



MOX-Report No. 43/2025

**Real-time optimal control of high-dimensional parametrized systems  
by deep-learning based reduced order models**

Tomasetto, M.; Manzoni, A.; Braghin, F.

MOX, Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

[mox-dmat@polimi.it](mailto:mox-dmat@polimi.it)

<https://mox.polimi.it>

# Real-time optimal control of high-dimensional parametrized systems by deep learning-based reduced order models

Matteo Tomasetto<sup>a</sup>, Andrea Manzoni<sup>b</sup>, Francesco Braghin<sup>a</sup>

<sup>a</sup>*Department of Mechanical Engineering, Politecnico di Milano, Via privata Giuseppe La Masa 1, 20156, Milano, Italy*

<sup>b</sup>*MOX – Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano, Italy*

---

## Abstract

Steering a system towards a desired target in a very short amount of time is a challenging task from a computational standpoint. Indeed, the intrinsically iterative nature of optimal control problems requires multiple simulations of the state of the physical system to be controlled. Moreover, the control action needs to be updated whenever the underlying scenario undergoes variations, as it often happens in applications. Full-order models based on, e.g., the Finite Element Method, do not meet these requirements due to the computational burden they usually entail. On the other hand, conventional reduced order modeling techniques such as the Reduced Basis method, despite their rigorous construction, are intrusive, rely on a linear superimposition of modes, and lack of efficiency when addressing nonlinear time-dependent dynamics. In this work, we propose a non-intrusive Deep Learning-based Reduced Order Modeling (DL-ROM) technique for the rapid control of systems described in terms of parametrized PDEs in multiple scenarios. In particular, optimal full-order snapshots are generated and properly reduced by either Proper Orthogonal Decomposition or deep autoencoders (or a combination thereof) while feedforward neural networks are exploited to learn the map from scenario parameters to reduced optimal solutions. Nonlinear dimensionality reduction therefore allows us to consider state variables and control actions that are both low-dimensional and distributed. After (i) data generation, (ii) dimensionality reduction, and (iii) neural networks training in the offline phase, optimal control strategies can be rapidly retrieved in an online phase for any scenario of interest. The computational speedup and the extremely high accuracy obtained with the proposed approach are finally assessed on different PDE-constrained optimization problems, ranging from the minimization of energy dissipation in incompressible flows modelled through Navier-Stokes equations to the thermal active cooling in heat transfer.

*Keywords:* optimal control, PDE-constrained optimization, parametrized systems, reduced order modeling, deep learning, fluid flows, heat transfer

---

## 1. Introduction

Controlling dynamical systems is one of the most challenging and widespread tasks in Applied Sciences and Engineering. For instance, reducing energy consumption, ensuring system stability, or attaining a desired configuration, are just three examples of problems that can be cast in the framework of Optimal Control Problems (OCPs). In order to steer the dynamics of a physical system towards a desired target, a suitable control action has to be properly selected and tuned. The optimal control strategy is typically found as a minimum point of a *loss* or *cost functional*. Minimization has to be constrained by the governing equation driving the considered dynamics in order to narrow the space of possible solutions on those that are physically admissible [44]. The governing equations may be, in general, algebraic systems, ordinary differential equations or, as considered in the following, Partial Differential Equations (PDEs). Specifically, we take into account PDEs parametrized by a vector of input (or scenario) parameters  $\boldsymbol{\mu}_s$ . Indeed, our main goal is to conceive an efficient numerical strategy to determine optimal controls in multiple scenarios of interest as fast as possible. OCPs governed by PDEs are traditionally solved by exploiting high-fidelity full-order models (FOMs) based on discretization techniques such as, e.g., the Finite Element Method (FEM). Despite their ubiquitous use – from fluid flows, backs to the 90s [26, 30, 25] to more recent applications in robotic swarms [68, 67, 69, 66] and acoustic and thermal cloaking [16, 14, 70, 63] – relying on FOMs becomes

unfeasible in the case of multiple scenarios, since an iterative optimization procedure and thus several PDE solves would be required for each new scenario.

Reduced Order Modeling (ROM) techniques [28, 54, 10, 43], such as the Reduced Basis (RB) method relying on Proper Orthogonal Decomposition (POD), have been largely exploited to reduce the dimensionality of optimality equations and thus enable faster resolutions [38, 39, 49, 5, 2, 70]. In the context of Model Predictive Control (MPC), POD and system identification techniques, such as Dynamic Mode Decomposition (DMD) [64], Sparse Identification of Nonlinear Dynamics (SINDy) [12, 13] or Eigensystem Realization Algorithm (ERA), have been employed to replace the physical constraint by a low-rank surrogate model [34, 29, 52]. Similar strategies are exploited to describe the dynamics of usually low-dimensional observables measured by sensors, that is, an approximation of the Koopman operator [51, 36, 37]. However, the linear and intrusive nature of the RB method does not make it suitable, or computationally affordable, whenever (i) the finite element matrices are not accessible, (ii) if the physical laws are nonlinear, or (iii) if the problems are not affine with respect to the scenario parameters.

In addition to ROM strategies, neural networks have been more recently considered in the literature to speed up the resolution of OCPs. For instance, in the MPC framework, the differential equations appearing as constraints are often approximated through neural network-based surrogate models [19, 15, 7, 50, 3], such as Recurrent Neural Networks, Physics-Informed Neural Networks or Input Convex Neural Networks, which replicate the dynamics of the available data, achieving a remarkable computational speedup. However, since no reduction techniques have been systematically applied in these cases, only low-dimensional states and controls can be taken into account to deal with feasible input and output neural networks dimensions.

To overcome the limitations of the previous ROM frameworks and thus allowing for greater flexibility and speedup, in this work we exploit non-intrusive ROM strategies to develop a framework able to solve any high-dimensional parametrized OCP in real-time for multiple scenarios. Differently from the RB method, non-intrusive techniques compress the available data through POD [27], deep convolutional autoencoders [21, 20], or a combination thereof [22], while the map going from scenarios to the reduced optimal solution can be modeled through a deep feedforward neural network. A combination of linear dimensionality reduction and neural networks has also been considered by [41] in the context of robust control, while convolutional autoencoders are taken into account, for instance, by [32] to reduce the dimensionality of uncontrolled state snapshots. In this work, instead, we properly reduce also high-dimensional control variables and we focus on controlled snapshots in order to avoid costly optimization procedures when computing optimal actions related to new scenarios.

OCPs have been increasingly addressed with Physics-Informed Neural Networks (PINNs) [57] modeling the space-time dependence of state variables. In addition to the misfit between model outputs and experimental data, PINN also encode the residuals of the physical equations in the loss function to be minimized as regularizing terms to achieve more consistent results. For example, [75, 31] solve OCPs approximating the control-to-state map through Physics-Informed Deep Operator Networks (PI-DeepONets [76, 40]). Moreover, all-at-once strategies based on PINNs, where optimal control, state, and adjoint variables are estimated with neural networks properly trained, generally speaking, minimizing the residuals of first-order optimality conditions, have been also investigated [47, 4, 18, 77]. Notably, akin to our approach, [18, 77] consider parametrized PDEs to quickly retrieve parameter-specific optimal control strategies: however, in contrast to our approach, the lack of dimensionality reduction techniques and a data-driven approach result in expensive training phases.

Another interesting and promising method to solve control problems is represented by Reinforcement Learning [71] and, more recently, Deep Reinforcement Learning (DRL). This framework aims to learn an optimal policy, that is, an optimal state-to-control map, in order to control a physical system from sensor-based measurements or simulated snapshots. This is achieved thanks to an iterative interaction between the agent – which has to decide the best possible control action to apply – and the environment, which is the dynamical system under investigation, possibly described by differential equations. At every episode, the agent applies a time-dependent action on the environment and receives the corresponding state and reward values: thanks to this information, it is possible to improve the policy towards the optimum trying to maximize the reward, i.e., to minimize the cost. This strategy has been exploited in several applications such as, e.g., flow control problems [60, 72, 55] or metamaterial design and cloaking [65, 45]. DRL algorithms usually deal only with low-dimensional variables due to sample inefficiency and computationally demanding

training phases. To restrict the number of agent-environment interactions, [78] exploit SINDy to build efficient and interpretable surrogate models for the environment, the reward and the policy, however only addressing low-dimensional dynamical systems and without considering systems described in terms of PDEs. [42], instead, handle high-dimensional observations through autoencoders in order to jointly reduce the problem dimensionality and learning an optimal policy more quickly in the context of fluid jets control for ball games. In this work, we consider similar reduction strategies to deal with high-dimensional state and control variables. Moreover, we focus on parametrized PDEs to easily handle state variability, as recently proposed by [8] in the context of DRL, and we exploit an offline-online decomposition as done by [62]. However, we do not consider feedback signals; rather, we only require to select the new scenario parameters to infer the corresponding optimal pair, with no need of state measurements online. For a complete overview on DRL see, e.g., [23, 73, 56] while, for a general introduction on hybrid approaches that combine machine learning techniques into physical problems, see, e.g., [11, 61, 74, 59].

The paper is organized as follows. Section 2 reviews parametrized optimal control problems and intrusive dimensionality reduction techniques, such as the RB method. Section 3 proposes a nonlinear and non-intrusive reduced order modeling strategy based on deep learning to solve optimal control problems of parametrized PDEs in real-time. Section 4 shows the performances of the proposed approach on three different applications ranging from flow control to active thermal cooling. Section 5 discusses some future development ideas about possible extensions of the proposed tool.

## 2. Optimal control of parametrized partial differential equations

This section briefly introduces the formulation of parametrized OCPs, the high-fidelity full-order model based on the finite element method, and standard dimensionality reduction techniques taken into account to speed up the resolution of OCPs.

### 2.1. Mathematical formulation of parametrized optimal control problems

Optimal control problems aim at finding the best control action capable of steering the dynamics of a (state) system as close as possible to a target configuration. This task is usually achieved by minimizing a cost functional subject to the constraint expressed by the state equation – this latter consisting of a differential problem – as follows:

$$\text{Given } \boldsymbol{\mu}_s, \quad \text{find } J_h(\mathbf{x}_h; \boldsymbol{\mu}_s) \rightarrow \min \quad \text{s.t.} \quad \mathbf{G}_h(\mathbf{x}_h; \boldsymbol{\mu}_s) = \mathbf{0}, \quad \mathbf{x}_h \in \mathcal{X}_{ad} \quad (1)$$

where  $\mathbf{x}_h = (\mathbf{y}_h, \mathbf{u}_h) \in \mathbb{R}^{N_h^y} \times \mathbb{R}^{N_h^u}$  are the optimization (state and control) variables, possibly subject to additional constraints (here  $\mathcal{X}_{ad} = \mathcal{Y}_{ad} \times \mathcal{U}_{ad}$ ). For the sake of simplicity, we directly take into account the discrete state and control variables, obtained by discretizing the state problem by means of a high-fidelity full-order model (FOM) built through, e.g., the finite element method (FEM). In practice, the infinite-dimensional state  $y$  and control  $u$  are approximated by  $y_h \in \mathcal{Y}_h$  and  $u_h \in \mathcal{U}_h$ , where  $\mathcal{Y}_h$  and  $\mathcal{U}_h$  are finite-dimensional spaces spanned by a basis of  $N_h^y$  and  $N_h^u$  elements, respectively – here  $h > 0$  denotes the discretization parameter related to the mesh size, so that the smaller  $h$ , the larger  $N_h^y$  and  $N_h^u$  and the more accurate and computational expensive the FEM approximation. In this way, we can retrieve the algebraic formulation of the differential equations in terms of  $\mathbf{y}_h \in \mathbb{R}^{N_h^y}$  and  $\mathbf{u}_h \in \mathbb{R}^{N_h^u}$ , that are the basis expansion coefficients of  $y_h$  and  $u_h$ . Notice that, whenever a nodal basis is chosen,  $\mathbf{y}_h$  and  $\mathbf{u}_h$  corresponds to the state and control values at the mesh nodes. For a complete presentation of FEM and optimal control problems, see, e.g., [53, 44]. In particular:

1. **State equation.** The governing equation  $\mathbf{G}_h(\mathbf{x}_h; \boldsymbol{\mu}_s) = \mathbf{0}$ , with  $\mathbf{G}_h \in \mathbb{R}^{N_h^y}$ , encodes the physical law describing the state dynamics in a region of interest  $\Omega$  and, possibly, over a time interval  $(0, T)$  for a given final time  $T > 0$ . In this work, we focus on (nonlinear and/or time-dependent) PDEs parametrized by a vector of  $p$  scenario parameters  $\boldsymbol{\mu}_s$  belonging to a parameter space  $\mathcal{P} \subset \mathbb{R}^p$ . For example,  $\boldsymbol{\mu}_s$  may (i) represent physical or material properties appearing in the equation, (ii) account for the geometrical variability of the region where the phenomenon takes place, or (iii) appear as a result of the parametrization of a given term in the PDE. Note that, in case of time-dependent problems, the time variable  $t$  is regarded as an additional scenario parameter. The considered PDEs

are thus affected by both the control action  $\mathbf{u}_h$  and the vector of parameters  $\boldsymbol{\mu}_s$ . The former is the quantity to tune in order to steer the dynamics towards the target, the latter represents the scenario variability that we want to address. Indeed, we aim to control the dynamics for multiple scenarios as fast as possible. The governing equation also includes suitable boundary and initial conditions on, respectively,  $\partial\Omega \times (0, T]$  and  $\Omega \times \{t = 0\}$ , in order to deal with well-posed problems and to ensure the existence and uniqueness of an optimal solution that depends continuously on the problem data.

2. **Cost functional.**  $J_h(\mathbf{x}_h; \boldsymbol{\mu}_s) : \mathbb{R}^{N_h^y} \times \mathbb{R}^{N_h^u} \rightarrow \mathbb{R}$  is a function encoding the objective to be achieved. The minimum point of  $J_h$  corresponds to the optimal state and control variables, also referred to as optimal pair. To guarantee the OCP well-posedness and avoid control strategies that are unfeasible from an energetic standpoint, additional regularization terms are typically added in the cost functional like, e.g., the  $L^2$ -norm of the control and its gradient [44]. Note that, since the information on the state of the system is often incomplete, a partial observation  $\mathbf{z}_h = C_h \mathbf{y}_h \in \mathbb{R}^{N_h^z}$ , with  $C_h : \mathbb{R}^{N_h^y} \rightarrow \mathbb{R}^{N_h^z}$ , can replace the state in the definition of  $J_h$ .

## 2.2. Parametrized optimal control problems: from full-order models to reduced order models

Solving PDE-constrained optimization problems is a challenging task due to the interplay between the optimization workflow and the approximation of the underlying PDE. In the unconstrained case (when no constraints are set on  $\mathbf{x}_h$ , that is,  $\mathcal{X}_{ad}$  is the entire space) a set of Karush-Kuhn-Tucker (KKT) optimality conditions can be obtained through the Lagrange multipliers' method, and reads as follows:

$$\mathbf{F}_h(\mathbf{x}_h, \mathbf{p}_h; \boldsymbol{\mu}_s) = \begin{bmatrix} \mathbf{g}_h(\mathbf{x}_h; \boldsymbol{\mu}_s) + \mathbf{C}_h^\top(\mathbf{x}_h; \boldsymbol{\mu}_s) \mathbf{p}_h \\ \mathbf{G}_h(\mathbf{x}_h; \boldsymbol{\mu}_s) \end{bmatrix} = \mathbf{0} \quad (2)$$

where  $\mathbf{p}_h \in \mathbb{R}^{N_h^p}$  is the discrete adjoint vector,

$$\mathbf{C}_h^\top(\mathbf{x}_h; \boldsymbol{\mu}_s) = [\partial_y \mathbf{G}_h(\mathbf{x}_h; \boldsymbol{\mu}_s) \quad \partial_u \mathbf{G}_h(\mathbf{x}_h; \boldsymbol{\mu}_s)] \in \mathbb{R}^{N_h^p \times (N_h^y + N_h^u)}$$

is the state operator Jacobian, whereas

$$\mathbf{g}_h = [\nabla_y J_h(\mathbf{x}_h; \boldsymbol{\mu}_s) \quad \nabla_u J_h(\mathbf{x}_h; \boldsymbol{\mu}_s)]^\top \in \mathbb{R}^{N_h^y + N_h^u}$$

is the gradient of  $J_h$  with respect to  $\mathbf{x}_h = (\mathbf{y}_h, \mathbf{u}_h)$ . Assuming that the same space is used to discretize both state and adjoint variables, we have  $N_h^p = N_h^y$ . To better highlight the structure of the KKT optimality conditions and the role of the adjoint equation, we can equivalently rewrite system (2) as follows:

$$\begin{cases} \nabla_y J_h((\mathbf{y}_h, \mathbf{u}_h); \boldsymbol{\mu}_s) + (\partial_y \mathbf{G}_h((\mathbf{y}_h, \mathbf{u}_h); \boldsymbol{\mu}_s))^\top \mathbf{p}_h = \mathbf{0} & \text{(adjoint equation)} \\ \nabla_u J_h((\mathbf{y}_h, \mathbf{u}_h); \boldsymbol{\mu}_s) + (\partial_u \mathbf{G}_h((\mathbf{y}_h, \mathbf{u}_h); \boldsymbol{\mu}_s))^\top \mathbf{p}_h = \mathbf{0} & \text{(optimality condition)} \\ \mathbf{G}_h((\mathbf{y}_h, \mathbf{u}_h); \boldsymbol{\mu}_s) = \mathbf{0} & \text{(state equation).} \end{cases} \quad (3)$$

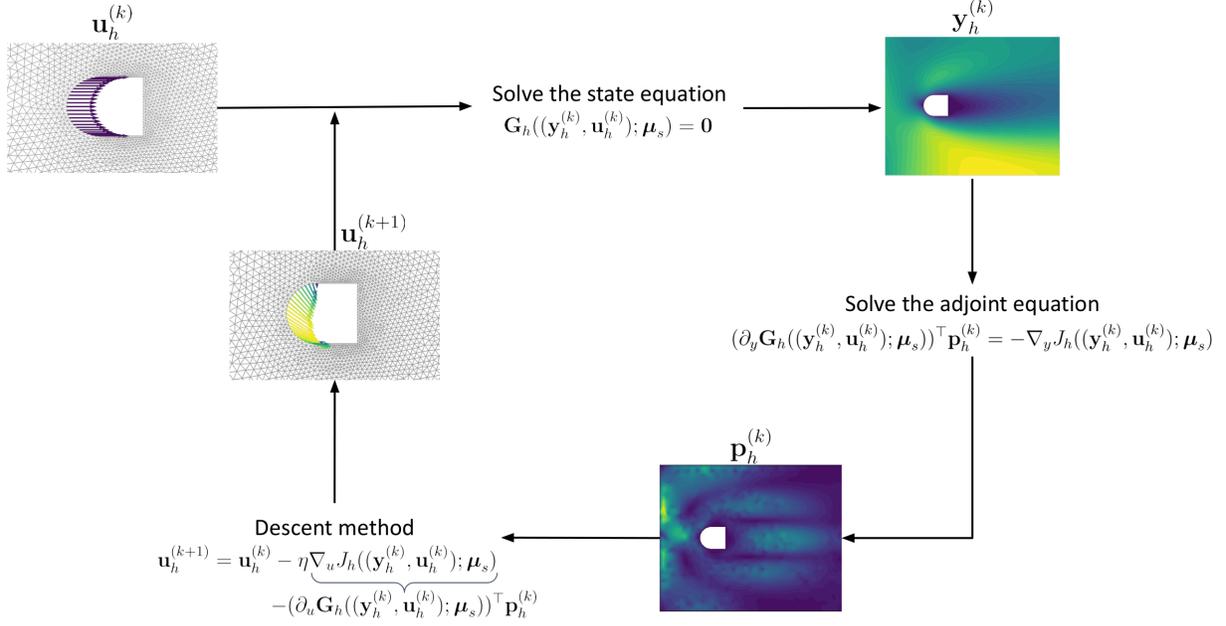
The optimality system (3) is usually a coupled, nonlinear (and possibly time-dependent) system that can in principle be solved through the Newton method – this corresponds to the so-called *direct method* in optimal control. Alternatively, the *indirect method* takes into account a two-step process where (i) the solution of the state and the adjoint problems and (ii) the minimization of the cost functional by means of a descent method are considered iteratively until convergence.

