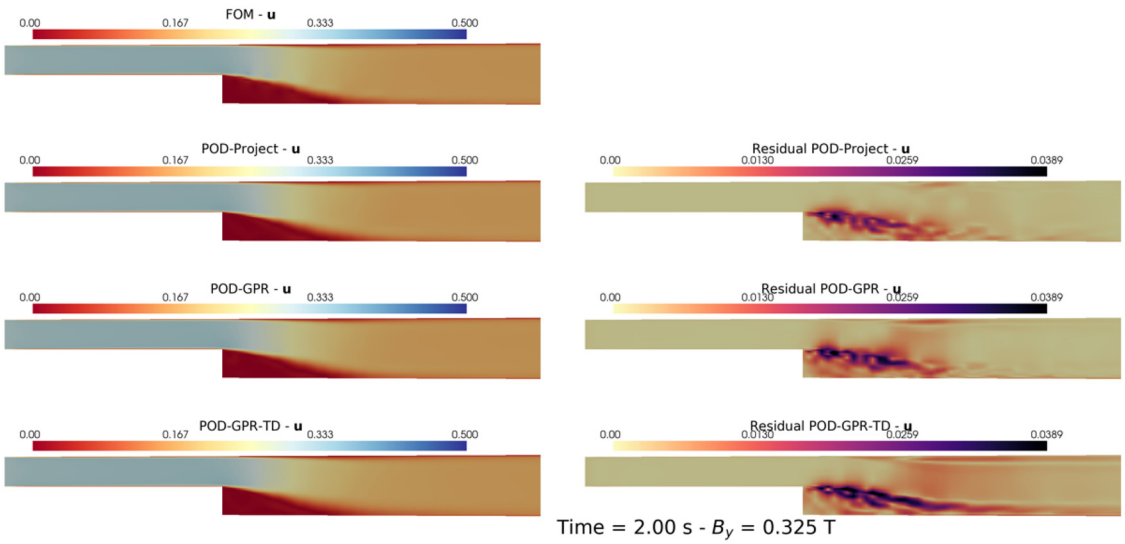

FIG. 7. Comparison of the ROM reconstruction, using ten modes, and the associated residual field for $B_y = 0.325$ T at $t = 2.00$ s.

flowing in breeding blankets. In the offline phase, the reduced model has been trained with several time-dependent snapshots in which different intensities of the applied perpendicular magnetic field were imposed, being the considered parameters represented by the time and vertical component of the magnetic field. This step encodes the spatial information in the reduced coordinate system provided by the reduced basis built by the POD procedure. During the online phase, the reduced coefficients have been estimated using two different supervised learning techniques, i.e., the monolithic GPR and the tensor-decomposed GPR. The results obtained with the two methods have been compared with the best-fit case in which the coefficients are computed by projection. All the explored methods proved to be able to accurately reconstruct the flow dynamics at any time instant and for any value of the imposed magnetic field belonging to the train range, with much lower computational costs compared to a full-order solution. The average CPU time needed to estimate modal coefficients for pressure and velocity is relatively low across the three methods applied. However, the POD-GPR-TD approach requires more CPU resources since it involves evaluating several GPR models during the online phase. The GPR-TD technique is particularly advantageous in the offline phase, where it significantly lowers the training computational burden. Once the reduced model is created, it allows for quick and inexpensive simulations, even for parameters not included in the training range. Thus, from an applicative standpoint, this study shows the

