POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

State Estimation in Magnetohydrodynamics: from interpolation to recurrent networks

Laurea Magistrale in Nuclear Engineering - Ingegneria Nucleare

Author: Claudio Scardino

Advisor: Prof. Antonio Cammi

Co-advisors: Dr. Stefano Riva, Dr. Carolina Introini

Academic year: 2025-2026

1. Introduction

This thesis work aims at studying the application of Data-Driven Reduced Order Models (DDROMs) within a Data Assimilation framework to estimate the temperature, pressure, and velocity fields characterizing a Magneto-Hydrodynamic (MHD) system for real-time monitoring purposes [2]. Due to the strong nonlinearities and multiphysics nature of the governing equations, high-fidelity solutions are computationally expensive and highly sensitive to input parameters. Accordingly, full-order simulations are not a feasible strategy for performing online monitoring of the different physical fields of an MHD system [4]. In this context, DDROMs are a promising alternative for simulating specific scenarios, based on the knowledge of high-fidelity solutions for other parameter combinations and online measurements coming from sparse sensors, as the computational cost and time are drastically reduced [1].
Different DDROM methods, ranging from machine learning algorithms to neural network–based architectures, are introduced and compared on a first two-dimensional case study in order to determine whether a particular technique is more accurate in performing state estimation than the others. Different sensor configurations are investigated, focusing either on multiple sensors freely distributed in the domain or on a sparse number of sensors constrained to the boundaries [2], as sensor placement in the bulk of the flow is often practically difficult due to harsh environments. In particular, the comparison has been conducted by considering different state estimation scenarios, such as time reconstruction mode, time-prediction mode, and parametric mode, in which the full temporal evolution of the state is computed for a new, unseen value of the input parameter, namely the magnetic field intensity. Once the SHRED has been determined to be the most accurate technique, it is applied to a more engineering-oriented three-dimensional case study, in which the flow of a fluid is hindered by two transverse cylinders acting as tubes of a heat exchanger to cool down the liquid metal [4]. For this test case, the parameters considered are the intensity of the magnetic field and its orientation, both individually and in combination.

Magneto-hydrodynamics describes the coupling between the electromagnetic field and a physical

system that is able to move in space, such as a fluid, and to carry electrical current as a conductor [4]; liquid metal immersed in an external magnetic field is a perfect example of an MHD system, as it is electrically conductive and in a fluid state at sufficiently high temperature.
The high-fidelity simulations have been conducted using a transient solver in OpenFOAM, called *magnetoHDFoam*[1], for a pseudo-compressible, buoyant, turbulent flow of a electrically conducting fluid, under the influence of a magnetic field using the PIMPLE (PISO-SIMPLE) solution for time-resolved and pseudo-transient simulations.

2. Data Driven Reduced Order modeling

Reduced order models are surrogate models of the Full Order Models (FOM), in which the physical fields are represented in a latent, low-dimensional mathematical space that approximates the high-dimensional physical space [1]. Data driven reduced order modeling builds the reduced space X_N based on different snapshots (i.e. specific solutions of the FOM, parametrized by a set of parameters $\boldsymbol{\mu}$) belonging to the solution manifolds $u \in \mathcal{U}$. In particular, starting from a set of full order solutions, N basis functions ϕ_i, depending only on the physical space $\mathbf{x}$, are determined, and the reduced space, also called *latent space*, is obtained as the *span* of the same basis functions, i.e., $X_N = \text{span}\{\phi_i\}_{i=1}^N$. Once the reduced space has been built, the snapshots and, in general, all the solutions belonging to the solution manifold can be projected onto the latent space as a linear combination of the basis functions weighted by the so-called *reduced coefficients* $\{\alpha_i\}_{i=1}^N$, which depend only on the set of parameters [1]:

$$u(\mathbf{x}; \boldsymbol{\mu}) \simeq u_N(\mathbf{x}; \boldsymbol{\mu}) = \sum_{i=1}^{N} \alpha_i(\boldsymbol{\mu})\phi_i(\mathbf{x}). \tag{1}$$