*Remark.* In the simplest case where a quadratic cost functional

$$J_h(\mathbf{y}_h, \mathbf{u}_h; \boldsymbol{\mu}_s) = \frac{1}{2}(\mathbf{y}_h - \mathbf{y}_d)^\top \mathbb{M}_h(\mathbf{y}_h - \mathbf{y}_d) + \frac{\beta}{2} \mathbf{u}_h^\top \mathbb{M}_h \mathbf{u}_h$$

and a linear, stationary state problem

$$\mathbf{G}_h(\mathbf{y}_h, \mathbf{u}_h; \boldsymbol{\mu}_s) = \mathbb{A}_h(\boldsymbol{\mu}_s) \mathbf{y}_h - \mathbf{f}_h(\boldsymbol{\mu}_s) - \mathbb{B}_h(\boldsymbol{\mu}_s) \mathbf{u}_h$$



**Figure 1.** *Test 1.1.* Indirect approach for a PDE-constrained optimization problem in the case of steady fluid flow control. Here  $\mathbf{u}_h^{(k)}$  is the suboptimal boundary control at the  $k$ -th iteration of the optimization loop;  $\mathbf{y}_h^{(k)}$  denotes the corresponding flow velocity computed by solving the steady Navier-Stokes equations;  $\mathbf{p}_h^{(k)}$  is the corresponding adjoint variable, resulting from the adjoint equation;  $\mathbf{u}_h^{(k+1)}$  is the boundary control updated through the steepest descent method with step size  $\eta$ .

are considered, with  $\mathbf{y}_d$  standing for the target state, (3) becomes a linear system featuring a saddle-point structure, of the following form:

$$\underbrace{\begin{bmatrix} \mathbb{M}_h & 0 & \mathbb{A}_h^\top \\ 0 & \beta \mathbb{M}_h & -\mathbb{M}_h \\ \mathbb{A}_h & -\mathbb{M}_h & 0 \end{bmatrix}}_{K_h(\boldsymbol{\mu}_s)} \underbrace{\begin{bmatrix} \mathbf{y}_h \\ \mathbf{u}_h \\ \mathbf{p}_h \end{bmatrix}}_{\mathbf{q}_h(\boldsymbol{\mu}_s)} = \underbrace{\begin{bmatrix} \mathbb{M}_h \mathbf{y}_d \\ \mathbf{0} \\ \mathbf{f}_h \end{bmatrix}}_{\mathbf{s}_h(\boldsymbol{\mu}_s)}. \quad (4)$$

where  $\mathbb{A}_h$  denotes the stiffness matrix,  $\mathbb{M}_h$  the mass matrix,  $\mathbb{B}_h$  the control-to-state matrix,  $\mathbf{f}_h$  the source term.

Figure 1 shows the rationale behind indirect approaches, focusing on the control problem we will tackle in Section 4.1.1 for a steady flow past an obstacle. Starting from an initial control strategy, here expressed as a boundary datum on the obstacle, the flow field is computed together with the adjoint variables in order to compute the gradient  $\nabla_u J_h(\mathbf{y}_h(\mathbf{u}_h), \mathbf{u}_h)$  and then update the control values toward the minimum of the loss function in case, e.g., a steepest descent method is employed. These steps have to be repeated iteratively until a suitable convergence criterion is met.

The loop sketched in Figure 1 thus requires a PDE solve at every optimization step in order to compute the state  $\mathbf{y}_h$  starting from the current control  $\mathbf{u}_h$ . When considering the algebraic formulation obtained through the FEM approximation, this is equivalent to solve a (possibly, nonlinear and/or time-dependent) system of  $N_h^y$  equations. Similarly, the adjoint problem requires the solution of a linear (possibly time-dependent) system of  $N_h^p$  equations. Especially when taking into account challenging small-scales dynamics,  $N_h^y$  (and  $N_h^p$ ) can be rather large, ultimately entailing a huge computational burden. Even more importantly, whenever optimal control strategies must be determined for different scenarios of interest, the entire optimization procedure has to be repeated almost from scratch. For these reasons, as usually suggested in many-query contexts, smart reduction strategies are strongly suggested to shrink the problem dimension and to speed up the OCPs resolutions.

The Reduced Basis (RB) method [28, 54] relying, e.g., on Proper Orthogonal Decomposition (POD), may be exploited to speed up computations when solving parametrized OCPs [49, 48, 54]. Thanks to the

RB method, the dimensions of state, control, and adjoint variables can be reduced through a projection onto linear subspaces of dimensions  $N_y$ ,  $N_u$ , and  $N_p$ , that is, we express

$$\mathbf{y}_h(\boldsymbol{\mu}_s) \approx \mathbb{V}_y \mathbf{y}_N(\boldsymbol{\mu}_s), \quad \mathbf{u}_h(\boldsymbol{\mu}_s) \approx \mathbb{V}_u \mathbf{u}_N(\boldsymbol{\mu}_s), \quad \mathbf{p}_h(\boldsymbol{\mu}_s) \approx \mathbb{V}_p \mathbf{p}_N(\boldsymbol{\mu}_s) \quad (5)$$

where the projection matrices  $\mathbb{V}_y \in \mathbb{R}^{N_h^y \times N_y}$ ,  $\mathbb{V}_u \in \mathbb{R}^{N_h^u \times N_u}$ , and  $\mathbb{V}_p \in \mathbb{R}^{N_h^p \times N_p}$  collect the basis elements spanning the low-dimensional RB subspaces. These latter are obtained through the Singular Values Decomposition (SVD) of the matrices collecting optimal state, control, and adjoint snapshots computed for suitably sampled scenarios  $\{\boldsymbol{\mu}_s^{(1)}, \dots, \boldsymbol{\mu}_s^{(N_s)}\} \in \mathcal{P}^{N_s}$  and, in case of time-dependent problems, on a time grid discretizing the interval  $[0, T]$ . In particular, RB subspaces for parametrized OCPs are the left singular vectors associated to the  $N_y$ ,  $N_u$ ,  $N_p$  largest singular values, respectively. On the other hand, the reduced coordinates are usually obtained through a (Petrov-)Galerkin projection of each equation of (3) onto the corresponding subspace  $\mathbb{V}_y$ ,  $\mathbb{V}_u$ , and  $\mathbb{V}_p$ . See, e.g., [58, 33, 6, 43] for some examples on the use of POD-based ROM techniques in the case of optimal control problems.

Compared to an indirect approach, in which each equation appearing in the system (3) is usually reduced independently, a suitable compatibility condition must be ensured between state and adjoint reduced subspaces when a direct approach is used. In particular, if a Galerkin projection is employed to generate the ROM, an aggregated trial space  $\mathbb{V}_{yp}$  made of both state and adjoint basis functions is introduced to preserve the stability of the resulting ROM. This latter would read as follows:

$$\underbrace{\mathbb{V}^\top \mathbf{F}_h(\mathbb{V} \mathbf{q}_N(\boldsymbol{\mu}_s))}_{\mathbf{F}_N(\mathbb{V} \mathbf{q}_N(\boldsymbol{\mu}_s))} = \mathbf{0} \quad \text{with} \quad \mathbf{q}_N(\boldsymbol{\mu}_s) = \begin{bmatrix} \mathbf{y}_N(\boldsymbol{\mu}_s) \\ \mathbf{u}_N(\boldsymbol{\mu}_s) \\ \mathbf{p}_N(\boldsymbol{\mu}_s) \end{bmatrix}, \quad \mathbb{V} = \begin{bmatrix} \mathbb{V}_{yp} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{V}_u & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{V}_{yp} \end{bmatrix}, \quad \mathbb{V}_{yp} = [\mathbb{V}_y \quad \mathbb{V}_p] \quad (6)$$

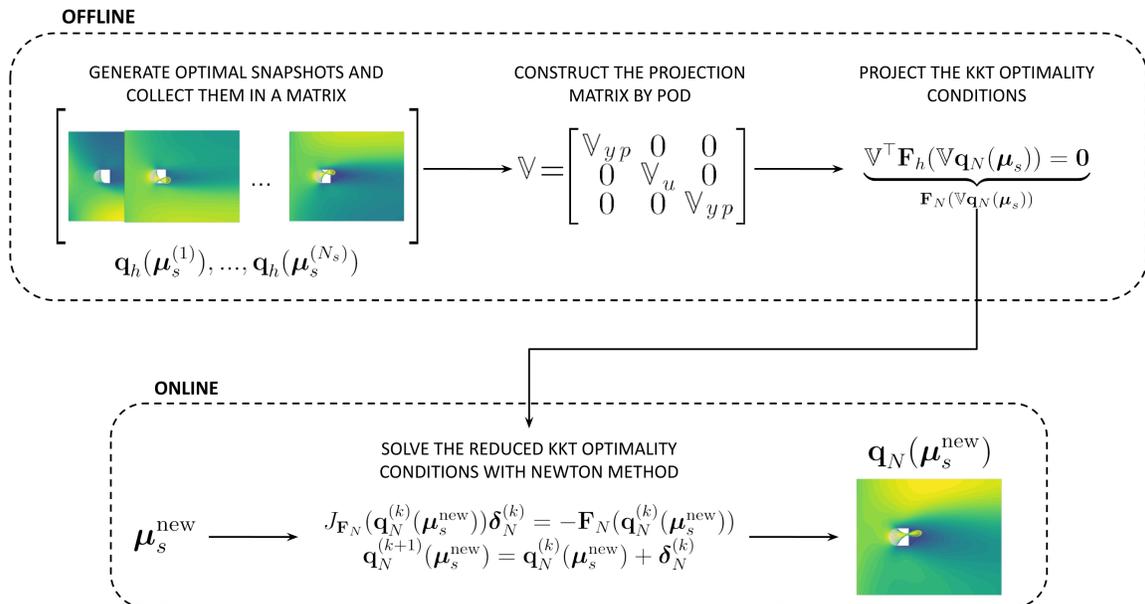
*Remark.* In the case of a linear-quadratic OCP, the RB method would yield the following reduced order model: given  $\boldsymbol{\mu}_s$ , solve

$$\underbrace{\mathbb{V}^\top K_h(\boldsymbol{\mu}_s) \mathbb{V}}_{K_N(\boldsymbol{\mu}_s)} \mathbf{q}_N(\boldsymbol{\mu}_s) = \underbrace{\mathbb{V}^\top \mathbf{s}_h(\boldsymbol{\mu}_s)}_{\mathbf{s}_N(\boldsymbol{\mu}_s)}.$$

See, e.g., [49, 48] for a series of applications of direct OCP solvers exploiting the RB method in the case of either advection-diffusion problems or Stokes flows, and a thorough discussion on the construction of RB spaces in those cases.

A direct OCP solver exploiting the RB method is summarized in Figure 2. For any new scenario, the optimal state, adjoint and control can be computed all-at-once by solving the reduced Karush-Kuhn-Tucker (KKT) system, that is obtained through a Galerkin projection. Thanks to the RB method, it is therefore possible to retrieve, online, the optimal solution corresponding to a new unseen scenario  $\boldsymbol{\mu}_s^{\text{new}}$  faster through the resolution of a smaller KKT system. The use of a direct approach, aiming at solving the parametrized KKT system through the RB method, can be found instead in [49, 48, 35].

Despite several results featuring a wide range of applications, the RB method lacks of efficiency whenever aiming at the rapid control of systems in multiple scenarios. Indeed, the RB method relies on the linear superimposition of modes: hence, the ROM dimension can become moderately large to ensure accuracy in case of more complex problems, therefore avoiding a very rapid solution of the ROM problem for any new parameter instance. This might happen, e.g., if the parameter-to-solution dependence is rather involved, or in the case of complex physical behaviors of the state system (e.g., in presence of strong nonlinearities or dominated advection regimes). Moreover, the intrinsic coupling between state, adjoint, and control variables adds a further, strong difficulty, requiring to reduce all the three fields involved in the system of optimality conditions. Furthermore, a projection-based ROM is an intrusive technique since the FOM structures, such as the matrix and the right-hand side of the KKT system appearing in (2), must be directly modified and projected to obtain the ROM structures appearing in (6). This operation is not always feasible since the arrays appearing in (2) may be not directly accessible; on the other hand, the use of hyper-reduction strategies in presence of either non-affine parameter dependencies or nonlinearities is almost out of reach for



**Figure 2.** *Test 1.1.* Offline-online decomposition required to solve PDE-constrained optimization problems faster exploiting a reduced basis solver, in the case of steady flow control. Offline: generation of optimal state, adjoint and control snapshots for random scenarios and reduction of the first-order optimality conditions in the KKT system. Online: resolution of the reduced KKT system to compute the optimal state, adjoint and control for a new scenario of interest. In general, the reduced KKT system is solved through the Newton method with reduced Jacobian matrix  $J_{F_N}$  and right-hand side  $F_N$  properly approximated.

OCPs given the coupled nature of these problems. Last, but not least, the inclusion of either control or state constraints is not straightforward in a ROM obtained through the RB method.

All these difficulties led us to draw an alternative ROM strategy to tackle the rapid solution of parametrized OCPs in multiple scenarios leveraging on deep learning algorithms, focusing in particular on deep learning-based reduced order models, as detailed in the following section.

### 3. Deep learning-based reduced order modeling for parametrized optimal control problems

This section presents a nonlinear and non-intrusive ROM strategy that allows to retrieve optimal states and controls in real-time for a given problem in multiple scenarios. Nonlinear and non-intrusive ROMs have been recently introduced – see, e.g., [27, 21, 22] – in order to overcome the limitations of the RB method highlighted in Section 2.2. In particular, instead of considering linear approximations relying on POD subspaces, snapshots can be compressed onto a low-dimensional space through nonlinear maps such as, e.g., autoencoders (AEs), that must be learnt. This allows to achieve the task of learning a reduced trail manifold in a much more effective way. Indeed, an autoencoder is a neural network architecture consisting of an encoder that compresses the available snapshots (to reach a very low-dimensional representation in the so-called latent space) and a decoder that recovers the full input information. The composition of the encoder and the decoder somehow approximates the identity map – imposing that the reconstructed data are as close as possible to the original ones. Differently from POD where the encoding and decoding actions are performed through multiplications by the matrices  $\mathbb{V}^\top$  and  $\mathbb{V}$ , the nonlinear activation functions within the hidden layers of the autoencoder network result in a nonlinear dimensionality reduction. For instance, [21, 20] propose a Deep Learning-based Reduced Order Model (DL-ROM) which exploits convolutional autoencoders to take advantage of the spatial structure of PDE solutions prioritizing the dependencies of nearby values. [22], instead, presents the so-called POD-DL-ROM framework where (i) the snapshots are initially compressed through POD and (ii) the resulting POD coefficients are further reduced by an autoencoder – see also [9] for further details. The reduction combining POD and autoencoders will henceforth be referred to as POD+AE.

Observe that, whenever POD is enough to retrieve an acceptable low-dimensional latent representation of the data, nonlinear reduction schemes may be avoided, as proposed by [27] with the POD-NN framework.

Another key ingredient to achieve flexible and efficient ROMs is non-intrusiveness. Specifically, instead of projecting, assembling and solving a reduced system of equations as in the case of the RB method, the solutions related to new scenarios are rather retrieved by means of surrogate models. For instance, deep feed-forward neural networks may be taken into account to generate the reduced solution in the latent space starting from the input parameters  $\boldsymbol{\mu}_s$  (and, in case of time-dependent problems, the time variable  $t$ ). For example, regarding the POD-NN framework, a neural network is exploited to approximate non-intrusively the map from  $\boldsymbol{\mu}_s$  (and possibly time) to the projected snapshots data in the POD subspace, thus avoiding the (usually, intrusive) assembling of a reduced problem. Note that the reduction step is crucial to deal with low-dimensional neural networks, allowing for lighter training and faster evaluations. By doing so, after training the solution associated with a new time instant  $t^{\text{new}}$  and a new scenario  $\boldsymbol{\mu}_s^{\text{new}}$  may be computed inexpensively, through a forward pass of the parameter-to-solution map and the decoder.

To rapidly solve a wide range of OCPs for multiple scenarios of interest, it is therefore possible to adapt and exploit the workflow of DL-ROMs introduced in [21, 20]. Specifically, we propose a deep learning-based ROM technique in the context of parametrized OCPs, as outlined in the following offline-online decomposition.