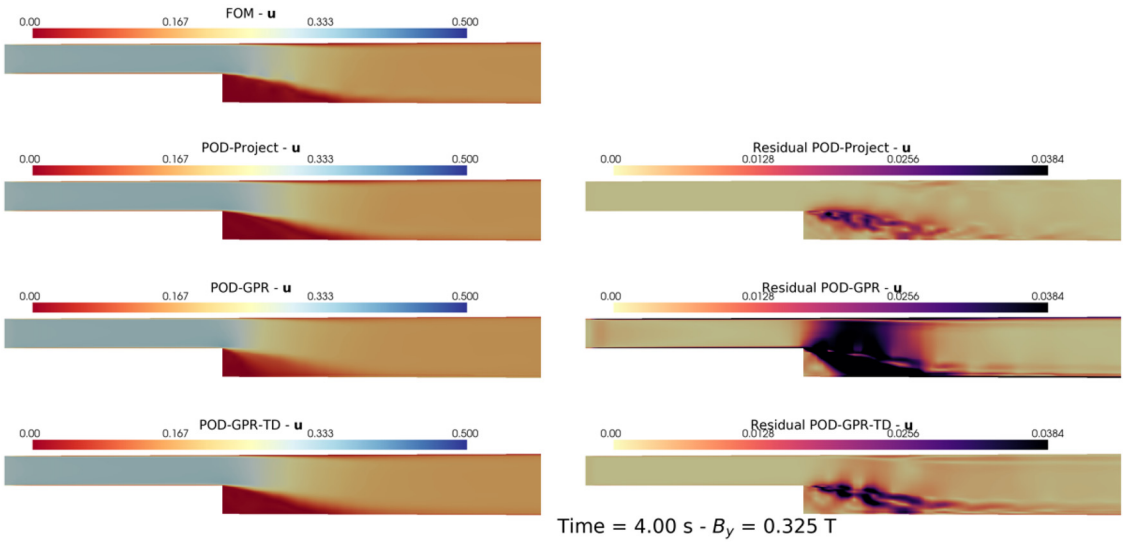

FIG. 8. Comparison of the ROM reconstruction, using ten modes, and the associated residual field for $B_y = 0.325$ T at $t = 4.00$ s.

possibility of generating reduced models representing the MHD effects exerted by residual magnetic fields of any intensity on operating fluids flowing within the blanket. The creation of these models could significantly aid in sensitivity analysis and uncertainty quantification by shifting the calculation to the reduced order level rather than the high-fidelity one. Furthermore, ROMs could facilitate the study of real-time simulations, which would be essential in contexts where controlling the effects of the magnetic field is critical. Future studies could explore the application of these techniques in scenarios involving more general magnetic fields, with arbitrary direction and orientation, as well as time-varying conditions. Additionally, the model could also be trained using data from experimental measurements, further enhancing its accuracy and applicability. Moreover, it is foreseen to apply these approaches to more complex geometries closer to real cases encountered in fusion reactors and to adopt other techniques aimed at combining the information of mathematical models with data directly taken from the physical system.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Matteo Lo Verso: Conceptualization (equal); Methodology (equal); Software (equal); Writing – original draft (equal). **Antonio Cammi:** Conceptualization (equal); Methodology (equal); Writing – review & editing (equal). **Stefano Riva:** Conceptualization (equal); Methodology (equal); Software (equal); Writing – original draft (equal). **Carolina Introini:** Writing – review & editing (equal). **Eric Cervi:** Writing – review & editing (equal). **Francesca Giacobbo:** Writing – review & editing (equal). **Laura Savoldi:** Writing – review & editing (equal). **Matteo Di Prinzio:** Writing – review & editing (equal). **Marco Caramello:** Writing – review & editing (equal). **Luciana Barucca:** Writing – review & editing (equal).

DATA AVAILABILITY

The results of this paper have been implemented in OpenFOAM adopting magnetoHDFoam, available in GitHub at https://github.com/ERMETE-Lab/MHD-magnetoHDFoam, whereas the ROM part has been performed with the pyforce package for Python, available in GitHub at https://github.com/ERMETE-Lab/ROSE-pyforce. Both are openly available under the MIT license.

NOMENCLATURE

Acronyms

BC	Boundary condition
DMD	Dynamic mode decomposition
FOM	Full-order model
GP	Gaussian process
GPR	Gaussian process regression
MCF	Magnetic confinement fusion
MHD	Magnetohydrodynamics
ODE	Ordinary differential equation
PDE	Partial differential equation
POD	Proper orthogonal decomposition
RB	Reduced basis
RBF	Radial basis function
ROM	Reduced order modeling
SVD	Singular value decomposition
TD	Tensor decomposition

