Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

# POD-Identification reduced order model of linear transport equations for control purposes

S. Kelbij Star*[1,2] | Francesco Belloni[1] | Gert Van den Eynde[1] | Joris Degroote[2]

[1]Institute for Advanced Nuclear Systems, SCK·CEN, Mol, Belgium
[2]Department of Flow, Heat and Combustion Mechanics, Ghent University, Ghent, Belgium

**Correspondence**
*Kelbij Star. Email: kelbij.star@sckcen.be

**Present Address**
Institute for Advanced Nuclear Systems, SCK·CEN, Boeretang 200, 2400 Mol, Belgium

**Summary**

Intrusive reduced order modeling techniques require access to the solver's discretization and solution algorithm, which are not available for most Computational Fluid Dynamics codes. Therefore, a non-intrusive reduction method that identifies the system matrix of linear fluid dynamical problems with a least-squares technique is presented. The methodology is applied to the linear scalar transport convection - diffusion equation for a 2D square cavity problem with a heated lid. The (time-dependent) boundary conditions are enforced in the obtained reduced order model (ROM) with a penalty method. The results are compared and the accuracy of the reduced order models is assessed against the full order solutions and it is shown that the reduced order model can be used for sensitivity analysis by controlling the non-homogeneous Dirichlet boundary conditions.

**KEYWORDS:**
Reduced order modeling, POD: Proper Orthogonal Decomposition, Galerkin, Advection-diffusion, Finite volume, Model reduction

## 1 | INTRODUCTION

Computational Fluid Dynamics (CFD) simulations are widely used in industry to solve fluid problems. However, running transient simulations using a full CFD approach is completely unfeasible for many engineering purposes due to the excessive amount of computational power needed, especially when a large number of parameters are to be tested for control purposes, sensitivity analyses or uncertainty quantification studies. Model reduction techniques have therefore been developed to approximate the (parametrized) Partial Differential Equations (PDEs) describing the fluid problem, which reduces the CPU time and computer memory usage [1,2,3,4].

Reduced order modeling (ROM) techniques can be applied to the standard discretization techniques as Finite Difference (FD), Finite Volume (FV) and spectral methods, but have been mostly developed for the Finite Element (FE) method. In industry, however, the FV method is often used, by commercial software and open-source codes, for the spatial discretization of the governing equations that describe the physical fluid model [5,6], as the method is robust [7] and preserves locally the conservation laws [8,9]. A set of reduced basis functions, containing the essential dynamics of the full order system, is often created with the Proper Orthogonal Decomposition (POD) method [10]. POD is commonly applied in fluid dynamics literature for this purpose, as it can also be applied to nonlinear models [3,4]. These POD basis functions are obtained through solving an eigenvalue problem on snapshots which are generated by sampling the full order model (FOM) at several moments in time. For unsteady problems, POD is typically combined with the Galerkin projection where the full order system is projected onto the low-dimensional subspace of POD modes and the difference with the snapshots is minimized [3,10,11] to obtain a system of time-dependent coefficients,

the reduced order model. However, the main issue of this reduced basis method is that knowledge of the solver's discretization and solution algorithm is required in order to perform the Galerkin projection and could therefore not be used for most (commercial) software. Instead, non-intrusive ROMs (NIROM) [15], are using for instance a sparse grid collocation approach [12,14] or interpolation [12,13,14,16,17] to calculate the POD coefficients. On the other hand, data-driven techniques, such as System Identification (SI), are using the input/output data of a dynamical system to identify a low-dimensional system that approximately describes the dynamics of a high-dimensional system [18] with a set of low-order ordinary differential equations (ODEs). Examples are the Dynamic Mode Decomposition (DMD), first introduced by Schmid [19] as a method for extracting coherent dynamic flow structures from a set of snapshots and Krylov-subspace projection-based ROM methods as Vervecken et al. [20] have shown for the convection - diffusion equation [21]. A disadvantage of SI methods is that the obtained reduced system does not have a physical meaning and consistency issues can occur for parameterized problems [22]. Therefore, a POD-based identification (POD-ID) method is proposed here, for which reduced system matrices of the same form as in the POD-Galerkin method are identified using a least-squares technique. A set of ODEs, still describing the physical model, is then obtained. The resulting ROM can be used for controlling the (time-dependent) non-homogeneous Dirichlet boundary conditions (BCs) instead of having to perform a high fidelity simulation for every BC of interest. The BCs are enforced in the ROM with a penalty factor [23]. The paper is organized as follows: In Section 2 the scalar transport equation is introduced for which the full order simulation is performed. The methodology of the POD-based identification method is addressed for parametrized boundary conditions together with BC enforcement method in Section 3. In Section 4 the ROM technique is tested for a numerical experiment and the results are provided and discussed in Section 5 and 6, respectively. Finally conclusions are drawn in Section 7 and an outlook for further improvements is provided.