### 3.1. Offline (training) phase

During the offline phase, the neural network architectures – i.e., the autoencoder for the sake of dimensionality reduction, and the feedforward neural network representing the parameter-to-solution map – must be trained. Prior to this task, a set of snapshots representing optimal (state and control) solutions for selected values of the input parameters  $\boldsymbol{\mu}_s$  must be computed exploiting a high-fidelity solver. More in detail, we perform the following steps.

- We compute  $N_s$  full-order optimal state and control snapshots by solving the system of first-order optimality conditions (3) for different random scenario parameters sampled in the parameter space  $\mathcal{P}$ , i.e.

$$\begin{aligned} & \{\mathbf{y}_h(\boldsymbol{\mu}_s^{(1)}), \dots, \mathbf{y}_h(\boldsymbol{\mu}_s^{(N_s)})\} \\ & \{\mathbf{u}_h(\boldsymbol{\mu}_s^{(1)}), \dots, \mathbf{u}_h(\boldsymbol{\mu}_s^{(N_s)})\} \end{aligned}$$

where the number of snapshots  $N_s$  strongly depends on several factors such as the scenario parameters dimension as well as the complexity of the parameter-to-solution map and the physics at hand. In case of time-dependent problems, for each sampled scenario, the optimal snapshots are computed on a time grid  $\{t_j\}_{j=1}^{N_t}$  discretizing the interval  $[0, T]$ , that is

$$\begin{aligned} & \{\mathbf{y}_h(t_1, \boldsymbol{\mu}_s^{(1)}), \dots, \mathbf{y}_h(t_{N_t}, \boldsymbol{\mu}_s^{(1)}), \dots, \mathbf{y}_h(t_1, \boldsymbol{\mu}_s^{(N_s)}), \dots, \mathbf{y}_h(t_{N_t}, \boldsymbol{\mu}_s^{(N_s)})\} \\ & \{\mathbf{u}_h(t_1, \boldsymbol{\mu}_s^{(1)}), \dots, \mathbf{u}_h(t_{N_t}, \boldsymbol{\mu}_s^{(1)}), \dots, \mathbf{u}_h(t_1, \boldsymbol{\mu}_s^{(N_s)}), \dots, \mathbf{u}_h(t_{N_t}, \boldsymbol{\mu}_s^{(N_s)})\} \end{aligned}$$

To generate controlled pairs, in this work we exploit high-fidelity OCP solvers based on a finite element discretization of the system of optimality conditions. Note that either direct or indirect approaches can be used to generate snapshots, in case a direct approach yields an algebraic problem hard to solve. Note also that optimal (state and control) snapshots are useful to restrict optimization procedures offline and thus to achieve real-time control capabilities in the subsequent online stage.

- We reduce the dimensionality of optimal state snapshots and, in case of high-dimensional controls, of optimal control snapshots, too. To reach this goal either POD, an autoencoder, or a combination of the two strategies (POD+AE), can be employed. If POD is taken into account, the reduction results in the matrix multiplications between snapshots and projection matrices, that is

$$\mathbf{y}_N = \mathbb{V}_y^\top \mathbf{y}_h, \quad \mathbf{u}_N = \mathbb{V}_u^\top \mathbf{u}_h$$

When considering DL-ROMs involving autoencoders on full-order snapshots, namely

$$\begin{aligned}\mathbf{y}_N &= \varphi_E^y(\mathbf{y}_h); & \mathbf{y}_h &= \varphi_D^y(\mathbf{y}_N) \\ \mathbf{u}_N &= \varphi_E^u(\mathbf{u}_h); & \mathbf{u}_h &= \varphi_D^u(\mathbf{u}_N)\end{aligned}$$

where  $\varphi_E^y, \varphi_D^y, \varphi_E^u$  and  $\varphi_D^u$  denotes the encoding and decoding networks applied respectively to state and control, the following mean squared errors are taken into account as training losses

$$\begin{aligned}J_{AE}^y &= \|\mathbf{y}_h - \mathbf{y}_{h,\text{rec}}\|^2 = \|\mathbf{y}_h - \varphi_D^y(\varphi_E^y(\mathbf{y}_h))\|^2 \\ J_{AE}^u &= \|\mathbf{u}_h - \mathbf{u}_{h,\text{rec}}\|^2 = \|\mathbf{u}_h - \varphi_D^u(\varphi_E^u(\mathbf{u}_h))\|^2\end{aligned}$$

where  $\|\cdot\|$  stands for the Euclidean norm. Instead, in the context of POD-DL-ROM where a POD+AE reduction is performed, the autoencoder is applied on the projection of the snapshot data onto a (potentially larger) POD space, namely

$$\begin{aligned}\mathbf{y}_N &= \varphi_E^y(\mathbb{V}_y^\top \mathbf{y}_h); & \mathbf{y}_h &= \mathbb{V}_y \varphi_D^y(\mathbf{y}_N) \\ \mathbf{u}_N &= \varphi_E^u(\mathbb{V}_u^\top \mathbf{u}_h); & \mathbf{u}_h &= \mathbb{V}_u \varphi_D^u(\mathbf{u}_N)\end{aligned}$$

and the training is performed with respect to the loss functions

$$\begin{aligned}J_{POD+AE}^y &= \|\mathbb{V}_y^\top \mathbf{y}_h - \mathbb{V}_y^\top \mathbf{y}_{h,\text{rec}}\|^2 = \|\mathbb{V}_y^\top \mathbf{y}_h - \varphi_D^y(\varphi_E^y(\mathbb{V}_y^\top \mathbf{y}_h))\|^2 \\ J_{POD+AE}^u &= \|\mathbb{V}_u^\top \mathbf{u}_h - \mathbb{V}_u^\top \mathbf{u}_{h,\text{rec}}\|^2 = \|\mathbb{V}_u^\top \mathbf{u}_h - \varphi_D^u(\varphi_E^u(\mathbb{V}_u^\top \mathbf{u}_h))\|^2.\end{aligned}$$

- We build and train a deep feed-forward neural network  $\varphi$  mapping the scenario parameters  $\boldsymbol{\mu}_s$  – and, in case of unsteady problems, the time  $t$  – onto the latent representations of optimal state and control, that is

$$\begin{bmatrix} \tilde{\mathbf{y}}_N \\ \tilde{\mathbf{u}}_N \end{bmatrix} = \varphi(t, \boldsymbol{\mu}_s).$$

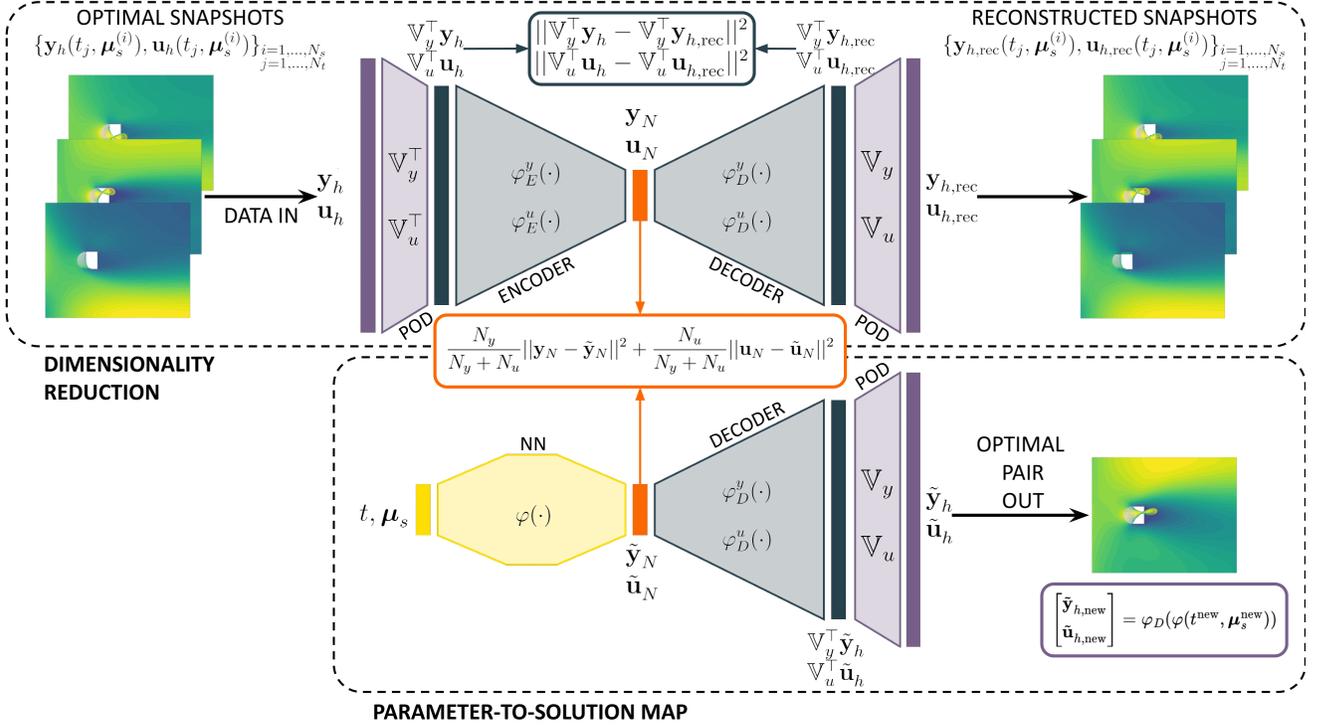
In particular, the loss function employed to train  $\varphi$  is

$$J_\varphi = \frac{N_y}{N_y + N_u} \|\mathbf{y}_N - \tilde{\mathbf{y}}_N\|^2 + \frac{N_u}{N_y + N_u} \|\mathbf{u}_N - \tilde{\mathbf{u}}_N\|^2$$

where the two weighting coefficients equally leverage state and control data reconstruction at the latent level.

Figure 3 visually represents the proposed architecture when taking into account a POD-DL-ROM. Note that, in general, we select the most efficient non-intrusive reduced order model – among POD-NN, DL-ROM, and POD-DL-ROM – for every specific OCP balancing lightweight and accuracy. In addition, according to the dynamics under investigation and the control action chosen, optimal state and control snapshots may be compressed exploiting two different reduction strategies. Moreover,  $\mathbf{y}_h$  and  $\mathbf{u}_h$  may be reduced independently, as done in the applications detailed in Section 4, or together with, e.g, a single autoencoder. The dimensionality reduction step is crucial to deal with high-dimensional state variables and distributed control fields, as of interest in the applications presented in Section 4: indeed, it allows to consider parameter-to-solution maps with compressed output sizes that are lighter to train and faster to evaluate online.

*Remark.* Note that, despite a set of solutions of the adjoint problem are generated while solving the system of optimality conditions during the snapshots' calculation, these data are discarded and no reduction is operated on the adjoint field, in the proposed workflow. This also represents a major difference compared to projection-based ROMs for optimal control problems, that must include the evaluation of the adjoint solution, thus yielding higher complexity and costs. Nevertheless, in case the adjoint solution is also required as a specific requirement of the application at hand, the workflow discussed so far can be easily adapted to generate the adjoint field, too.



**Figure 3.** *Test 1.2.* Deep learning-based reduced order model architecture to solve parametrized optimal control problems in real-time. Optimal state and control snapshots are generated and reduced through a combination of POD and autoencoders. Moreover, the map  $\varphi$  from time and scenario parameters to reduced optimal pair is modeled through a feed-forward neural network. After properly training the networks in the offline phase, the full-order optimal pair corresponding to a new time instant  $t^{\text{new}}$  and a new scenario  $\mu_s^{\text{new}}$  is retrieved online by a forward pass through  $\varphi$  and the decoders.

### 3.2. Online phase

During the online phase, once the proposed architecture has been trained, it is possible to evaluate the optimal (state and control) solution for new scenarios. In particular, we can retrieve the optimal pair associated with a new scenario  $\mu_s^{\text{new}}$  and, in case of time-dependent problems, a new time instant  $t^{\text{new}}$  through a forward pass of the trained neural networks, that is

$$\begin{bmatrix} \tilde{\mathbf{y}}_{h,\text{new}} \\ \tilde{\mathbf{u}}_{h,\text{new}} \end{bmatrix} = \varphi_D \left( \begin{bmatrix} \tilde{\mathbf{y}}_{N,\text{new}} \\ \tilde{\mathbf{u}}_{N,\text{new}} \end{bmatrix} \right) = \varphi_D(\varphi(t^{\text{new}}, \mu_s^{\text{new}}))$$

where, for the sake of compactness,  $\varphi_D$  combines the state and control decoders. Specifically, depending on the reduction technique taken into account,  $\varphi_D$  results in

$$\varphi_D \left( \begin{bmatrix} \tilde{\mathbf{y}}_{N,\text{new}} \\ \tilde{\mathbf{u}}_{N,\text{new}} \end{bmatrix} \right) = \begin{cases} \begin{bmatrix} \mathbb{V}_y & 0 \\ 0 & \mathbb{V}_u \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{y}}_{N,\text{new}} \\ \tilde{\mathbf{u}}_{N,\text{new}} \end{bmatrix} & \text{in case a POD-NN is used} \\ \begin{bmatrix} \varphi_D^y(\tilde{\mathbf{y}}_{N,\text{new}}) \\ \varphi_D^u(\tilde{\mathbf{u}}_{N,\text{new}}) \end{bmatrix} & \text{in case a DL-ROM is used} \\ \begin{bmatrix} \mathbb{V}_y & 0 \\ 0 & \mathbb{V}_u \end{bmatrix} \begin{bmatrix} \varphi_D^y(\tilde{\mathbf{y}}_{N,\text{new}}) \\ \varphi_D^u(\tilde{\mathbf{u}}_{N,\text{new}}) \end{bmatrix} & \text{in case a POD-DL-ROM is used.} \end{cases}$$

The proposed architecture is extremely flexible thanks to its non-intrusive data-driven nature and the independence of the surrogate models on the particular problem faced. Indeed, it can be applied to a wide range of parametrized OCPs, as demonstrated throughout Section 4, and can be extended to deal with different data sources, such as sensors or recordings.

## 4. Numerical results

This section is devoted to the numerical results obtained with the proposed architecture when tackling three distinct parametric OCPs in two spatial dimensions. Section 4.1 focuses on nonlinear OCPs involving the minimization of the energy dissipated by a fluid flow passing an obstacle in a channel considering both steady and time-dependent Navier-Stokes equations to model its dynamics. Section 4.2, instead, takes into account an active thermal cooling problem where the temperature of an object has to be kept equal to a reference value. To perform error analysis and fairly assess the performance of the proposed machinery, we consider a test dataset  $\{\mathbf{y}_h(t_j, \boldsymbol{\mu}_s^{(i)}), \mathbf{u}_h(t_j, \boldsymbol{\mu}_s^{(i)})\}_{(i,j) \in I_{test}}$  where  $I_{test} \subset \{1, \dots, N_s\} \times \{1, \dots, N_t\}$  and  $|I_{test}| \approx 0.2N_tN_s$ . In particular, the following mean relative errors are considered for evaluation purposes:

$$\varepsilon_{\text{rel}, L^2}^y = \frac{1}{|I_{test}|} \sum_{(i,j) \in I_{test}} \sqrt{\frac{\int_{\Omega} \|y_h(t_j, \boldsymbol{\mu}_s^{(i)}) - \tilde{y}_h(t_j, \boldsymbol{\mu}_s^{(i)})\|^2 d\Omega}{\int_{\Omega} \|y_h(t_j, \boldsymbol{\mu}_s^{(i)})\|^2 d\Omega}}$$

$$\varepsilon_{\text{rel}, L^2}^u = \frac{1}{|I_{test}|} \sum_{(i,j) \in I_{test}} \sqrt{\frac{\int_{\Omega} \|u_h(t_j, \boldsymbol{\mu}_s^{(i)}) - \tilde{u}_h(t_j, \boldsymbol{\mu}_s^{(i)})\|^2 d\Omega}{\int_{\Omega} \|u_h(t_j, \boldsymbol{\mu}_s^{(i)})\|^2 d\Omega}}$$

where  $y_h$  and  $u_h$  are the ground truth finite element functions associated with the coefficients  $\mathbf{y}_h$  and  $\mathbf{u}_h$ . Instead,  $\tilde{y}_h$  and  $\tilde{u}_h$  are the corresponding approximations associated with  $\mathbf{y}_{h,\text{rec}}$  and  $\mathbf{u}_{h,\text{rec}}$  when evaluating the ability of the chosen reduction technique, while they are linked to  $\tilde{\mathbf{y}}_h$  and  $\tilde{\mathbf{u}}_h$  when interested in the prediction errors committed by the proposed architecture. Notice that the Euclidean norm  $\|\cdot\|$  is useful to take into account both scalar and vector functions.