In particular, the basis functions encode the most relevant spatial patterns, while the reduced coefficients encode the parametric dependence. Once the basis functions are known, it is sufficient to compute the reduced coefficients in order to perform the state estimation. Data Assimilation (DA) is the mathematical connection between the physical model, i.e. the set of PDEs describing the physical phenomenon, and the real world context characterized by empirical measurements $\mathbf{y}$ coming from sparse sensors. In the DDROM framework, the problem is shifted to the determination, starting from these measurements, of the *latent dynamic* and hence of the reduced coefficients $\alpha_i(\boldsymbol{\mu})\ \forall i \in \{1, \cdots, N\}$, once the reduced space (i.e. the span of the basis functions $\phi_i(\mathbf{x})$) has been determined, together with the mathematical map $\mathcal{M}$ between the latent dynamic and measurements for the *training* snapshots. In particular, DDROM in combination with DA is composed of two sequential phases:

- *Offline (training) phase*: starting from the training snapshots, the reduced basis is built, and the map $\mathcal{M}$ between the training measurements and the latent dynamics obtained by projecting the training snapshots onto the reduced basis is established.
- *Online (test) phase*: state estimation from new available data coming from sensors is performed by simply exploiting the learned map $\mathcal{M}$, applying it to these new measurements, $\mathcal{M}(\mathbf{y}) = \{\alpha_i(\boldsymbol{\mu})\}_{i=1}^N$, thus determining the new latent dynamics and, finally, performing the linear combination of the basis functions with the reduced coefficients.

In this thesis work, two different reduced bases are investigated: the one built in the *Generalized Empirical Interpolation Method* (GEIM) [3] and the one built according to the *Singular Value Decomposition* (SVD) [1]. GEIM is the DDROM technique used in the thesis as the state of the art among linear methods.

2.1. SHallow REcurrent Decoder (SHRED)

SHallow REcurrent Decoder (SHRED) is a Neural Network (NN) composed of a *Long Short Term Memory* (LSTM) network and a *Shallow Decoder* network (SDN) [5]. The LSTM network, composed of two hidden layers, processes temporal measurements that have been pre-handled according to the *Time-Lagged-Embedding* (TLE) procedure. TLE consists of grouping sequential temporal measurements into mini-batches of dimensions equal to the *lag* hy-

[1] Available at `https://github.com/ERMETE-Lab/MHD-magnetoHDFoam?tab=readme-ov-file`

perparameter, L, such that the estimation of the POD coefficients at time t depends on the sensors history up to time $t - L\Delta t$, where Δt is the sampling interval. The Shallow Decoder Network is a feedforward neural network composed of two hidden layers that maps the latent space produced by the LSTM onto the r POD coefficients at time t. A key feature of SHRED is its ability to estimate the POD coefficients of all fields using measurements from a single observable field, i.e., enabling indirect state estimation. The Universal Approximation Theorem for Operators and Takens' embedding theorem are the two theoretical pillars on which POD-SHRED is based. Moreover, following the rationale of separation of variables, instead of knowing the initial conditions over the entire domain, the entire temporal history is known at a few points corresponding to the sensor positions; accordingly, SHRED is agnostic to the sensor positions. POD-SHRED is a generalization of the POD-DeepONet as POD-SHRED encodes the spatial patterns through the SVD basis, substituting the *trunk* network of DeepONet, and encodes the temporal and parametric dynamics with the LSTM and a shallow decoder architecture, substituting the *branch* network.

3. Comparison between DDROM methods on a 2D case study

For the first two-dimensional benchmark case, a channel with obstacles on the top and bottom walls acting as a heat exchanger, with the magnetic field intensity as the only external parameter, GEIM, TR-GEIM, POD-DeepONet, and POD-SHRED are compared in terms of state estimation accuracy for all combinations of sensor number and placement (free or boundary-constrained). In particular, the possible combinations are:

- Multiple-free sensors configuration: 20 sensors available from the entire spatial domain.
- Multiple-constrained sensors configuration: 20 sensors available from specific regions of the spatial domain.
- Three-free sensors configuration: only 3 sensors available from the entire spatial domain.
- Three-constrained sensors configuration: only 3 sensors available from specific regions of the spatial domain.

Regardless of the configuration, regarding the POD-SHRED and POD-DeepONet indirect reconstruction is always performed, as all the fields are estimating only measuring the temperature one, while for the GEIM and TRGEIM, only direct estimation is feasible and hence each field is measured. The mean relative error (mre) has been defined as the performance metric for a given ROM method under a specific sensor configuration, expressed as follows:

$$mre(\boldsymbol{u}_j^\star) = \left\langle \frac{\|\boldsymbol{r}_j(t, B^\star)\|_2}{\|\boldsymbol{u}_j(t, B^\star)\|_2} \right\rangle_{t, B^\star \in \Xi^B_{test}} \qquad (2)$$

where $\boldsymbol{u}_j^\star$ is the state estimation for the j-th field (e.g. velocity, temperature and pressure), $\boldsymbol{r}_j$ is the absolute residual between the full order solution and the reduced solution and $\langle\cdot\rangle_{t,B^\star\in\Xi^B_{test}}$ represents the mean over test time and parameters.