Greek Symbols

$\{\alpha_n\}_{n=1}^{N}$	Modal/reduced coefficients
Γ	Boundary of the spatial domain
ε	Relative error
ε_{POD}	POD projection error lower bound
η	Fluid electrical resistivity (Ω·m)
$\{\eta_n\}_{n=1}^{N_s}$	POD eigenvectors
θ	GP kernel hyper-parameters vector
$\kappa_\theta(\mathbf{x}, \mathbf{x}')$	Kernel (covariance function) of the prior distribution of f
$\Lambda^{(n)}$	Singular values matrix
$\{\lambda\}_{n=1}^{N_s}$	POD eigenvalues
$\boldsymbol{\mu}$	Parameters vector
μ	Fluid viscosity (Pa·s)
μ_0	Magnetic permeability (H/m)
Ξ^{test}	Test set
Ξ^{train}	Training set
ρ	Fluid density (kg/m^3)
σ^2	Sample variance
$\phi^{(n)}$	Parametric modes
$\{\varphi_n\}_{n=1}^{N}$	Basis functions
$\psi^{(n)}$	Time modes
Ω	Spatial domain

Latin Symbols

$\mathbb{A}_n$	Matrix for the nth coefficient α_n
$\mathbf{B}$	Magnetic field ($\mathscr{T}$)
$\mathscr{B}$	Magnetic field parametric domain
$\mathbb{C}$	POD correlation matrix
c	Constant for wave function initialization
$\mathscr{D}$	Training set
d	Size of the GRP training set $\mathscr{D}$
E	Average absolute error
$\mathbb{E}(\mathbf{x})$	Mean of the prior distribution of f
f	Generic (to-be-inferred) input–output relationship
h	Step height (m)
I	Relative information content
l	Truncation rank
M	Dimension of the GPR test set
$\mathscr{N}_h$	Dimension of the spatial grid
N	Dimension of the reduced space (*rank*)
N_s	Number of available training solutions
N_t	Number of sampled time instants
P	Conditional density function of y
p	Fluid pressure (Pa)
r	Residual field
Re	Fluid Reynolds number
$\mathbb{S}$	Snapshot matrix

$\mathscr{T}$	Time parametric domain
$\mathscr{U}$	Solution manifold
u	Training solutions (*snapshots*)
$\mathbf{u}$	Fluid velocity (m/s)
X_N	Reduced space
$\mathbb{X}$	Input training set
$(\mathbf{x}_i, y_i)$	Input-output pair
$\mathbf{y}$	Output training set