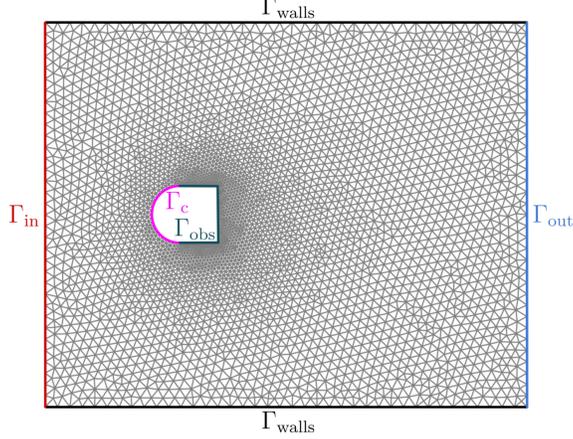
### 4.1. Flow control problems

In this section, we test the proposed real-time OCP solver in the case of nonlinear Navier-Stokes equations, which are resumed in the following system:

$$\begin{cases} \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} - \nu \Delta \mathbf{v}(\mathbf{x}, t) + (\mathbf{v}(\mathbf{x}, t) \cdot \nabla) \mathbf{v}(\mathbf{x}, t) + \nabla p(\mathbf{x}, t) = 0 & \text{in } \Omega \times (0, T) \\ \text{div } \mathbf{v}(\mathbf{x}, t) = 0 & \text{in } \Omega \times (0, T) \\ \mathbf{v}(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, t) & \text{on } \Gamma_c \times (0, T] \\ \mathbf{v}(\mathbf{x}, t) = \mathbf{0} & \text{on } \Gamma_{\text{obs}} \times (0, T] \\ \mathbf{v}(\mathbf{x}, t) = \mathbf{v}_{\text{in}}(\boldsymbol{\mu}_s) & \text{on } \Gamma_{\text{in}} \times (0, T] \\ \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = 0 & \text{on } \Gamma_{\text{walls}} \times (0, T] \\ (\nu \nabla \mathbf{v}(\mathbf{x}, t) - p(\mathbf{x}, t)) \mathbf{n}(\mathbf{x}) \cdot \mathbf{t}(\mathbf{x}) = 0 & \text{on } \Gamma_{\text{walls}} \times (0, T] \\ (\nu \nabla \mathbf{v}(\mathbf{x}, t) - p(\mathbf{x}, t)) \mathbf{n}(\mathbf{x}) = 0 & \text{on } \Gamma_{\text{out}} \times (0, T] \\ \mathbf{v}(\mathbf{x}, 0) = \mathbf{0} & \text{in } \Omega \times \{t = 0\}. \end{cases} \quad (7)$$

In the equations above,  $\nu$  denotes the kinematic viscosity, while  $\mathbf{n}$  and  $\mathbf{t}$  are the unit vectors normal and tangential to the domain boundary, respectively. We consider a rectangular channel  $\Omega = [0, 25] \times [0, 20]$  with an obstacle inside, inspired to configurations already explored in the literature several times [17, 48, 44]. The domain discretization in Figure 4 is generated through `gmsh` utilities [24] and consists of 3716 vertices and 7203 elements. The state variable in Equation (7) includes both the velocity vector  $\mathbf{v}(\mathbf{x}, t)$  ( $m s^{-1}$ ) and the pressure  $p(\mathbf{x}, t)$  ( $Pa$ ). However, for simplicity, only the velocity is considered in the following, disregarding the pressure field – although also this latter field could be retrieved by means of a slight adaptation of the proposed framework.

When dealing with the full-order model, the state is discretized with  $\mathbb{P}_2 - \mathbb{P}_1$  finite elements, resulting in  $N_h^v = 29270$  degrees of freedom for the velocity. The control action is instead represented by the time-dependent Dirichlet velocity datum  $\mathbf{u}$  on the rounded edge of the obstacle boundary  $\Gamma_c$ , which can be interpreted as fluid injection or absorption.  $\mathbf{u}$  is thus a boundary control depending on  $N_h^u = 98$  degrees of freedom at the high-fidelity level when discretized through FEM. As far as boundary and initial conditions are concerned, the following are taken into account: homogeneous Dirichlet conditions on the right portion



**Figure 4.** *Test 1.* Flow control. Mesh exploited to generate high-fidelity snapshots both in steady and time-dependent flow control test cases, along with the boundaries considered in Navier-Stokes equations. In particular,  $\Gamma_{\text{in}}$  in red is the inflow boundary where the fluid enters the channel,  $\Gamma_{\text{out}}$  in blue is the outflow,  $\Gamma_{\text{walls}}$  in black are the walls bounding the channel from the top and the bottom,  $\Gamma_{\text{obs}}$  in green and the control region  $\Gamma_c$  in magenta are the two portions of the obstacle boundary.

of the boundary  $\Gamma_{\text{obs}}$ , homogeneous Neumann boundary conditions on the outflow  $\Gamma_{\text{out}}$ , free-slip boundary conditions on the upper and lower walls  $\Gamma_{\text{walls}}$ , and homogeneous initial conditions for  $t = 0$  seconds. On the inflow boundary  $\Gamma_{\text{in}}$ , Dirichlet boundary conditions are imposed with velocity datum given by

$$\mathbf{v}_{\text{in}} = \|\mathbf{v}\|_{\text{in}} [\cos(\alpha_{\text{in}}), 0.01x_2(20 - x_2) \sin(\alpha_{\text{in}})]^\top$$

where  $\alpha_{\text{in}}$  (radians) denotes the angle of attack and  $\|\mathbf{v}\|_{\text{in}}$  ( $ms^{-1}$ ) the inflow intensity.

In this setting, we aim to minimize the energy dissipated by the fluid, that can be expressed as the  $L^2$  norm of the velocity gradient. The loss function is enriched by two regularizing terms dealing with the  $L^2$  norm of  $\mathbf{u}$  and its gradient, in order to stabilize the control problem and to avoid extremely expensive and energetic optimal control strategies. In the end, the selected loss function is

$$J(\mathbf{v}(\mathbf{x}, t), \mathbf{u}(\mathbf{x}, t)) = \int_0^T \int_{\Omega} \|\nabla \mathbf{v}(\mathbf{x}, t)\|^2 d\Omega dt + \beta \int_0^T \int_{\Gamma_c} \|\mathbf{u}(\mathbf{x}, t)\|^2 d\Gamma_c dt + \beta_g \int_0^T \int_{\Gamma_c} \|\nabla \mathbf{u}(\mathbf{x}, t)\|^2 d\Gamma_c dt$$

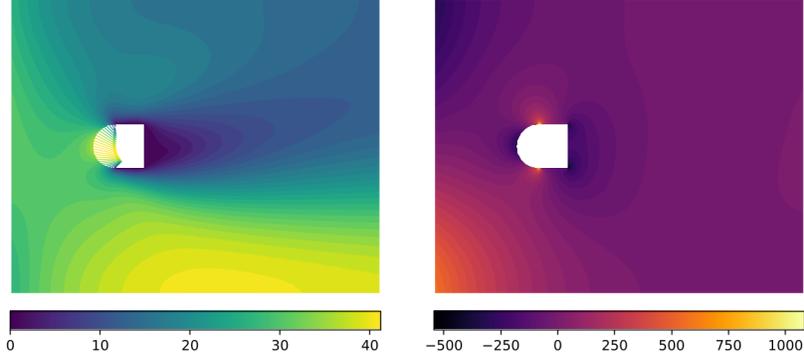
where the time integrals are considered only for time-dependent settings and the coefficients  $\beta$  and  $\beta_g$  are introduced to balance the integrals in the cost functional that may have different magnitudes. In particular,  $\beta = \beta_g = 10^{-2}$  in order to prioritize the minimization of energy dissipation. Note that, after discretizing Equation (7) through FEM, it is possible to obtain the discretized cost functional  $J_h$  introduced in Section 2.

#### 4.1.1. Flow control for steady Navier-Stokes equations

We first consider the steady Navier-Stokes equations, reported here for the sake of completeness

$$\begin{cases} -\nu \Delta \mathbf{v}(\mathbf{x}) + (\mathbf{v}(\mathbf{x}) \cdot \nabla) \mathbf{v}(\mathbf{x}) + \nabla p(\mathbf{x}) = 0 & \text{in } \Omega \\ \text{div } \mathbf{v}(\mathbf{x}) = 0 & \text{in } \Omega \\ \mathbf{v}(\mathbf{x}) = \mathbf{u}(\mathbf{x}) & \text{on } \Gamma_c \\ \mathbf{v}(\mathbf{x}) = \mathbf{0} & \text{on } \Gamma_{\text{obs}} \\ \mathbf{v}(\mathbf{x}) = \mathbf{v}_{\text{in}}(\boldsymbol{\mu}_s) & \text{on } \Gamma_{\text{in}} \\ \mathbf{v}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0 & \text{on } \Gamma_{\text{walls}} \\ (\nu \nabla \mathbf{v}(\mathbf{x}) - p(\mathbf{x}) \mathbf{n}(\mathbf{x})) \cdot \mathbf{t}(\mathbf{x}) = 0 & \text{on } \Gamma_{\text{walls}} \\ (\nu \nabla \mathbf{v}(\mathbf{x}) - p(\mathbf{x}) \mathbf{n}(\mathbf{x})) \cdot \mathbf{n}(\mathbf{x}) = 0 & \text{on } \Gamma_{\text{out}} \end{cases}$$

with  $\nu = 10m^2s^{-1}$ . The two parameters characterizing the inflow are regarded as scenario parameters, that is  $\boldsymbol{\mu}_s = [\alpha_{\text{in}}, \|\mathbf{v}\|_{\text{in}}]^\top$ . Following the procedure detailed in Section 3, we generate  $N_s = 50$  optimal pairs related to random inflow intensities and angles of attack uniformly sampled in the intervals  $(10.0, 130.0)ms^{-1}$



**Figure 5.** *Test 1.1.* Steady flow control. Optimal control, velocity and pressure obtained for  $\|\mathbf{v}_{\text{in}}\| = 31.16\text{ms}^{-1}$  and  $\alpha_{\text{in}} = -0.53$  radians through the high-fidelity OCP solver. The velocity on  $\Omega$  is depicted through a scalar field with colours corresponding to its norm, while the control on  $\Gamma_c$  is represented through a vector field.

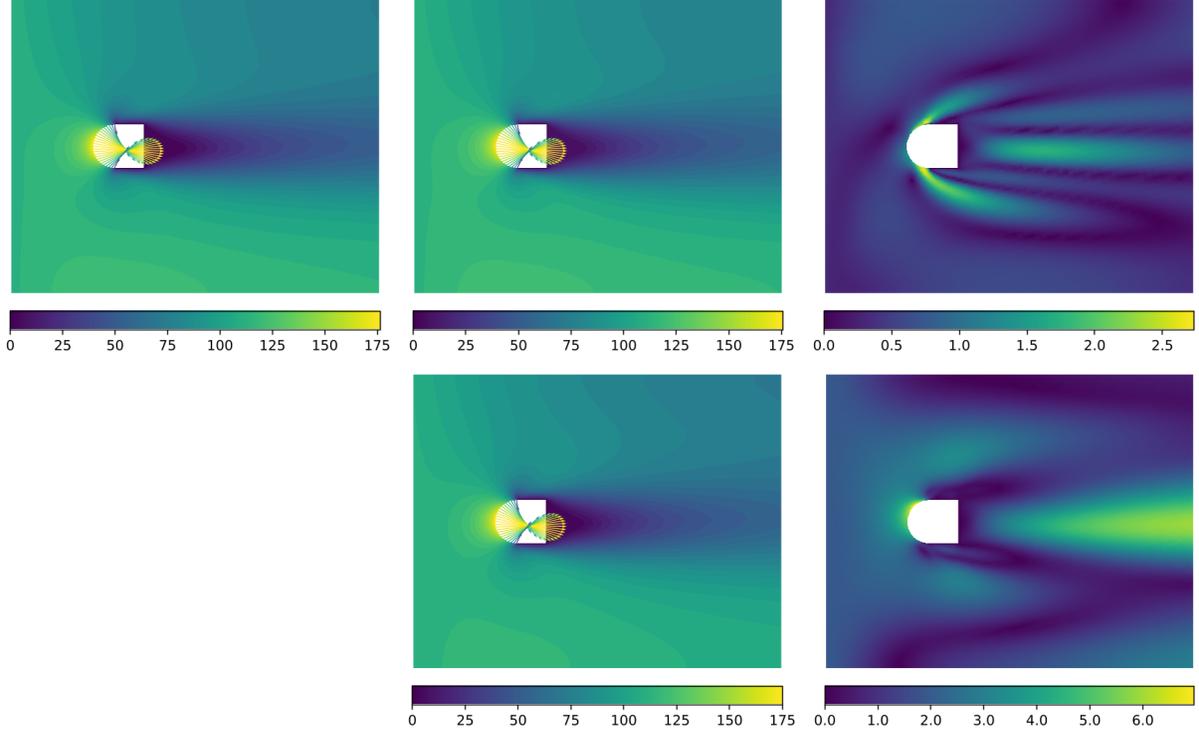
and  $(-1.0, 1.0)$  radians, respectively. The Reynolds number is thus very low and varies between 3 and 39. The forward resolution of Equation (7) is performed through the Newton method implemented in `fenics` [1], while optimal state and control snapshots are calculated with an OCP solver based on finite elements and the adjoint method implemented in `dolfin-adjoint` [46]. In particular, the high-fidelity solver takes, on average, 11 minutes to retrieve every snapshot. Out of all simulated data, 40 optimal pairs are exploited for dimensionality reduction and to train neural networks (training set), while the remaining ones are used for testing purposes (test set). The optimal snapshot generated for  $\|\mathbf{v}_{\text{in}}\| = 31.16\text{ms}^{-1}$  and  $\alpha_{\text{in}} = -0.53$  radians is shown in Figure 5.

To build a faster but still reliable OCP solver following the architecture presented in Section 3, we start reducing the snapshots dimensionality through POD. In particular, when reducing vector fields, POD is applied component-wise, i.e. the  $x_1$  and  $x_2$  components of the velocity are reduced separately, while considering the same number of modes for the two components. In this test case, nonlinear reduction strategies based on autoencoders are not necessary thanks to a fast singular value decay that allows to achieve acceptable reduction errors with few POD modes. Specifically, considering  $N_{\mathbf{u}} = 6$  and  $N_{\mathbf{v}} = 10$ , the  $L^2$  mean relative error committed when trying to reconstruct test data is 0.73% on the state and 0.70% on the control. Top rows of Figure 6 and Figure 7 display the ground truth and the reconstructed velocity fields, that are the snapshots compressed and reconstructed through POD, related to two different choices of scenario parameters in the test set.

To rapidly retrieve the optimal state and control values for different inflow velocities and angles of attack, here we must learn the map from the scenario parameters onto the POD coefficients of the optimal pairs, thus following a POD-NN strategy. This is done considering a feed-forward neural network  $\varphi$  having 3 hidden layers with, respectively, 50, 100, 50 neurons and exploiting leaky Relu as activation function. After building and training the neural network in 26.84 seconds through `pytorch` utilities taking into account the L-BFGS optimization algorithm, the predicted test data are accurate up to a  $L^2$  mean relative error equal to 1.53% for the optimal state and 1.81% for the optimal control. The bottom rows of Figure 6 and Figure 7 display the ground truth optimal solutions related to two different choices of scenario parameters in the test set and the corresponding POD-NN predictions given by

$$\begin{bmatrix} \tilde{\mathbf{v}}_h \\ \tilde{\mathbf{u}}_h \end{bmatrix} = \begin{bmatrix} \nabla_y & 0 \\ 0 & \nabla_u \end{bmatrix} \varphi(\alpha_{\text{in}}, \|\mathbf{v}\|_{\text{in}})$$

To emphasize the importance of dimensionality reduction when tackling high-dimensional problems, we fit a neural network mapping the scenario parameters directly into the high-fidelity optimal state and control variables, resulting in an output dimension equal to  $N_h^{\mathbf{v}} + N_h^{\mathbf{u}} = 29368$ . To achieve a  $L^2$  mean relative error of about 4% for both state and control, which is acceptable but higher than the one obtained with POD-NN, a neural network involving 135 times more weights, resulting in a training time 400 times longer, with respect to the one required for POD-NN shall be required.



**Figure 6.** *Test 1.1.* Steady flow control. First row: high-fidelity optimal snapshot, POD reconstruction and reconstruction error corresponding to the test scenario parameters  $\|\mathbf{v}_{\text{in}}\| = 114.79\text{ms}^{-1}$  and  $\alpha_{\text{in}} = -0.17$  radians. Second row: POD-NN prediction and prediction error corresponding to the test scenario parameters  $\|\mathbf{v}_{\text{in}}\| = 114.79\text{ms}^{-1}$  and  $\alpha_{\text{in}} = -0.17$  radians. The velocity on  $\Omega$  is depicted through a scalar field with colors corresponding to its norm, while the control on  $\Gamma_c$  is represented through a vector field.

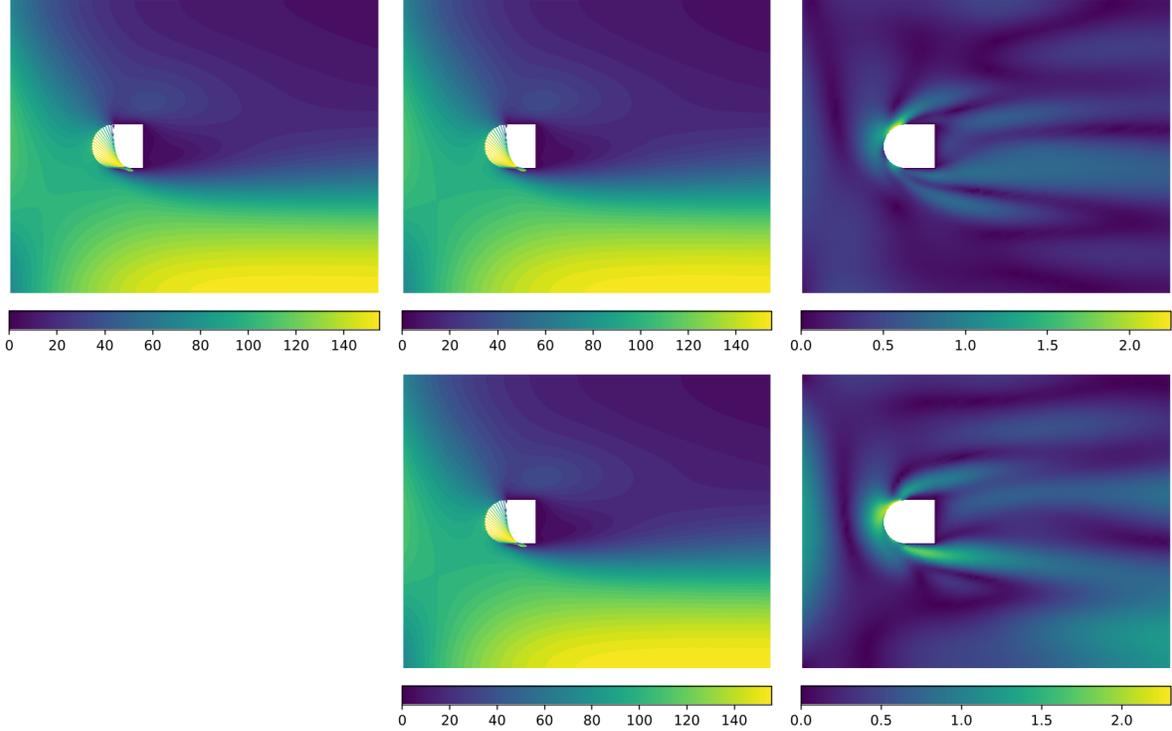
Figure 8 proposes a comparison between the reconstruction errors offered by POD and the prediction errors committed by POD-NN for different latent dimensions, while keeping fixed the other hyperparameters such as the  $\varphi$  architecture and the training set dimension. The decay of the POD reconstruction errors on test data, both for the state and the control, are superlinear. The POD-NN errors on test data are higher than those provided by POD due to the approximation given by the neural network  $\varphi$ , while keeping similar superlinear decays with respect to the latent dimension both for the control and for the state approximation. Furthermore, it is possible to assess a saturation in the POD-NN performance for large latent dimensions.