3.1. Temporal reconstruction mode

In this scenario, once the value of the external magnetic field as $B = 0.01$ T has been fixed, the only parameter for which the state has to be reconstructed is the time. In this context, the capabilities of different DDROM methods to estimate the different fields for a time instant unseen in the training phase, is tested in interpolation regime. Analyzing the different mean relative errors, the following conclusions can be stated. Neural network-based ROM techniques perform remarkably better than GEIM and TR-GEIM, regardless of the number of sensors and their configuration. In particular, the mean relative error committed by GEIM using only three sensors is more than five times that committed by POD-SHRED in estimating the velocity field ($\sim 6\%$ vs. $\sim 30\%$). This difference becomes even larger when considering the pressure and temperature fields, for which the SHRED error is almost ten times lower than that of GEIM ($\sim 2\%$ vs, $\sim 20\%$ on average on different configurations). The lowest possible error is achieved by SHRED in the state estimation of the temperature field $mre = 1\%$ in the multisensor-free configuration, since this is the exact field on which the SHRED neural network was trained, performing indeed direct state estimation rather than indirect state estimations for the other two fields. Regarding the regularization of TR-GEIM compared to GEIM, it can be correctly noted that, since the standard deviation of the noise distribution $\sigma = 0.05$, is not negligible, the error committed by TR-GEIM is lower than that of GEIM ($\sim 25\%$ vs, $\sim 30\%$ on average on different configurations for the velocity field). Most importantly, the superiority of SHRED compared to DeepONet is evident, as the mean relative error for SHRED is half of that for DeepONet ($\sim 3\%$ vs, $\sim 7\%$ on average on different configurations for the pressure field).

3.2. Temporal Prediction mode

In this scenario, the only parameter considered is again the time, and prediction mode consists of the state estimation of test snapshots for time instants $t^\star$ never seen in the training phase such that $t^\star > t_{N_t^{train}}$, where $t_{N_t^{train}}$ is the last time instant chronologically simulated in the training set. In this context, neural network–based ROM techniques perform only slightly better than GEIM and TR-GEIM, and, in general, the state estimation task is far more difficult, leading to higher average relative errors (5% for temperature, 20% for velocity, and 10% for pressure). In particular, it should be mentioned that, for the pressure field, GEIM performs slightly better than SHRED (9% vs 11% on average across different configurations). This trend can be explained by the relative sensor positions. Indeed, since GEIM (and TR-GEIM) sensors are chosen to be optimal for the interpolation problem, the state estimation for the pressure field may be more sensitive to sensor positions, as pressure is a more uniform field and hence more easily interpolated. This may be due to the fact that, in contrast, SHRED measures only the temperature field, which is more chaotic than the pressure field, and therefore performs indirect reconstruction, leading to slightly higher errors.

3.3. Parametric mode

Parametric mode consists of estimating the entire temporal evolution of the state for test snapshots corresponding to values of the magnetic field intensity B_{test} that were never seen during the training phase. Investigating the mean relative error it can be stated that the POD-SHRED based state estimation is significantly more accurate than those based on interpolation strategies, such as GEIM and TR-GEIM (average error of 10% vs 20% for the velocity field, 2% vs.10% for the pressure field, 3% vs 9% for the temperature field). Moreover, SHRED performs better than DeepONet, although the margin is not as large as observed in the reconstruction mode (on average for the POD-DeepONet the mre is around 12% for the velocity, 4% for the pressure and 4% for the temperature). As a general remark, multiple sensors freely distributed in the physical domain lead to the lowest mean relative error, since more information is available to learn the correlation between temporal measurements and the latent dynamics. Nevertheless, boundary-constrained sensors for SHRED result in only slightly worse reconstruction, proving SHRED's agnosticism to sensor positioning. Concluding the comparison, POD-SHRED results as the most accurate DDROM technique, and thanks to the possibility of using different state estimations obtained from different positions of only three sensors for each configuration to compute an ensemble reconstruction, with an associated standard deviation field that estimates the reconstruction residual, SHRED is selected as the technique to be used in the more engineering-oriented second case study.