REFERENCES

[1]J. P. Freidberg, *Ideal MHD* (Cambridge University Press, 2014).
[2]H. Zohm, G. Gantenbein, A. Isayama, A. Keller, R. L. Haye, M. Maraschek, A. Mück, K. Nagasaki, S. Pinches, and E. Strait, "Mhd limits to tokamak operation and their control," Plasma Phys. Controlled Fusion **45**, A163 (2003).
[3]L. Bühler, "Liquid metal magnetohydrodynamics for fusion blankets," in *Magnetohydrodynamics*, Fluid Mechanics And Its Applications Vol. 80 (Springer, Dordrecht, 2007).
[4]G. Ferrero, R. Testoni, and M. Zucchetti, "Impact assessment of radiative heat transport in arc-class reactor FLiBe liquid immersion blanket," Nucl. Sci. Eng. **198**, 898–816 (2024).
[5]D. Biskamp and D. Biskamp, *Nonlinear Magnetohydrodynamics* (Cambridge University Press, 1997), Vol. 1.
[6]T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza, "Model order reduction in fluid dynamics: Challenges and perspectives," in *Reduced Order Methods for Modeling and Computational Reduction* (Springer International Publishing, Cham, 2014), pp. 235–273.
[7]G. Rozza, M. Hess, G. Stabile, M. Tezzele, F. Ballarin, C. Gräßle, M. Hinze, S. Volkwein, F. Chinesta, P. Ladeveze, Y. Maday, A. Patera, and J. Farhat Char, *Model Order Reduction: Volume 2: Snapshot-Based Methods and Algorithms* (De Gruyter, 2020).
[8]A. Quarteroni, A. Manzoni, and F. Negri, *Reduced Basis Methods for Partial Differential Equations: An Introduction*, 1st ed., UNITEXT (Springer Cham, 2015).
[9]S. Lorenzi, A. Cammi, L. Luzzi, and G. Rozza, "POD-Galerkin method for finite volume approximation of Navier–Stokes and RANS equations," Comput. Methods Appl. Mech. Eng. **311**, 151–179 (2016).
[10]M. Tezzele, N. Demo, A. Mola, and G. Rozza, "An integrated data-driven computational pipeline with model order reduction for industrial and applied mathematics," in *Novel Mathematics Inspired by Industrial Challenges* (Springer International Publishing, Cham, 2022), pp. 179–200.
[11]A. Charalampopoulos and T. Sapsis, "Uncertainty quantification of turbulent systems via physically consistent and data-informed reduced-order models," Phys. Fluids **34**, 075120 (2022).
[12]A. Moni, W. Yao, and H. Malekmohamadi, "Data-driven reduced-order modeling for nonlinear aerodynamics using an autoencoder neural network," Phys. Fluids **36**, 016105 (2024).
[13]R. Taylor, J. N. Kutz, K. Morgan, and B. A. Nelson, "Dynamic mode decomposition for plasma diagnostics and validation," Rev. Sci. Instrum. **89**, 053501 (2018).
[14]A. A. Kaptanoglu, K. D. Morgan, C. J. Hansen, and S. L. Brunton, "Characterizing magnetized plasmas with dynamic mode decomposition," Phys. Plasmas **27**, 032108 (2020).
[15]A. A. Kaptanoglu, K. D. Morgan, C. J. Hansen, and S. L. Brunton, "Physics-constrained, low-dimensional models for magnetohydrodynamics: First-principles and data-driven approaches," Phys. Rev. E **104**, 015206 (2021).
[16]S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," Proc. Nat. Acad. Sci. U.S.A. **113**, 3932–3937 (2016).
[17]Y. Yang, Y. Xue, W. Zhao, S. Yao, C. Li, and C. Wu, "Fast flow field prediction of three-dimensional hypersonic vehicles using an improved Gaussian process regression algorithm," Phys. Fluids **36**, 016129 (2024).
[18]G. Ortali, N. Demo, and G. Rozza, "A Gaussian process regression approach within a data-driven POD framework for engineering problems in fluid dynamics," Math. Eng. **4**, 1–16 (2021).
[19]L. Cicci, S. Fresca, M. Guo, A. Manzoni, and P. Zunino, "Uncertainty quantification for nonlinear solid mechanics using reduced order models with Gaussian process regression," Comput. Math. Appl. **149**, 1–23 (2023).
[20]Y. Maday, O. Mula, A. T. Patera, and M. Yano, "The generalized empirical interpolation method: Stability theory on Hilbert spaces with an application to the Stokes equation," Comput. Methods Appl. Mech. Eng. **287**, 310–334 (2015).
[21]A. Cammi, S. Riva, C. Introini, L. Loi, and E. Padovani, "Indirect field reconstruction and sensor positioning in circulating fuel reactors using data-driven model order reduction," in *Proceeding of the* 2023 International Congress on Advances in Nuclear Power Plants, Gyeongju, Korea (2023), pp. 1–12.
[22]M. Guo and J. S. Hesthaven, "Reduced order modeling for nonlinear structural analysis using Gaussian process regression," Comput. Methods Appl. Mech. Eng. **341**, 807–826 (2018).