The POD-NN architecture is then exploited to solve the parametric OCP for new values of scenario parameters unseen during the data generation and training phases. In particular, we consider  $\|\mathbf{v}_{\text{in}}\| = 100\text{ms}^{-1}$  and several angles of attack in a grid spanning  $(-1, 1)$  radians. Figure 9 shows the results obtained for three different angles of attack considered. Each result is computed online in less than 0.002 seconds thanks to a forward pass through the neural network  $\varphi$  and the POD decoding in order to predict full-order optimal pairs.

Figure 10 presents, instead, two different analyses that are helpful in understanding the benefits and behavior of the proposed method. In particular, the left panel investigates the trend of  $L^2$  mean relative test error for different training set dimensions  $N_{\text{train}}$  keeping fixed the other hyperparameters, while a similar analysis is repeated in the right panel for different  $\varphi$  complexities, i.e. for different number of neural network weights  $N_{\text{weights}}$ . The superlinear decay obtained in the first case suggests that an increase in the number of training data has a higher impact on the goodness of fit of the proposed model rather than an increase of the number of  $\varphi$  weights, which entails a sublinear decay.

#### 4.1.2. Flow control for time-dependent Navier-Stokes equations

We now consider the time-dependent case detailed in Equation (7) with kinematic viscosity  $\nu = 1\text{m}^2\text{s}^{-1}$  and  $\|\mathbf{v}_{\text{in}}\| = 10\text{ms}^{-1}$  in order to focus on a low Reynolds number ( $\text{Re} = 30$ ). The scenario parameters

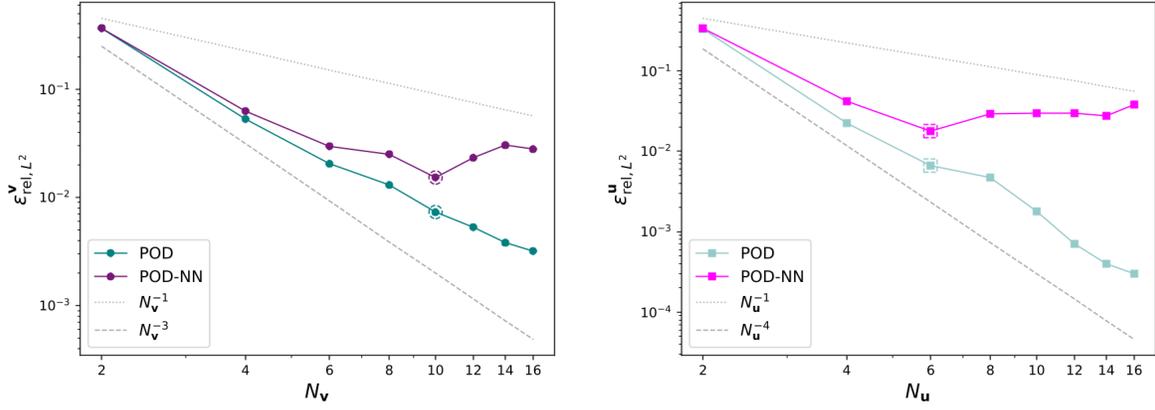


**Figure 7.** *Test 1.1.* Steady flow control. First row: high-fidelity optimal snapshot, POD reconstruction and reconstruction error corresponding to the test scenario parameters  $\|\mathbf{v}_{\text{in}}\| = 112.71\text{ms}^{-1}$  and  $\alpha_{\text{in}} = -0.88$  radians. Second row: POD-NN prediction and prediction error corresponding to the test scenario parameters  $\|\mathbf{v}_{\text{in}}\| = 112.71\text{ms}^{-1}$  and  $\alpha_{\text{in}} = -0.88$  radians. The velocity on  $\Omega$  is depicted through a scalar field with colours corresponding to its norm, while the control on  $\Gamma_c$  is represented through a vector field.

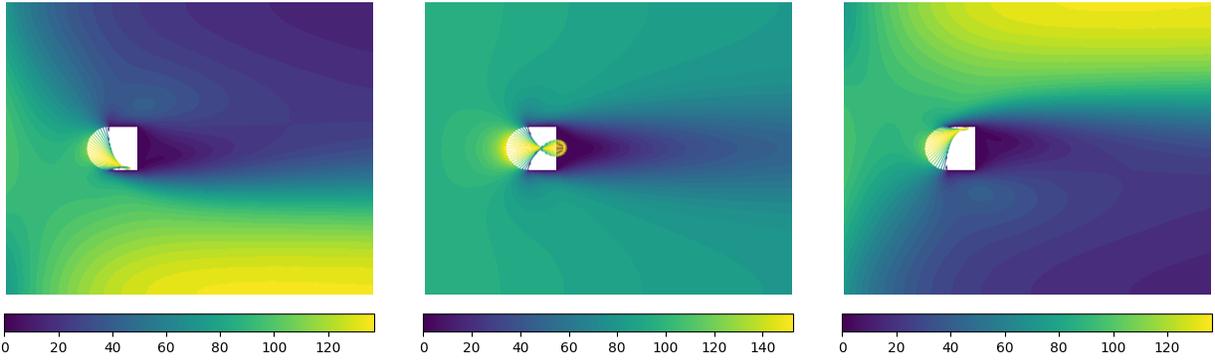
are therefore the inflow angle  $\alpha_{\text{in}}$  and, as always happens in time-dependent settings, the time variable  $t$ . We generate 22 time-dependent optimal trajectories through the full-order OCP solver based on FEM implemented in `dolfin-adjoint`, solving the forward problem in Equation (7) with the incremental Chorin-Temam projection method. Every trajectory considers a constant angle of attack randomly sampled in the interval  $(-1.0, 1.0)$  radians and it is recorded in 10 different time steps uniformly distributed in the interval  $[0.05, 0.5]$  seconds, i.e. the final time is  $T = 0.5$  seconds and the time step is equal to 0.05 seconds. The snapshots are then shuffled and divided into training set (180 snapshots) and test set (40 snapshots). Every optimal trajectory, which consists of 10 different snapshots in time, is computed solving the nonlinear time-dependent full-order OCP requiring on average a computational time of 20 minutes. An example of optimal control, velocity and pressure related to  $\alpha_{\text{in}} = 0.45$  radians and  $t = 0.5$  seconds is provided in Figure 11: for visualization purposes, the scale considered for the control vectors is 10 times bigger with respect to the figures in Section 4.1.1.

In this setting, the  $L^2$  mean relative test errors entailed when reconstructing state and control through POD shows an exponential decay, as visible in Figure 12. However, due to a slow decay in correspondence of small latent dimensions, we need a higher number of POD modes with respect to the test case in Section 4.1.1 to achieve the same accuracy: specifically, with  $N_{\mathbf{u}} = 10$  and  $N_{\mathbf{v}} = 18$ , the  $L^2$  mean relative reconstruction errors on test data are equal to 0.26% on the state and 0.63% on the control.

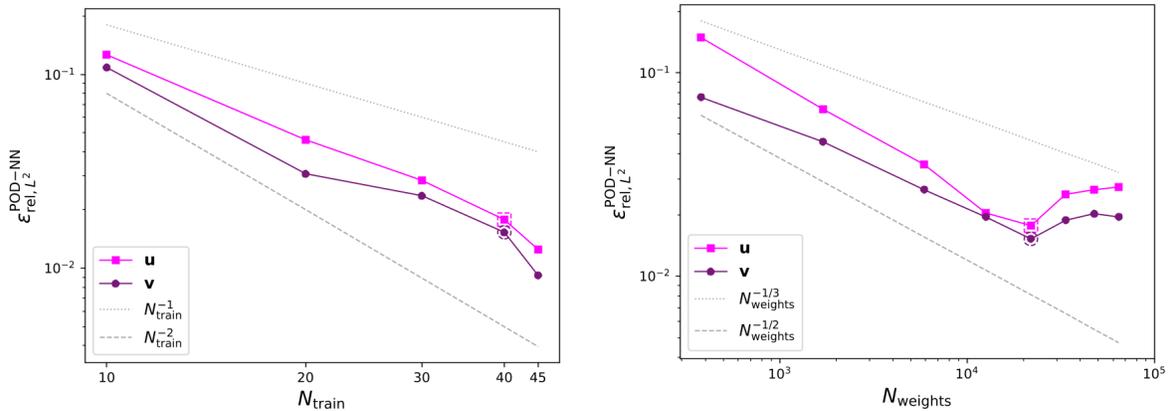
Nonlinear reduction strategies may be therefore considered to further reduce the problem dimensionality and allow for a lighter, more accurate and easily trainable neural network  $\varphi$  with smaller output dimensions. As proposed by [22] and resumed in Section 3, it is possible to combine an initial linear reduction step through POD and a nonlinear autoencoder to further compress the data. In particular, we consider 80 and 140 POD modes for the control and the state variables, respectively, which returns very low reconstruction errors on test data. Thanks to the nonlinear reduction provided by fully-connected autoencoders, the latent dimensions are then shrunk down to  $N_{\mathbf{u}} = N_{\mathbf{v}} = 6$ . The encoder responsible for the reduction of the state



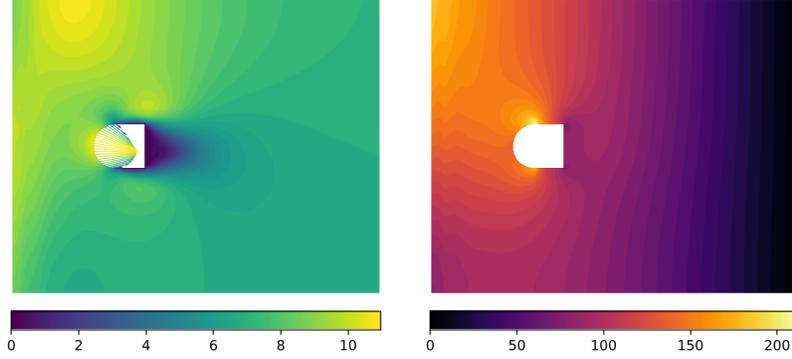
**Figure 8.** *Test 1.1.* Steady flow control. Left:  $L^2$  mean relative error on  $\mathbf{v}$  committed by POD (reconstruction error) and POD-NN (prediction error) for different latent dimensions  $N_{\mathbf{v}}$ . Right:  $L^2$  mean relative error on  $\mathbf{u}$  committed by POD (reconstruction error) and POD-NN (prediction error) for different latent dimensions  $N_{\mathbf{u}}$ . The latent dimensions selected in the numerical example are highlighted with dashed markers.



**Figure 9.** *Test 1.1.* Steady flow control. Optimal state and control provided by POD-NN corresponding to the scenario parameters  $\|\mathbf{v}_{\text{in}}\| = 100 \text{ms}^{-1}$  and  $\alpha_{\text{in}} = -0.75, 0, 0.75$  radians. The velocity on  $\Omega$  is depicted through a scalar field with colours corresponding to its norm, while the control on  $\Gamma_c$  is represented through a vector field.



**Figure 10.** *Test 1.1.* Steady flow control. Left:  $L^2$  mean relative test error decay with respect to training set dimension  $N_{\text{train}}$ . Right:  $L^2$  mean relative test error decay with respect to the number of  $\varphi$  weights  $N_{\text{weights}}$ . The hyperparameters selected in the numerical example are highlighted with dashed markers.

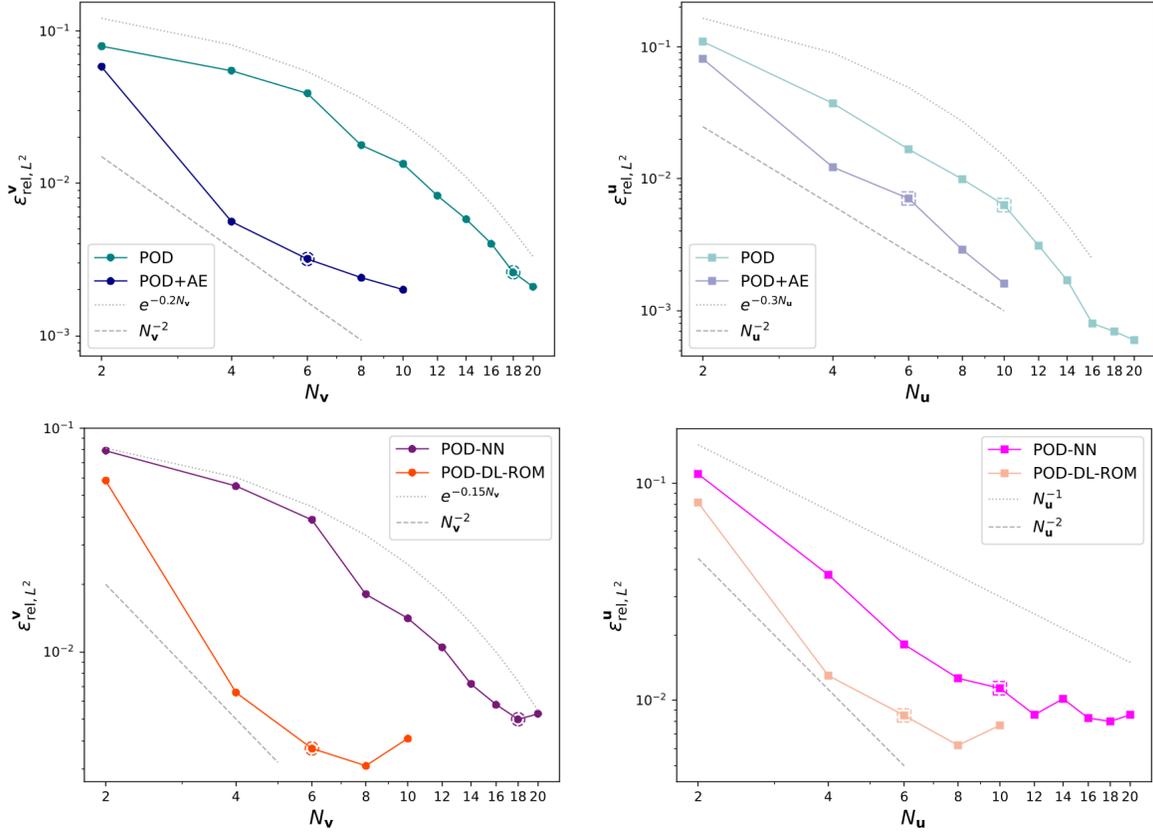


**Figure 11.** *Test 1.2.* Time-dependent flow control. Optimal control, velocity and pressure obtained for  $\alpha_{\text{in}} = 0.45$  radians and  $t = 0.5$  seconds through a high-fidelity full-order OCP solver. The velocity on  $\Omega$  is depicted through a scalar field with colours corresponding to its norm, while the control on  $\Gamma_c$  is represented through a vector field.

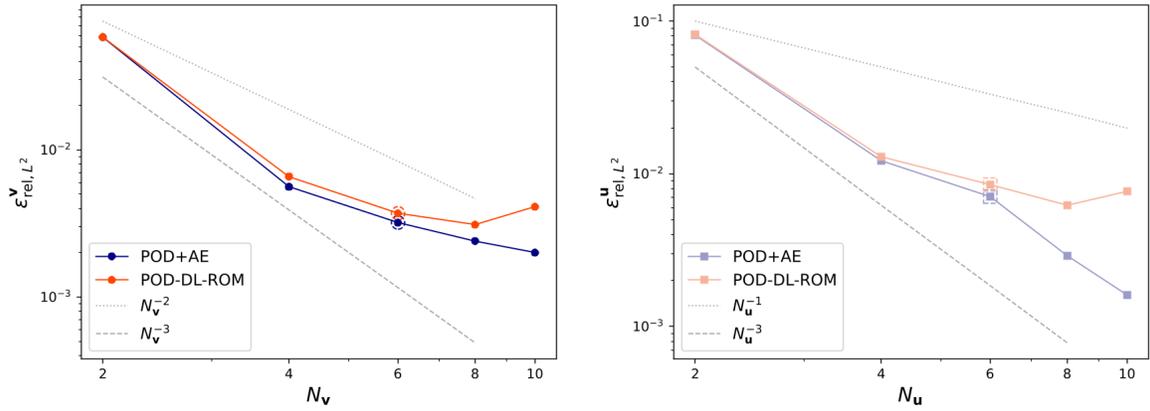
POD modes has 1 hidden layer with 50 neurons, while the corresponding decoder exploits 2 hidden layers with 50 neurons each. As far as the compression of control POD modes is concerned, the same number of hidden layers are considered, but the number of neurons per layer is decreased to 25 thanks to the smaller input-output dimensions. Thanks to the joint reduction provided by POD and autoencoders, it is possible to achieve a very low-dimensional latent representation of the system variables with high-accuracy: indeed, the  $L^2$  mean relative reconstruction errors on test data are 0.32% for the state and 0.70% for the control. The first rows of Figure 14 and Figure 15 display the ground truth and the reconstructed optimal pair related to two different test scenarios, confirming the high precision of the reduction procedure. The first row of Figure 12 shows, instead, the  $L^2$  mean relative errors committed when reconstructing test data by POD and POD+AE, that is the combination of POD and autoencoders, for different latent dimensions, while keeping fixed the other hyperparameters. It is possible to assess that a nonlinear reduction technique allows to achieve more accurate results while considering smaller latent dimensions, entailing a lighter  $\varphi$  architecture that is faster to train and to evaluate online.