3.4. SHRED as an Operator Learning framework

A detailed analysis of the mean relative error in the parametric mode for SHRED, as a function of the number of sensors employed, the lag hyperparameter, and the measured scalar field (temperature or pressure), has been conducted, and important conclusions can be drawn. With reference to Fig. 1, where the mean relative error (%) for pressure state estimation is shown as a function of the number of sensors and their typology (i.e., what they actually measure), as expected, increasing the number of sensors leads to a decreasing error trend until it saturates.

The unexpected result is that the saturation error for pressure estimation is lower when the measured scalar field is temperature rather than pressure. Moreover, the decreasing trend is more regular for temperature measurements than for pressure measurements, for which an oscillatory

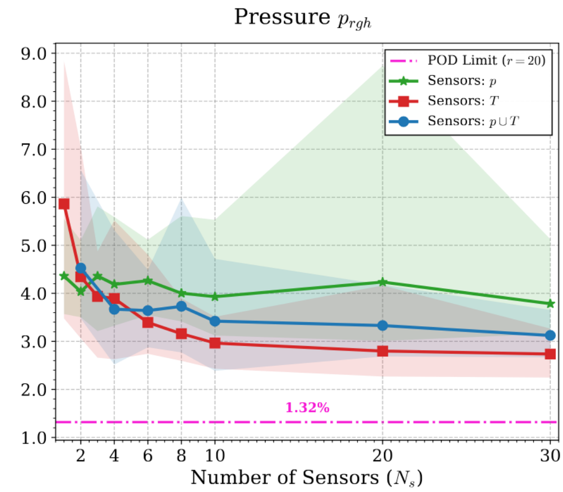

Figure 1: Mean relative error (%) for pressure state estimation.

behavior can be observed: the error remains almost constant as a function of the number of sensors, such that the mean relative error is nearly the same whether one sensor or 30 sensors are used. All these considerations lead to the conclusion that SHRED is not merely a tool for fitting data, but that it actually captures the underlying physics of the MHD equations as implemented in magnetoHDFoam. Indeed, temperature is correctly recognized as the most important variable to understand the flow dynamics. The neural network understands the causality between the fields (a change in temperature leading to a change in density, which leads to a change in pressure through the Poisson equation), mapping the cause, a change in temperature to the effect, a change in the pressure field. In this sense, by measuring T to reconstruct p_{rgh}, SHRED is learning the *inverse operator* of the Laplacian $(\nabla^2)^{-1}$, which physically links pressure as the response to temperature as the source through density variations.
Regarding the decreasing trend of the curves, SHRED recognizes the elliptic nature of the Poisson equation when pressure measurements are used, as the mean relative error does not depend on the number of sensors: even a single point, according to the convolution with Green's function (the solution of the Poisson equation), encodes information about the entire spatial pressure field. In contrast, for temperature measurements, SHRED recognizes the advection–diffusion nature of the energy transport equation, and hence the error decreases as the number of sensors increases. Moreover, SHRED is energy-consistent in its predictions: by analyzing the Pearson correlation coefficients between the POD coefficients of the temperature field and the velocity field, an anticorrelation is observed, such that when the velocity, and hence the kinetic energy, is overestimated, the internal energy, and thus the temperature, is underestimated. In this way, even when the neural network makes prediction errors, the total energy is nevertheless conserved.

4. Single and Double Parametric Analysis

The second, more engineering-oriented benchmark case is a three-dimensional duct in which the flow of liquid metal (Pb–17Li) is hindered by the presence of two transversal cylinders that act as a heat exchanger. This test case simulates the dynamics of the liquid metal used inside the blanket of a fusion reactor to dissipate the heat generated by fusion reactions, and it is strongly influenced by the external magnetic field produced by the magnets used for magnetic confinement.
The POD-SHRED's capability to infer the dynamics of the flow when a single parameter is varied (either the inclination angle or the intensity of the external magnetic field), or when both are varied in combination, is tested. According to the dominant toroidal component of the magnetic field with respect to the poloidal one, the flow dynamics are simulated for inclination angles spanning from 5° up to 30° with respect to the transversal plane. Regarding the intensity of the magnetic field, values range from low magnitudes (~ 0.02 T), which allow turbulent flow regimes, to high magnitudes (~ 0.350 T), for which the magnetic braking and laminarization effects are observed.
The results of the mean relative error show remarkable accuracy in the reconstruction for both the single and double parametric analyses. In particular, regarding state estimation when a single parameter is considered, lower errors are obtained in the reconstruction of all physical fields when the parameter is the inclination angle ($mre < 3\%$) rather than when the parameter considered is the intensity of the magnetic field ($mre < 6\%$). This behavior is consistent with the difficulty of the reconstruction task, which is