[23]M. Guo and J. S. Hesthaven, "Data-driven reduced order modeling for time-dependent problems," Comput. Methods Appl. Mech. Eng. **345**, 75–99 (2019).
[24]D. Xiao *et al.*, "Non-intrusive reduced order models and their applications," Ph.D. thesis (Imperial College London, 2016).
[25]D. Xiao, F. Fang, C. Pain, and I. Navon, "A parameterized non-intrusive reduced order model and error analysis for general time-dependent nonlinear partial differential equations and its applications," Comput. Methods Appl. Mech. Eng. **317**, 868–889 (2017).
[26]S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, 2nd ed. (Cambridge University Press, USA, 2022).
[27]C. Williams and C. Rasmussen, "Gaussian processes for regression," in *Advances in Neural Information Processing Systems*, edited by D. Touretzky, M. C. Mozer, and M. Hasselmo (MIT Press, 1995), Vol. 8, pp. 1–7.
[28]C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning.*, Adaptive Computation and Machine Learning (MIT Press, 2006).
[29]A. Cammi, S. Riva, C. Introini, L. Loi, and E. Padovani, "Data-driven model order reduction for sensor positioning and indirect reconstruction with noisy data: Application to a Circulating Fuel Reactor," Nucl. Eng. Des. **421**, 113105 (2024).
[30]S. Riva, C. Introini, and A. Cammi, "Multi-physics model bias correction with data-driven reduced order techniques: Application to nuclear case studies," Appl. Math. Modell. **135**, 243–268 (2024).
[31]M. W. Scroggs, I. A. Baratta, C. N. Richardson, and G. N. Wells, "Basix: A runtime finite element basis evaluation library," J. Open Source Softw. **7**, 3982 (2022).
[32]M. S. Alnaes, A. Logg, K. B. Ølgaard, M. E. Rognes, and G. N. Wells, "Unified form language: A domain-specific language for weak formulations of partial differential equations," ACM Trans. Math. Softw. **40**, 1 (2014).
[33]I. A. Baratta, J. P. Dean, J. S. Dokken, M. Habera, J. S. Hale, C. N. Richardson, M. E. Rognes, M. W. Scroggs, N. Sime, and G. N. Wells (2023). "DOLFINx: The next generation FEniCS problem solving environment," Zenodo. http://www.doi.org/10.5281/zenodo.10447666
[34]M. W. Scroggs, J. S. Dokken, C. N. Richardson, and G. N. Wells, "Construction of arbitrary order finite element degree-of-freedom maps on polygonal and polyhedral cell meshes," ACM Trans. Math. Softw. **48**, 1 (2022).
[35]L. Boccaccini, G. Aiello, J. Aubert, C. Bachmann, T. Barrett, A. Del Nevo, D. Demange, L. Forest, F. Hernandez, P. Norajitra *et al.*, "Objectives and status of eurofusion demo blanket studies," Fusion Eng. Des. **109–111**, 1199–1206 (2016).
[36]U. Müller and L. Bühler, *Magnetofluiddynamics in Channels and Containers* (Springer Science & Business Media, 2001).
[37]H. G. Weller, G. Tabor, H. Jasak, and C. Fureby, "A tensorial approach to computational continuum mechanics using object-oriented techniques," Comput. Phys. **12**, 620–631 (1998). publisher: AIP.
[38]M. Lo Verso, C. Introini, E. Cervi, F. Giacobbo, and A. Cammi, "A novel OpenFOAM library for magneto-hydrodynamics studies in the nuclear fusion field," in Proceedings of the NUTHOS 2024 Conference, 2024.
[39]C. Introini, S. Riva, S. Lorenzi, S. Cavalleri, and A. Cammi, "Non-intrusive system state reconstruction from indirect measurements: A novel approach based on hybrid data assimilation methods," Ann. Nucl. Energy **182**, 109538 (2023).
[40]B. Armaly, F. Durst, J. Pereira, and B. Schönung, "Experimental and theoretical investigation of backward-facing step flow," J. Fluid Mech. **127**, 473–496 (1983).
[41]S. Scharnowski, I. Bolgar, and C. Kähler, "Control of the recirculation region of a transonic backward-facing step flow using circular lobes," in Proceedings of

the Ninth International Symposium on Turbulence and Shear Flow Phenomena (TSFP-9) (2015), pp. 1–6.
[42]P. M. Nadge and R. N. Govardhan, "High Reynolds number flow over a backward-facing step: Structure of the mean separation bubble," Exp. Fluids **55**, 1657 (2014).
[43]D. Lee and H. Choi, "Magnetohydrodynamic turbulent flow in a channel at low magnetic Reynolds number," J. Fluid Mech. **439**, 367–394 (2001).
[44]P. Davidson, *Turbulence: An Introduction for Scientists and Engineers* (Oxford University Press, 2015).
[45]L. Sirovich, "Turbulence and the dynamics of coherent structures part II: Symmetries and transformations," Quart. Appl. Math. **45**, 573–582 (1987).
[46]R. Everson and L. Sirovich, "Karhunen–Loève procedure for gappy data," J. Opt. Soc. Am. A **12**, 1657 (1995).