To rapidly compute the optimal state and control for different time instants and inflow velocities, we model the map from these parameters onto the reduced optimal pairs through a feed-forward neural network  $\varphi$ , thus considering a POD-DL-ROM approach. In particular,  $\varphi$  consists of 2 hidden layers with 150 neurons each, and leaky Relu is exploited as activation function. After training the neural network with the L-BFGS optimization algorithm, the  $L^2$  mean relative errors on test data are equal to 0.37% for the optimal state and 0.85% for the optimal control. The computational time required to train the autoencoders and the neural network  $\varphi$  amounts to 4 minutes and 11.52 seconds. The second row of Figure 12 compares the  $L^2$  mean relative errors committed when predicting test data through POD-NN and POD-DL-ROM for different latent dimensions: coherently with the corresponding reconstruction errors, a POD-DL-ROM allows to obtain more accurate results with smaller latent dimensions. Moreover, Figure 13 shows the reconstruction and prediction errors of POD-DL-ROM: while both show a superlinear decay with respect to the latent dimension, the prediction errors are slightly higher than the reconstruction ones due to the approximation of the parameter-to-solution map through the neural network  $\varphi$ . The second rows of Figure 14 and Figure 15 display, instead, the predictions obtained by POD-DL-ROM compared to the ground truth optimal solutions related to two different choices of test scenario parameters: both the control and the velocity are retrieved with high accuracy and no substantial differences can be assessed between ground truth data and POD-DL-ROM predictions.

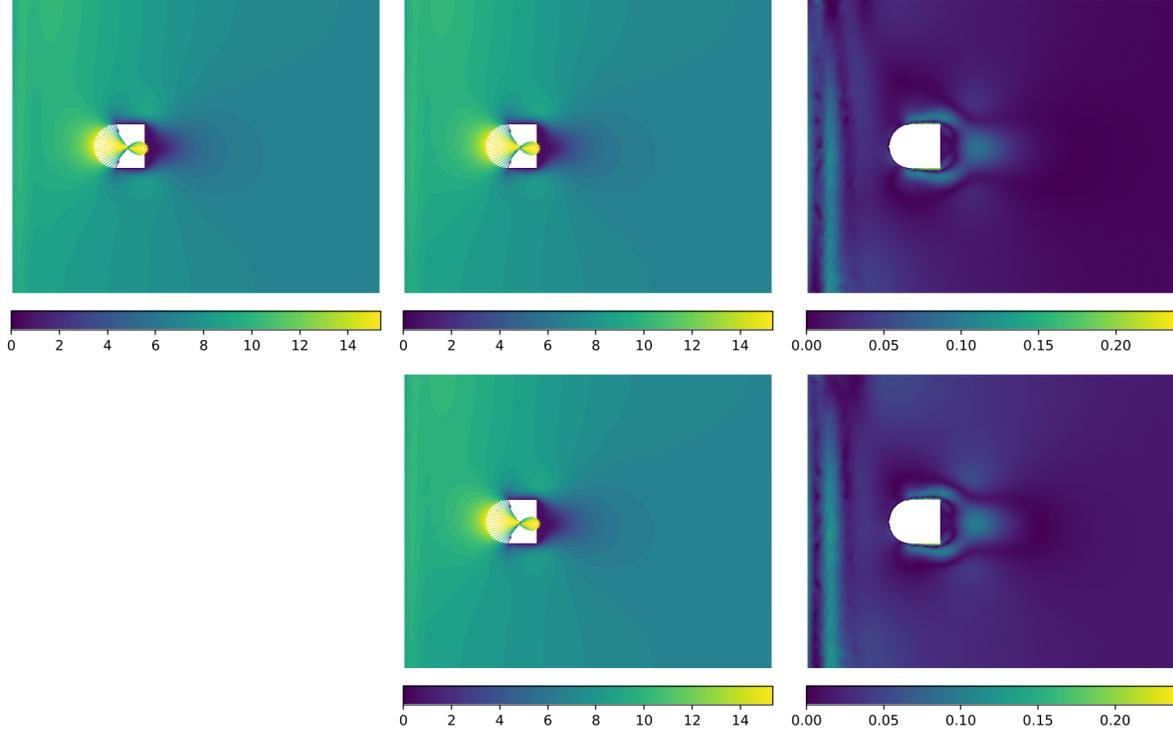
The POD-DL-ROM architecture is then exploited to solve the time-dependent parametrized OCP for new and unseen time instants and angles of attack sampled in the intervals  $[0.05, 0.5]$  seconds and  $(-1.0, 1.0)$  radians, respectively. Figure 16 shows the time-dependent optimal pairs computed via POD-DL-ROM when considering  $\alpha_{\text{in}} = 0.5$  radians at three different time instants. Note that, in order to find each of these solutions, we just need a forward pass of the neural network  $\varphi$  and a decoding step through the autoencoders and POD that lasts at most 0.01 seconds. Note also that, the forward pass  $\varphi_D(\varphi(t^{\text{new}}, \boldsymbol{\mu}_s^{\text{new}}))$  infers the time-dependent optimal pair – i.e. the minimizer of the loss function computed in the interval  $(0, T)$  with



**Figure 12.** *Test 1.2.* Time-dependent flow control. Top:  $L^2$  mean relative reconstruction error on  $\mathbf{v}$  and  $\mathbf{u}$  committed by POD and POD+AE for different latent dimensions  $N_{\mathbf{v}}$  and  $N_{\mathbf{u}}$ . Bottom:  $L^2$  mean relative prediction error on  $\mathbf{v}$  and  $\mathbf{u}$  committed by POD-NN and POD-DL-ROM for different latent dimensions  $N_{\mathbf{v}}$  and  $N_{\mathbf{u}}$ . The latent dimensions selected in the numerical example are highlighted with dashed markers.



**Figure 13.** *Test 1.2.* Time-dependent flow control. Left:  $L^2$  mean relative error on  $\mathbf{v}$  committed by POD+AE (reconstruction error) and POD-DL-ROM (prediction error) for different latent dimensions  $N_{\mathbf{v}}$ . Right:  $L^2$  mean relative error on  $\mathbf{u}$  committed by POD+AE (reconstruction error) and POD-DL-ROM (prediction error) for different latent dimensions  $N_{\mathbf{u}}$ . The latent dimensions selected in the test case are highlighted with dashed markers.



**Figure 14.** *Test 1.2.* Time-dependent flow control. First row: high-fidelity optimal snapshot, POD+AE reconstruction and reconstruction error corresponding to the test scenario parameters  $t = 0.25$  seconds and  $\alpha_{\text{in}} = 0.34$  radians. Second row: POD-DL-ROM prediction and prediction error corresponding to the test scenario parameters  $t = 0.25$  seconds and  $\alpha_{\text{in}} = 0.34$  radians. The velocity on  $\Omega$  is depicted through a scalar field with colours corresponding to its norm, while the control on  $\Gamma_c$  is represented through a vector field.

fixed final time  $T$  – evaluated at  $t = t^{\text{new}}$ .

#### 4.2. Active thermal cooling

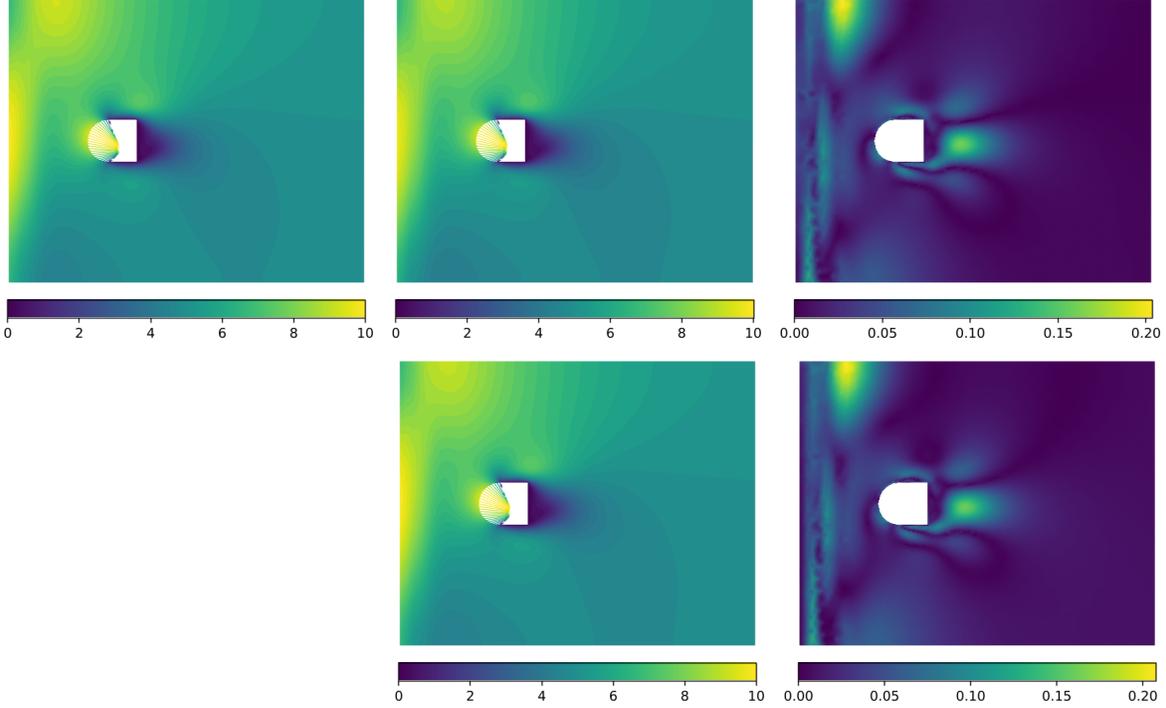
This section presents the numerical results obtained by applying the proposed real-time OCP solver to an active thermal cooling problem. We consider a square domain  $\Omega = (-1, 1) \times (-1, 1)$  properly discretized thanks to `gmsh` utilities with the presence of a circular obstacle centered at  $(0, 0)$  with radius equal to 0.15. The physical system under investigation is described by the steady heat equation, that is

$$\begin{cases} -\nu \Delta y(\mathbf{x}) = s(\mathbf{x}; \boldsymbol{\mu}_s) + u(\mathbf{x}) \mathbb{1}_{\Omega_c} & \text{in } \Omega \\ -\nu \nabla y(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = y(\mathbf{x}) & \text{on } \Gamma_d \\ -\nu \nabla y(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = \gamma(y(\mathbf{x}) - y_{\text{ext}}) & \text{on } \Gamma_{\text{obs}} \end{cases} \quad (8)$$

where  $\nu = 1m^2s^{-1}$  is the fixed material-specific thermal diffusivity and  $\mathbf{n}$  is the unit vector normal to the domain boundary. The state variable  $y$  is a scalar-quantity denoting the temperature (Kelvin degrees) at every domain location: since linear finite elements are considered, the vector  $\mathbf{y}_h$  collects the temperature values at these mesh nodes and in this case the dimension of the state space is  $N_h^y = 51665$ . The obstacle is heated by an external source  $s$  ( $Ks^{-1}$ ) modeled through the Gaussian function

$$s = 5000 \exp\{-40[(x_1 - x_1^s)^2 + (x_2 - x_2^s)^2]\}$$

centered at the point  $(x_1^s, x_2^s)$ . To allow heat exchange between the obstacle and the surrounding environment, the Newton’s law of cooling is taken into account as boundary condition on the obstacle boundary  $\Gamma_{\text{obs}}$ , where  $y_{\text{ext}} = 125K$  is the external temperature, that is proportional to the product between the intensity and the variance of  $s$ , and  $\gamma = 1ms^{-1}$  is the fixed heat transfer coefficient multiplied by the volumetric heat capacity.



**Figure 15.** *Test 1.2.* Time-dependent flow control. First row: high-fidelity optimal snapshot, POD+AE reconstruction and reconstruction error corresponding to the test scenario parameters  $t = 0.45$  seconds and  $\alpha_{\text{in}} = 0.85$  radians. Second row: POD-DL-ROM prediction and prediction error corresponding to the test scenario parameters  $t = 0.45$  seconds and  $\alpha_{\text{in}} = 0.85$  radians. The velocity on  $\Omega$  is depicted through a scalar field with colours corresponding to its norm, while the control on  $\Gamma_c$  is represented through a vector field.

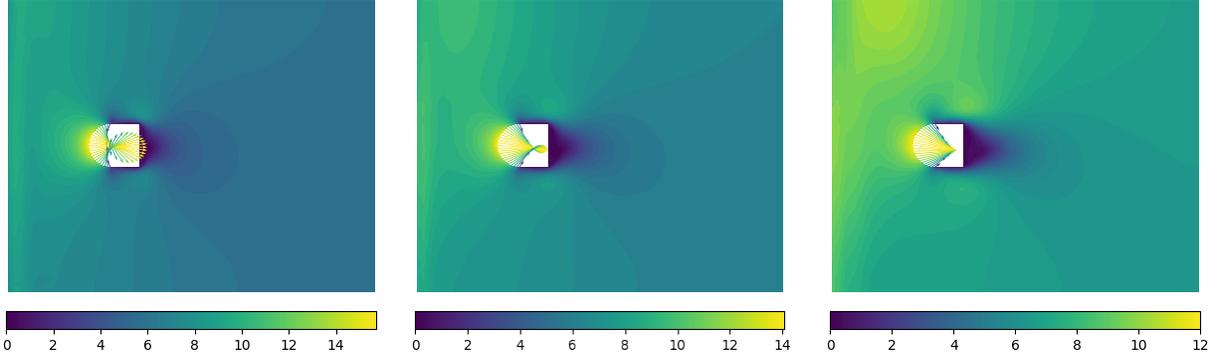
Moreover, a homogeneous Robin boundary condition is applied on the external boundary  $\Gamma_d$  in order to approximate an unbounded domain to first-order, as proposed by [70]. The problem setting is visible in the left panel of Figure 17.

In this context, our objective is to keep the temperature on the obstacle constant and equal, for simplicity, to the reference value  $0K$  while considering different locations of the external heat source  $s$ . The scenario parameters are therefore the coordinates of the source center: for convenience, the polar coordinates are taken into account, that is  $\boldsymbol{\mu}_s = [\vartheta_s, r_s]^\top$ , where  $x_1^s = r_s \cos(\vartheta_s)$  and  $x_2^s = r_s \sin(\vartheta_s)$ . To solve this parametric OCP, we consider an active control action, that may be interpreted as a cooler. In particular,  $u$  is a space-varying source term taking values in  $\Omega_c$ , that is the annular region with inner and outer radius equal respectively to 0.2 and 0.3, as visible in the left panel of Figure 17. The control variable is strongly high-dimensional since  $N_h^u = 7721$  degrees of freedom are exploited to discretize the control region. In order to achieve the target temperature on  $\Gamma_{\text{obs}}$  while considering feasible and regular control actions, the cost function is defined as

$$J(y(\mathbf{x}), u(\mathbf{x})) = \int_{\Gamma_{\text{obs}}} |y(\mathbf{x})|^2 d\Gamma_{\text{obs}} + \beta \int_{\Omega_c} |u(\mathbf{x})|^2 d\Omega_c + \beta_g \int_{\Omega_c} \|\nabla u(\mathbf{x})\|^2 d\Omega_c$$

where the coefficients  $\beta$  and  $\beta_g$  are set equal to  $1e-8$ . To better understand the problem at hand, Figure 17 shows a comparison between the uncontrolled and the controlled temperature fields for a particular scenario: it is possible to assess that an optimal space-varying control strategy allows to cool the obstacle and keep its temperature equal to the reference value.

As far as the data generation process is concerned, 100 optimal snapshots are computed through the OCP high-fidelity solver based on FEM provided by `dolphin-adjoint` considering random source coordinates  $\vartheta_s, r_s$  uniformly sampled in the intervals  $(-\frac{\pi}{2}, \frac{\pi}{2})$  radians and  $(0.4, 0.9)$ . The full-order solver requires, on average, 4 minutes and 22 seconds to retrieve one optimal snapshot while exploiting L-BFGS as optimization algorithm and a tolerance equal to  $1e-10$ . The data are then divided into training set (80 snapshots) and test set (20 snapshots) for model evaluation purposes.



**Figure 16.** *Test 1.2.* Time-dependent flow control. Optimal state and control provided by POD-DL-ROM corresponding to the scenario parameters  $\alpha_{\text{in}} = 0.5$  radians and  $t = 0.159, 0.276, 0.453$  seconds.

In this context, dimensionality reduction is strongly recommended due to the huge dimensionality of state and control variables. In particular, a linear reduction scheme based on POD is enough to reduce the optimal control: indeed, considering  $N_u = 7$ , the  $L^2$  mean relative reconstruction error obtained on test data results equal to 0.17%. The second row of Figure 18 displays the singular value decay and the two most-energetic POD modes – i.e. the ones associated with the two largest singular values – that capture the most of the control snapshots variance.

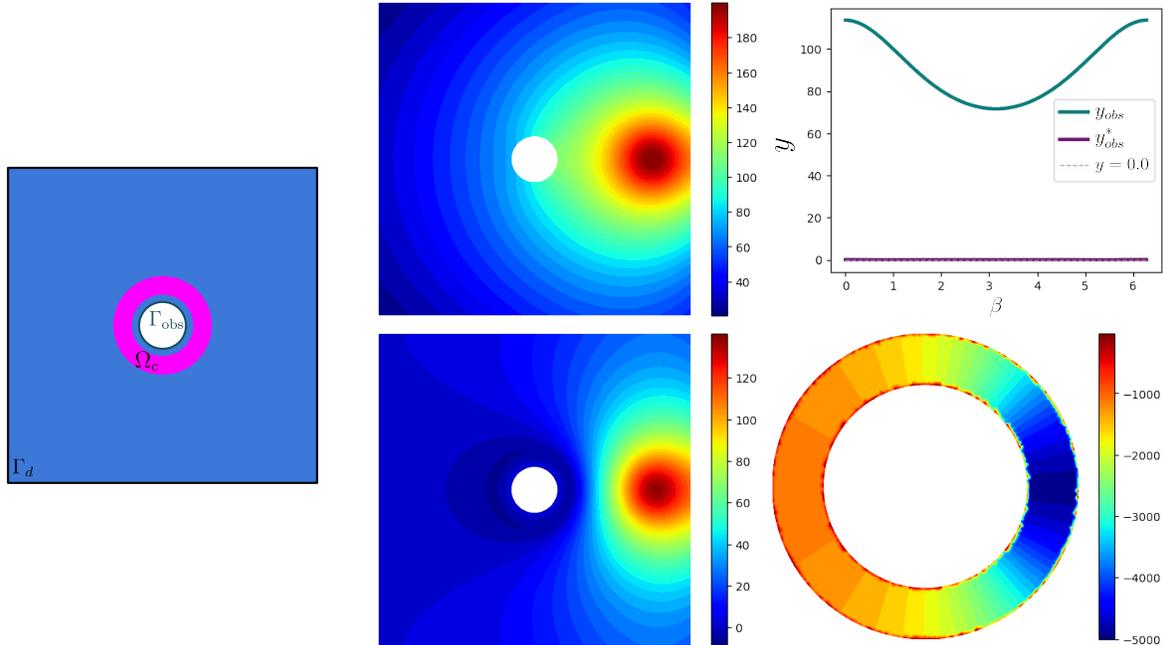
While POD is accurate to project the control snapshots in a low-dimensional latent space, the same is not true for state snapshots compression. Indeed, due to a slower singular values decay, as visible in Figure 18, more than 37 POD modes would be necessary in order to achieve acceptable reconstruction errors. Therefore, following the POD-DL-ROM strategy, 72 POD modes are initially computed and then compressed thanks to a nonlinear autoencoder into a six-dimensional latent representation, that is  $N_y = 6$ . The encoder consists of 2 layers with, respectively, 70 and 30 neurons, while 3 layers having 30, 50 and 70 neurons are considered in the decoder architecture, with leaky Relu as activation function. After training the autoencoder with the L-BFGS optimization algorithm, the  $L^2$  mean relative reconstruction error committed on test data is equal to 2%. The first row in Figure 19 displays the ground truth and the reconstructed temperature field corresponding to the test scenario parameters  $\vartheta_s = -0.22$  radians and  $r_s = 0.44$ . The first row of Figure 20 shows, instead, the ground truth and the corresponding optimal control compressed and reconstructed through POD related to the same choice of scenario parameters.