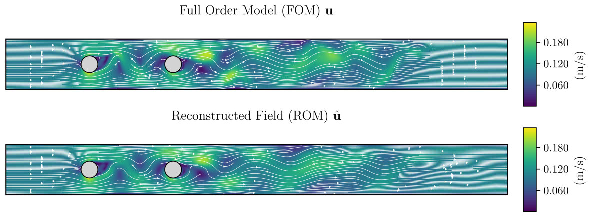

Figure 2: Full Order vs. SHRED reconstructed velocity field for $\alpha = 5°$, $B = 0.05$ T at $t = 1.5$ s.

embedded in the dependence of the solution on the specific parameter. In particular, changing the inclination angle does not change the energy introduced into the system by the magnetic field; only the spatial distribution is modified. Therefore, since the studied angle range is limited, the dependence of the solution on the angle is well encoded in the data, and SHRED reconstructs it accurately. Regarding the magnetic field intensity as a parameter, the energy injected into the system changes according to its value, and a change in the flow regime is expected; this makes the task more difficult for the neural network, since the dependence of the solution on the parameter is more pronounced: according to the non-linearity of the MHD equations, even small changes in the parameter value can lead to large changes in the solution, as happens, for example, for a magnetic field intensity of 0.1 T, where a transitional regime from turbulent to laminar is observed.

In a pioneering context, such as double-parametric state estimation for an MHD problem, SHRED demonstrates notably high accuracy. Except for some pathological cases, i.e., specific combinations of magnetic field intensity and inclination angle where the dynamics are not well encoded by the training data, the mean relative error across all fields is, on average, around 5%. Nevertheless, even for a pathological case, for example $B = 0.3$ T and inclination angle $\alpha = 5°$, the highest mean relative error, which occurs for the velocity field, is around 10%, indicating that the overall flow behavior is still successfully captured.

5. Conclusions

In this thesis, different DDROM methods have been applied and compared based on their accuracy in state estimation for a first two-dimensional MHD test case, establishing the overall superiority of POD-SHRED across all configurations and operational modes. The analysis of the decreasing trend of the mean relative error as a function of the number of sensors, depending on whether the measured field is pressure or temperature, shows an actual learning capability of SHRED, as the structure of the advection–diffusion equations for temperature and the ellipticity of the Poisson equation for pressure are recognized, and as SHRED effectively learns to approximate the inverse Laplacian operator to link the temperature measurements to the pressure field.

In conclusion, when applied to a more engineering oriented test case, SHRED proves to be a remarkably accurate tool to estimate the state of a highly non-linear system, such as an MHD problem, even when a double-parametric analysis is performed.

The final, appealing goal is to build a direct digital twin of the experimental facilities and, based on this, exploit experimental measurements within a closed-loop feedback control strategy, where a reinforcement learning algorithm dynamically adjusts the magnetic field to obtain a desired system response.

References

[1] Steven L Brunton and J Nathan Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, USA, 2nd edition, 2022.

[2] Antonio Cammi, Stefano Riva, Carolina Introini, Lorenzo Loi, and Enrico Padovani. Data-driven model order reduction for sensor positioning and indirect reconstruction with noisy data: Application to a circulating fuel reactor. *Nuclear Engineering and Design*, 421:113105, 2024.

[3] Yvon Maday and Olga Mula. A Generalized Empirical Interpolation Method: Application of Reduced Basis Techniques to Data Assimilation. *Springer INdAM Series*, pages 221–235, 2013. ISBN: 9788847025929 Publisher: Springer Milan.

[4] S. Molokov, R. Moreau, K. Moffatt, and L. Bühler. Liquid metal magnetohydrodynamics for fusion blankets. In *Magnetohydrodynamics: Historical Evolution and Trends*, pages 171–194. Springer Dordrecht, 2007.

[5] Jan P. Williams, Olivia Zahn, and J. Nathan Kutz. Sensing with shallow recurrent decoder networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 480(2298):20240054, 2024.