## 2 | THE CONVECTION - DIFFUSION SCALAR TRANSPORT EQUATION

In this work the unsteady convection-diffusion scalar transport equation in incompressible form is considered, given by

$$\begin{cases} \frac{\partial T}{\partial t} + \nabla \cdot (\langle \boldsymbol{u} \rangle T) - \nabla^2(\mathcal{D}T) = 0 & \text{in } \Omega, \\ T(\boldsymbol{x}, t) = f(t) & \text{on } \Gamma_{lid}, \\ T(\boldsymbol{x}, t) = 0 & \text{on } \Gamma_0, \\ T(\boldsymbol{x}, 0) = T_0 & \text{in } \Omega, \end{cases} \tag{1}$$

where $\langle \boldsymbol{u} \rangle$ is a steady velocity field, $T$ is the transported scalar, $T_0$ is the initial scalar field and $\mathcal{D}$ is the diffusion coefficient divided by the fluid density and the heat capacity, which are both constant. The boundary of the domain, $\Omega$, is divided in two parts: $\Gamma = \Gamma_0 \cup \Gamma_{lid}$, where $\boldsymbol{u} = g(\boldsymbol{x})$ on $\Gamma_{lid}$ and $\boldsymbol{u} = (0,0)$ on $\Gamma_0$. It is important that the given steady background velocity field, $\langle \boldsymbol{u} \rangle$, satisfies the continuity equation $\nabla \cdot \boldsymbol{u} = 0$ for incompressible flows. This flow field can be obtained from a standard CFD model and hence its calculation is not described further. Discretizing the transport equation in space and rearranging in matrix form leads to the following system of equations

$$\begin{cases} \dot{\boldsymbol{T}} + \boldsymbol{C}\boldsymbol{T} - \mathcal{D}\boldsymbol{B}\boldsymbol{T} = 0, \\ \boldsymbol{T}(0) = \boldsymbol{T}_0, \end{cases} \tag{2}$$

where the dot indicates the time derivative, $\boldsymbol{C}$ and $\boldsymbol{B}$ are the convective and diffusive matrices, respectively. $\boldsymbol{T}_0$ is the initial condition.

## 3 | THE REDUCED ORDER MODEL

In this section the methodology of the proposed POD-based IDentification method, as a non-intrusive reduced-order method (NIROM) for linear CFD problems, is described.

### 3.1 | Generation of the POD-basis

The main assumption of the POD method is that the system's dynamics are governed by a reduced number of dominant modes and that there exists an approximation of the transported scalar, $T_r(\boldsymbol{x}, t)$, so that the full order solution, $\boldsymbol{T}(\boldsymbol{x}, t)$, can be expressed

as a linear combination of orthogonal spatial modes, $\phi_i(\boldsymbol{x})$, multiplied by time-dependent coefficients, $a_i(t)$, as follows

$$T(\boldsymbol{x},t) \approx T_r(\boldsymbol{x},t) = \sum_{i=1}^{N_r} \phi_i(\boldsymbol{x}) a_i(t) \tag{3}$$

where $N_r$ is the dimension of the reduced basis space[3,10,11]. The spatial modes $\phi_i(\boldsymbol{x})$ are determined using a snapshot technique where a snapshot matrix $\boldsymbol{Y}$ is generated, in this case by numerical simulations, containing a set of solutions of $T$, internal field plus boundary points, at some selected times $t^n$ for n = 0,..., $N_t$.