As detailed in Section 3, the last ingredient required to obtain a non-intrusive real-time optimal solver for this parametric OCP is the parameter-to-solution map  $\varphi$ . Specifically,  $\varphi$  is modeled through a feed-forward neural network having 2 layers with 50 neurons each and exploiting leaky Relu as activation function. The input layer of  $\varphi$  is augmented considering, in addition to the scenario parameters  $\vartheta_s$  and  $r_s$ , meaningful quantities such as  $r_s \cos \vartheta_s$  and  $r_s \sin \vartheta_s$ . POD-DL-ROM is able to achieve a  $L^2$  mean relative error on test data equal to 3.83% for the optimal state and 1.10% for the optimal control. The overall training of  $\varphi$  and the autoencoder lasts 4 minutes and 41 seconds. Instead, the evaluation time required to predict the optimal pair related to a new scenario of interest is equal to, on average, 0.007 seconds, thus providing a solver 40000 times faster with respect to the full-order one. The bottom rows of Figure 19 and Figure 20 display a comparison between the ground truth optimal solutions  $\mathbf{y}_h, \mathbf{u}_h$  and the POD-DL-ROM predictions obtained by

$$\begin{bmatrix} \tilde{\mathbf{y}}_h \\ \tilde{\mathbf{u}}_h \end{bmatrix} = \varphi_D(\varphi(\vartheta_s, r_s))$$

related to two different choices of scenario parameters in the test set, where  $\varphi_D$  resumed the decoding actions applied separately to state and control.

The POD-DL-ROM architecture is then tested online with respect to new scenario parameters unseen during training. Figure 21 displays the results obtained for  $r_s = 0.65$  and three different values of  $\vartheta_s$ . On average, the temperature on the obstacle is equal to  $0.0011K$ , thus very close to the target value ( $0K$ ). The left panel of Figure 22 shows, instead, the optimal values of the loss function  $J_h(\tilde{\mathbf{y}}_h, \tilde{\mathbf{u}}_h; \boldsymbol{\mu}_s)$  evaluated at optimal states and controls computed for several scenarios in a uniform  $100 \times 100$  grid covering the



**Figure 17.** *Test 2.* Active thermal cooling. Left: subdomains and boundaries considered in the active thermal cooling problem. In particular,  $\Gamma_d$  in black is the external boundary,  $\Gamma_{obs}$  in green is the obstacle boundary and  $\Omega_c$  in magenta is the annular region around the obstacle where the control source takes place. Center: temperature fields computed in absence of control action (top) and considering optimal control values (bottom) with thermal source located at  $(0.75, 0.0)$ . Top right: comparison of the temperature profiles at the obstacle boundary points  $(0.15 \cos(\beta), 0.15 \sin(\beta))$  for different angles  $\beta$  in the controlled ( $y_{obs}^*$ ) and uncontrolled ( $y_{obs}$ ) setups. Bottom right: space-varying optimal control action considering the thermal source located at  $(0.75, 0.0)$ .

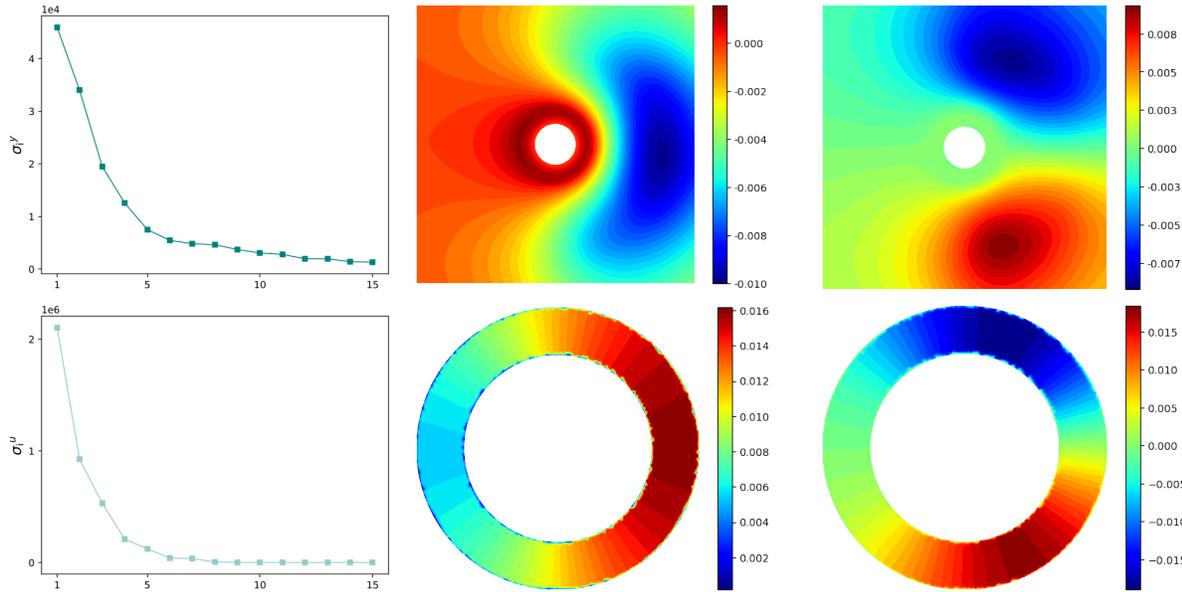
parameter space. Note that, since 10000 optimal pairs are required in order to perform this parametric analysis, a full-order solver would need approximately 30 days of computation. Instead, thanks to our fast-evaluable architecture, all the optimal solutions are computed in 2.14 seconds. From the loss function behaviour in the parameter space, it is clear that the closer the heat source to the obstacle – i.e. the smaller  $r_s$  – the higher the loss due to more energetic control terms required. The same analysis is performed looking at the optimal  $L^2$  norm of the temperature on the obstacle boundary, that is

$$\|\tilde{y}_h(\boldsymbol{\mu}_s)\|_{L^2(\Gamma_{obs})} = \sqrt{\int_{\Gamma_{obs}} \tilde{y}_h^2(\mathbf{x}; \boldsymbol{\mu}_s) d\Gamma_{obs}}$$

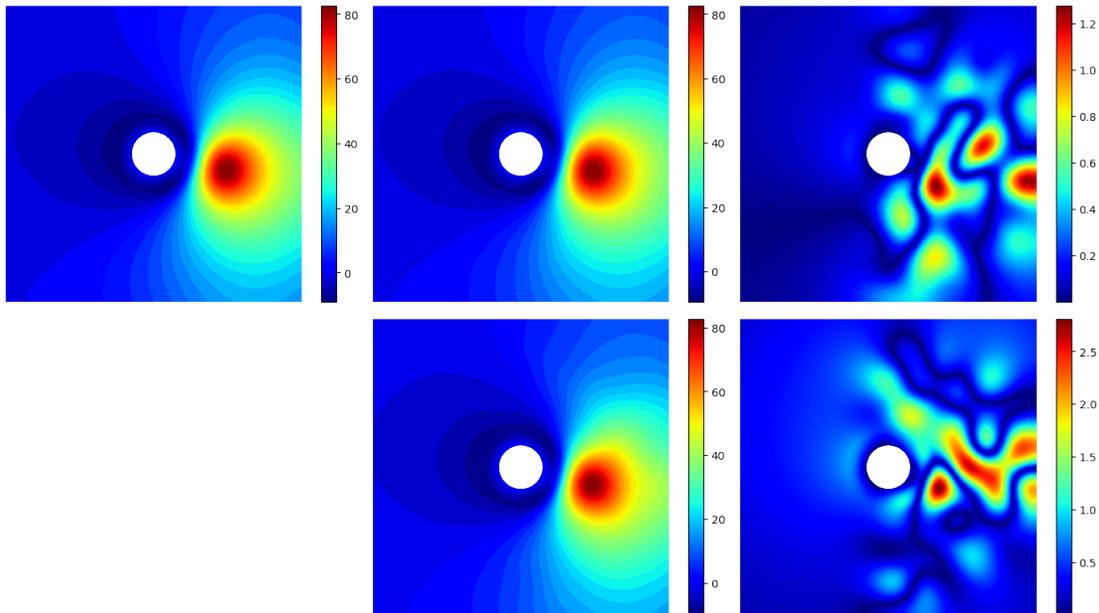
where  $\tilde{y}_h$  is the function associated with the FEM coefficients  $\tilde{\mathbf{y}}_h$ . As visible in the right panel of Figure 22, the control strategies predicted by POD-DL-ROM allow to attain obstacle temperatures close to the target in the whole parameter space and, thus, to properly cool the obstacle. In particular, since the loss function is a sum of terms being minimized, slightly higher obstacle temperatures are naturally obtained corresponding to more energetic control strategies.

## 5. Conclusions

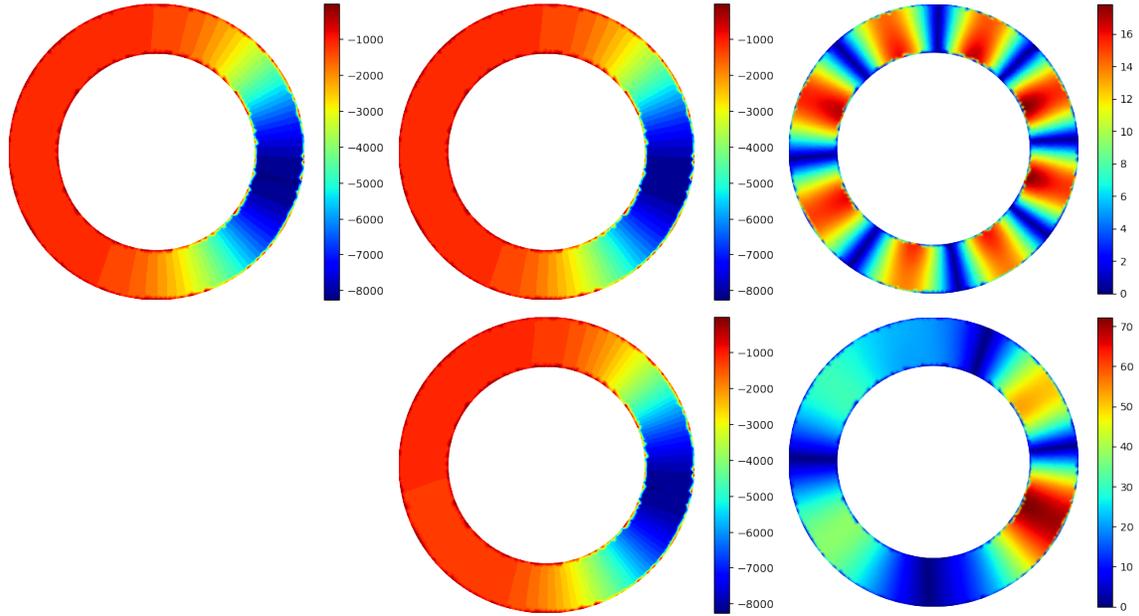
In this work we propose an efficient and reliable non-intrusive data-driven method to solve parametrized OCPs. Indeed, in all applications we considered, the mismatch between the predicted optimal solutions and the corresponding ground truth is always less than 4% when taking into account new unseen scenarios. Moreover, the online stage is really fast – it lasts at most 0.01 seconds – regardless of the problem complexity and dimension. Along with speed and accuracy, another key feature of the presented approach is its extreme versatility: in fact, as demonstrated throughout the numerical results in Section 4, it is possible to handle a broad spectrum of parametrized OCPs, ranging from scalar to vector, from linear to nonlinear and from steady to time-dependent problems.



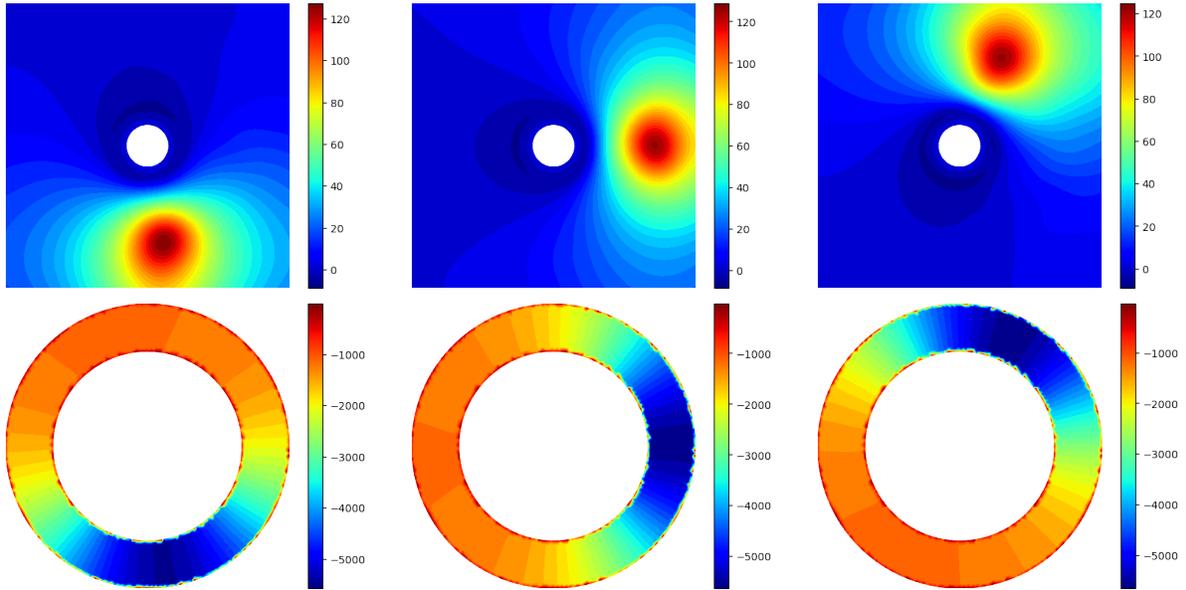
**Figure 18.** *Test 2.* Active thermal cooling. Left: decay of the singular values computed through SVD applied to the state (top) and control (bottom) snapshots matrices. Center: POD mode associated with the highest singular value of the state (top) and control (bottom). Right: POD mode associated with the second highest singular value of the state (top) and control (bottom).



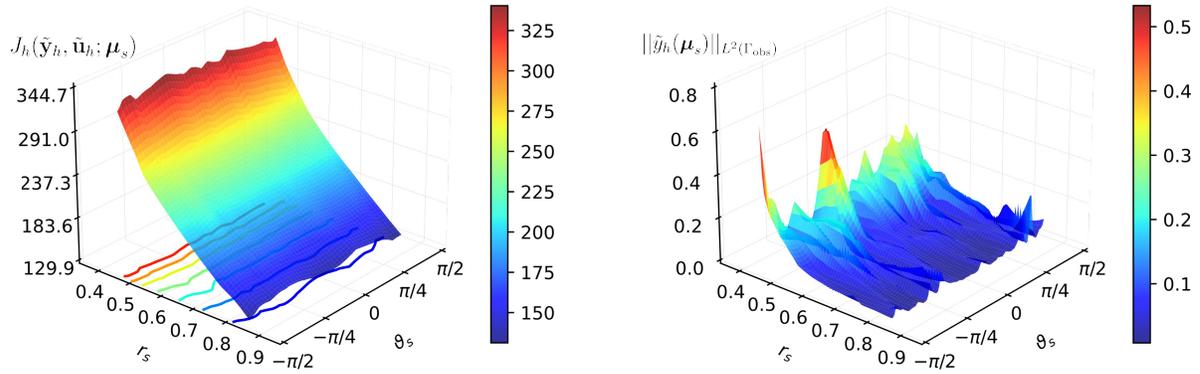
**Figure 19.** *Test 2.* Active thermal cooling. First row: high-fidelity optimal snapshot, POD+AE reconstruction and reconstruction error corresponding to the test scenario parameters  $\vartheta_s = -0.22$  radians and  $r_s = 0.44$ . Second row: POD-DL-ROM prediction and prediction error corresponding to the test scenario parameters  $\vartheta_s = -0.22$  radians and  $r_s = 0.44$ .



**Figure 20.** *Test 2.* Active thermal cooling. First row: high-fidelity optimal snapshot, POD reconstruction and reconstruction error corresponding to the test scenario parameters  $\vartheta_s = -0.22$  radians and  $r_s = 0.44$ . Second row: POD-DL-ROM prediction and prediction error corresponding to the test scenario parameters  $\vartheta_s = -0.22$  radians and  $r_s = 0.44$ .



**Figure 21.** *Test 2.* Active thermal cooling. Optimal state and control provided by POD-DL-ROM corresponding to the scenario parameters  $r_s = 0.65$  and  $\vartheta_s = -1.42, 0.03, 1.13$  radians.



**Figure 22.** *Test 2.* Active thermal cooling. Left: optimal loss function values computed online considering 10000 different scenario parameters in the parameter space. Contour lines of the loss function surface are available on the  $\vartheta_s$ - $r_s$  plane. Right:  $L^2$  norm of the obstacle temperatures computed online considering 10000 different scenario parameters in the parameter space.

An additional key point of the proposed strategy is the dimensionality reduction step that, differently from several control strategies available in the literature, allows us to easily deal with high-dimensional state variables and boundary or distributed controls. Furthermore, different non-intrusive reduced order models, including POD-NNs, DL-ROMs, and POD-DL-ROMs, can be taken into account in order to exploit the most effective reduction strategy for each OCP at hand, extending the current state-of-the-art on non-intrusive ROMs to control problems.

Several extensions of the proposed framework can be considered, even if they are left to future works. For instance, instead of dealing with high-dimensional states simulated through FEM solvers, it is in principle possible to exploit sensor data or images representing the underlying dynamics. A further, interesting enhancement of the proposed framework is to consider data-driven or physics-informed surrogate models approximating the forward problem to speed up the data generation process, which can be computationally intensive, although necessary only once during the offline phase. Moreover, uncertain parameters can be incorporated into the parameterized PDE to enable real-time robust control strategies. Feedback signals may also be considered if additional measurements are available online, as well as attention mechanisms may be integrated into the parameter-to-solution map in order to achieve better results when extrapolating in time.

## Acknowledgments

AM acknowledges the Project “Reduced Order Modeling and Deep Learning for the real-time approximation of PDEs (DREAM)” (Starting Grant No. FIS00003154), funded by the Italian Science Fund (FIS) - Ministero dell’Università e della Ricerca and the project FAIR (Future Artificial Intelligence Research), funded by the NextGenerationEU program within the PNRR-PE-AI scheme (M4C2, Investment 1.3, Line on Artificial Intelligence). MT and AM are members of the Gruppo Nazionale Calcolo Scientifico-Istituto Nazionale di Alta Matematica (GNCS-INdAM).

## References

- [1] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. Rognes, and G. Wells. The fenics project version 1.5. *Archive of Numerical Software*, 3(100):9–23, 2015.
- [2] D. Amsallem, M. Zahr, Y. Choi, and C. Farhat. Design optimization using hyper-reduced-order models. *Structural and Multidisciplinary Optimization*, 51(4):919–940, 2015.
- [3] E. A. Antonelo, E. Camponogara, L. O. Seman, J. P. Jordanou, E. R. de Souza, and J. F. Hübner. Physics-informed neural nets for control of dynamical systems. *Neurocomputing*, 579:127419, 2024.

- [4] J. Barry-Straume, A. Sarshar, A. A. Popov, and A. Sandu. Physics-informed neural networks for PDE-constrained optimization and control. arXiv:2205.03377, 2022.
- [5] P. Benner, E. Sachs, and S. Volkwein. Model order reduction for PDE constrained optimization. *International Series of Numerical Mathematics*, 165:303–326, 2014.
- [6] M. Bergmann and L. Cordier. Optimal control of the cylinder wake in the laminar regime by trust-region methods and pod reduced-order models. *J. Comput. Phys.*, 227(16):7813–7840, 2008.
- [7] K. Bieker, S. Peitz, S. L. Brunton, J. N. Kutz, and M. Dellnitz. Deep model predictive flow control with limited sensor data and online learning. *Theoretical and Computational Fluid Dynamics*, 34(4):577–591, 2020.
- [8] N. Botteghi and U. Fasel. Parametric PDE Control with Deep Reinforcement Learning and Differentiable L0-Sparse Polynomial Policies. arXiv:2403.15267, 2024.
- [9] S. Brivio, S. Fresca, N. R. Franco, and A. Manzoni. Error estimates for pod-dl-roms: a deep learning framework for reduced order modeling of nonlinear parametrized pdes enhanced by proper orthogonal decomposition. *Advances in Computational Mathematics*, 50(3):33, 2024.
- [10] S. Brunton and J. N. Kutz. *Data-driven science and engineering: machine learning, dynamical systems, and control*. Cambridge University Press, 2019.
- [11] S. L. Brunton, B. R. Noack, and P. Koumoutsakos. Machine Learning for Fluid Mechanics. *Annual Review of Fluid Mechanics*, 52:477–508, 2020.
- [12] 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 of the United States of America*, 113(15):3932–3937, 2016.
- [13] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Sparse Identification of Nonlinear Dynamics with Control (SINDYc). *IFAC-PapersOnLine*, 49(18):710–715, 2016.
- [14] P. Chen, M. R. Haberman, and O. Ghattas. Optimal design of acoustic metamaterial cloaks under uncertainty. *Journal of Computational Physics*, 431:110114, 2021.
- [15] Y. Chen, Y. Shi, and B. Zhang. Optimal Control Via Neural Networks: A Convex Approach. arXiv:1805.11835, 2018.
- [16] S. Cominelli, D. E. Quadrelli, C. Sinigaglia, and F. Braghin. Design of arbitrarily shaped acoustic cloaks through partial differential equation-constrained optimization satisfying sonic-metamaterial design requirements. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 478(2257), 2022.
- [17] L. Dede’. Optimal flow control for Navier–Stokes equations: Drag minimization. *International Journal for Numerical Methods in Fluids*, 55:347–366, 2007.
- [18] N. Demo, M. Strazzullo, and G. Rozza. An extended physics informed neural network for preliminary analysis of parametric optimal control problems. *Computers and Mathematics with Applications*, 143(May):383–396, 2023.
- [19] A. Draeger, S. Engell, and H. Ranke. Model predictive control using neural networks. *IEEE Control Systems Magazine*, 15(5):61–66, 1995.
- [20] N. R. Franco, A. Manzoni, and P. Zunino. A Deep Learning approach to Reduced Order Modeling of parameter dependent Partial Differential Equations. *Mathematics of Computation*, 92(340):483–524, 2023.
- [21] S. Fresca, L. Dede’, and A. Manzoni. A Comprehensive Deep Learning-Based Approach to Reduced Order Modeling of Nonlinear Time-Dependent Parametrized PDEs. *Journal of Scientific Computing*, 87(2):1–36, 2021.

- [22] S. Fresca and A. Manzoni. POD-DL-ROM: Enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition. *Computer Methods in Applied Mechanics and Engineering*, 388:114181, 2022.
- [23] P. Garnier, J. Viquerat, J. Rabault, A. Larcher, A. Kuhnle, and E. Hachem. A review on deep reinforcement learning for fluid mechanics. *Computers and Fluids*, 225, 2021.
- [24] C. Geuzaine and J.-F. Remacle. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009.
- [25] M. D. Gunzburger. *Perspectives in flow control and optimization*. SIAM, 2002.
- [26] M. D. Gunzburger, L. Hou, and T. P. Svobodny. Analysis and finite element approximation of optimal control problems for the stationary navier-stokes equations with distributed and neumann controls. *Mathematics of Computation*, 57(195):123–151, 1991.
- [27] J. Hesthaven and S. Ubbiali. Non-intrusive reduced order modeling of nonlinear problems using neural networks. *Journal of Computational Physics*, 363, 2018.
- [28] J. S. Hesthaven, G. Rozza, and B. Stamm. Certified Reduced Basis Methods for Parametrized Partial Differential Equations. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, pages 1–131, 2015.
- [29] M. K. Hickner, U. Fasel, A. G. Nair, B. W. Brunton, and S. L. Brunton. Data-Driven Unsteady Aeroelastic Modeling for Control. *AIAA Journal*, 61(2):780–792, 2023.
- [30] M. Hinze and K. Kunisch. Second order methods for optimal control of time-dependent fluid flow. *SIAM Journal on Control and Optimization*, 40(3):925–946, 2001.
- [31] R. Hwang, J. Y. Lee, J. Y. Shin, and H. J. Hwang. Solving PDE-Constrained Control Problems Using Operator Learning. *Proceedings of the 36th AAAI Conference on Artificial Intelligence, AAAI 2022*, 36(Phase 1):4504–4512, 2022.
- [32] T. Ishize, H. Omichi, and K. Fukagata. Flow control by a hybrid use of machine learning and control theory. arXiv:2311.08624, 2023.
- [33] K. Ito and S. S. Ravindran. Reduced basis method for optimal control of unsteady viscous flows. *Int. J. Comput. Fluid Dyn.*, 15(2):97–113, 2001.
- [34] E. Kaiser, J. N. Kutz, and S. L. Brunton. Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2219), 2018.
- [35] M. Kärcher and M. Grepl. A certified reduced basis method for parametrized elliptic optimal control problems. *ESAIM: Control Optim. Calc. Var.*, 20(2):416–441, 2014.
- [36] S. Klus, F. Nüske, S. Peitz, J.-H. Niemann, C. Clementi, and C. Schütte. Data-driven approximation of the Koopman generator: Model reduction, system identification, and control. *Physica D: Nonlinear Phenomena*, 406:132416, 2020.
- [37] M. Korda and I. Mezić. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *Automatica*, 93:149–160, 2018.
- [38] K. Kunisch and S. Volkwein. Control of the Burgers Equation by a Reduced-Order Approach Using Proper Orthogonal Decomposition. *Journal of Optimization Theory and Applications*, 102(2):345–371, 1999.
- [39] K. Kunisch and S. Volkwein. Proper orthogonal decomposition for optimality systems. *ESAIM: Modélisation mathématique et analyse numérique*, 42(1):1–23, 2008.

- [40] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021.
- [41] D. Luo, T. O’Leary-Roseberry, P. Chen, and O. Ghattas. Efficient PDE-Constrained optimization under high-dimensional uncertainty using derivative-informed neural operators. arXiv:2305.20053, 2023.
- [42] P. Ma, Y. Tian, Z. Pan, B. Ren, and D. Manocha. Fluid directed rigid body control using deep reinforcement learning. *ACM Transactions on Graphics*, 37(4), 2018.
- [43] A. Manzoni and S. Pagani. A certified RB method for PDE-constrained parametric optimization problems. *Communications in Applied and Industrial Mathematics*, 10(1):123–152, 2019.
- [44] A. Manzoni, A. Quarteroni, and S. Salsa. *Optimal Control of Partial Differential Equations*. Springer Cham, 2021.
- [45] M. Mirzakhloo, S. Esmailzadeh, and M.-r. Alam. Active cloaking in Stokes flows via reinforcement learning. *Journal of Fluid Mechanics*, 903(A34), 2020.
- [46] S. K. Mitusch, S. W. Funke, and J. S. Dokken. dolfin-adjoint 2018.1: automated adjoints for fenics and firedrake. *Journal of Open Source Software*, 4(38):1292, 2019.
- [47] S. Mowlavi and S. Nabi. Optimal control of PDEs using physics-informed neural networks. *Journal of Computational Physics*, 473:111731, 2023.
- [48] F. Negri, A. Manzoni, and G. Rozza. Reduced basis approximation of parametrized optimal flow control problems for the stokes equations. *Computers & Mathematics with Applications*, 69(4):319–336, 2015.
- [49] F. Negri, G. Rozza, A. Manzoni, and A. Quarteroni. Reduced basis method for parametrized elliptic optimal control problems. *SIAM Journal on Scientific Computing*, 35(5):A2316–A2340, 2013.
- [50] S. Peitz and K. Bieker. On the universal transformation of data-driven models to control systems. *Automatica*, 149:110840, 2023.
- [51] S. Peitz and S. Klus. Koopman operator-based model reduction for switched-system control of PDEs. *Automatica*, 106:184–191, 2019.
- [52] J. L. Proctor, S. L. Brunton, and J. N. Kutz. Dynamic Mode Decomposition with Control. *SIAM Journal on Applied Dynamical Systems*, 15(1):142–161, 2016.
- [53] A. Quarteroni. *Numerical Models for Differential Problems*. Springer Cham, 2017.
- [54] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations: An introduction*. Springer Cham, 2015.
- [55] J. Rabault, M. Kuchta, A. Jensen, and U. Réglade. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *Journal of Fluid Mechanics*, 865:281–302, 2019.
- [56] J. Rabault, F. Ren, W. Zhang, H. Tang, and H. Xu. Deep reinforcement learning in fluid mechanics: A promising method for both active flow control and shape optimization. *Journal of Hydrodynamics*, 32(2):234–246, 2020.
- [57] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.
- [58] S. S. Ravindran. A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. *Int. J. Numer. Meth. Fluids*, 34(5):425–448, 2000.

- [59] F. Ren, H.-b. Hu, and H. Tang. Active flow control using machine learning: A brief review. *Journal of Hydrodynamics*, 32(2):247–253, 2020.
- [60] F. Ren, C. Wang, and H. Tang. Bluff body uses deep-reinforcement-learning trained active flow control to achieve hydrodynamic stealth. *Physics of Fluids*, 33(9):093602, 2021.
- [61] O. San, A. Rasheed, and T. Kvamsdal. Hybrid analysis and modeling, eclecticism, and multifidelity computing toward digital twin revolution. *GAMM Mitteilungen*, 44(2), 2021.
- [62] C. Sánchez-Sánchez and D. Izzo. Real-Time Optimal Control via Deep Neural Networks: Study on Landing Problems. *Journal of Guidance, Control, and Dynamics*, 41(5):1122–1135, 2018.
- [63] R. Saporiti, C. Sinigaglia, A. Manzoni, and F. Braghin. An optimal control strategy to design passive thermal cloaks of arbitrary shape. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 480(2294):20230646, 2024.
- [64] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, 2010.
- [65] T. Shah, L. Zhuo, P. Lai, A. De La Rosa-Moreno, F. Amirkulova, and P. Gerstoft. Reinforcement learning applied to metamaterial design. *The Journal of the Acoustical Society of America*, 150(1):321–338, 2021.
- [66] C. Sinigaglia, F. Braghin, and S. Berman. Optimal Control of Velocity and Nonlocal Interactions in the Mean-Field Kuramoto Model. *Proceedings of the American Control Conference, 2022-June*:290–295, 2022.
- [67] C. Sinigaglia, A. Manzoni, and F. Braghin. Density Control of Large-Scale Particles Swarm Through PDE-Constrained Optimization. *IEEE Transactions on Robotics*, 38(6):3530–3549, 2022.
- [68] C. Sinigaglia, A. Manzoni, F. Braghin, and S. Berman. Indirect Optimal Control of Advection-Diffusion Fields through Distributed Robotic Swarms. *IFAC-PapersOnLine*, 55(30):299–304, 2022.
- [69] C. Sinigaglia, A. Manzoni, F. Braghin, and S. Berman. Robust optimal density control of robotic swarms. arXiv:2205.12592, 2022.
- [70] C. Sinigaglia, D. E. Quadrelli, A. Manzoni, and F. Braghin. Fast active thermal cloaking through PDE-constrained optimization and reduced-order modelling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 478(2258), 2022.
- [71] R. S. Sutton and A. G. Barto. *Reinforcement Learning: an introduction*. The MIT Press, 2018.
- [72] S. Verma, G. Novati, and P. Koumoutsakos. Efficient collective swimming by harnessing vortices through deep reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, 115(23):5849–5854, 2018.
- [73] C. Vignon, J. Rabault, and R. Vinuesa. Recent advances in applying deep reinforcement learning for flow control: Perspectives and future directions. *Physics of Fluids*, 35(3), 2023.
- [74] R. Vinuesa, S. L. Brunton, and B. J. McKeon. The transformative potential of machine learning for experiments in fluid mechanics. *Nature Reviews Physics*, 5(9):536–545, 2023.
- [75] S. Wang, M. A. Bhourri, and P. Perdikaris. Fast PDE-constrained optimization via self-supervised operator learning. arXiv:2110.13297, 2021.
- [76] S. Wang, H. Wang, and P. Perdikaris. Learning the solution operator of parametric partial differential equations with physics-informed DeepONets. *Science Advances*, 7(40):1–9, 2021.
- [77] P. Yin, G. Xiao, K. Tang, and C. Yang. Aonn: An adjoint-oriented neural network method for all-at-once solutions of parametric optimal control problems. *SIAM Journal on Scientific Computing*, 46(1):C127–C153, 2024.

- [78] N. Zolman, U. Fasel, J. N. Kutz, and S. L. Brunton. SINDy-RL: Interpretable and Efficient Model-Based Reinforcement Learning. arXiv:2403.09110, 2024.

## MOX Technical Reports, last issues

Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 41/2025** Torzoni, M.; Maisto, D.; Manzoni, A.; Donnarumma, F.; Pezzulo, G.; Corigliano, A.  
*Active digital twins via active inference*
- 42/2025** Franco, N. R.; Manzoni, A.; Zunino, P.; Hesthaven, J. S.  
*Deep orthogonal decomposition: a continuously adaptive neural network approach to model order reduction of parametrized partial differential equations*
- 40/2025** Tentori, C.A.; Gregorio, C.; ...; Ieva, F.; Della Porta, M.G.  
*Clinical and Genomic-Based Decision Support System to Define the Optimal Timing of Allogeneic Hematopoietic Stem-Cell Transplantation in Patients With Myelodysplastic Syndromes*
- 37/2025** Spreafico, M.; Ieva, F.; Fiocco, M.  
*Causal effect of chemotherapy received dose intensity on survival outcome: a retrospective study in osteosarcoma*
- Gimenez Zapiola, A.; Boselli, A.; Menafoglio, A.; Vantini, S.  
*Hyper-spectral Unmixing algorithms for remote compositional surface mapping: a review of the state of the art*
- 35/2025** Perotto, S.; Ferro, N.; Speroni, G.; Temellini, E.  
*Anisotropic recovery-based error estimators and mesh adaptation for real-life engineering innovation*
- 34/2025** Bucelli, M.; Dede', L.  
*Coupling models of resistive valves to muscle mechanics in cardiac fluid-structure interaction simulations*
- 32/2025** De Sanctis, M.F.; Di Battista, I.; Arnone, E.; Castiglione, C.; Palummo, A.; Bernardi, M.; Ieva, F.; Sangalli, L.M.  
*Exploring nitrogen dioxide spatial concentration via physics-informed multiple quantile regression*
- 33/2025** Di Battista, I.; De Sanctis, M.F.; Arnone, E.; Castiglione, C.; Palummo, A.; Sangalli, L.M.  
*A semiparametric space-time quantile regression model*
- 31/2025** Botteghi, N.; Fresca, S.; Guo, M.; Manzoni, A.  
*HypeRL: Parameter-Informed Reinforcement Learning for Parametric PDEs*