$$\boldsymbol{Y} = [T(\boldsymbol{x},t^0), T(\boldsymbol{x},t^1), \dots, T(\boldsymbol{x},t^{N_t})] \in \mathbb{R}^{N_x \times (N_t+1)} \tag{4}$$

where $N_x$ is the spatial dimension and $T(\boldsymbol{x},t^0)$ is the initial condition. The snapshots do not necessarily have to be collected at every time step for which the full order solution is calculated. As the POD modes are orthogonal to each other, $\left\langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j \right\rangle_{L_2(\Omega)} = \delta_{ij}$, the POD basis, $E^{POD}$, is optimal when the difference between the snapshots and the projection of the snapshots on the basis functions is minimal for a certain norm. The $L_2$-norm is preferred for discrete numerical schemes[24] with $\langle \cdot, \cdot \rangle_{L_2(\Omega)}$ the $L^2$ inner product of the functions over the domain $\Omega$. The minimization problem is given by

$$E_{N_r}^{POD} = \arg\min \frac{1}{N_t+1} \sum_{n=0}^{N_t} \left\| T(\boldsymbol{x},t^n) - \sum_{i=1}^{N_r} \langle T(\boldsymbol{x},t^n), \phi_i(\boldsymbol{x}) \rangle_{L_2} \phi_i(\boldsymbol{x}) \right\|_{L_2}^2 . \tag{5}$$

One way to compute the POD modes is by applying the Singular Value Decomposition (SVD) to the snapshot matrix, $\boldsymbol{Y} = U\Sigma V^T$. However, the SVD approach is computationally more expensive than solving the eigenvalue problem, especially when the dimension of the grid, used to discretize the domain, is increased. It is then recommended to solve an eigenvalue problem, using the correlation matrix $\boldsymbol{C_{corr}} \in \mathbb{R}^{(N_t+1)\times(N_t+1)}$ of the snapshots, to determine the basis functions. The eigenvalue problem is given by

$$\boldsymbol{C_{corr}} \boldsymbol{Q} = \boldsymbol{Q} \lambda \tag{6}$$

where $\boldsymbol{C_{corr}}_{i,j} = \left\langle T(\boldsymbol{x},t^i), T(\boldsymbol{x},t^j) \right\rangle_{L_2(\Omega)}$ is the correlation matrix, $Q$ is a square matrix of eigenvectors and $\lambda$ is a vector containing the eigenvalues. The POD modes, $\phi_i$, can then be constructed as follows

$$\phi_i(\boldsymbol{x}) = \frac{1}{(N_t+1)\sqrt{\lambda_i}} \sum_{n=0}^{N_t} T(\boldsymbol{x},t^n) Q_{i,n} \qquad \text{for } i = 1, ..., N_r. \tag{7}$$

For a more detailed explanation of this method the reader is referred to[25,26].

## 3.2 | POD-based Identification

In case of the classical POD-Galerkin method, the scalar function, $T$, is replaced by the approximation, $T_r$, in equation (1) and by applying the Galerkin projection onto the reduced basis, the following ROM is obtained

$$\dot{\boldsymbol{a}} + \boldsymbol{C_r} \boldsymbol{a} - \mathcal{D} \boldsymbol{B_r} \boldsymbol{a} = 0 \tag{8}$$

where

$$\boldsymbol{B_r}_{i,j} = \left\langle \nabla \phi_i, \nabla \phi_j \right\rangle_{L_2(\Omega)} \tag{9}$$

$$\boldsymbol{C_r}_{i,j} = \left\langle \phi_i, \nabla \cdot (\langle u \rangle \phi_j) \right\rangle_{L_2(\Omega)} \tag{10}$$

The main issue of this method is, however, that knowledge of the solver's discretization and solution algorithm is required in order to perform the Galerkin projection. Furthermore the full order matrices, $\boldsymbol{B}$ and $\boldsymbol{C}$, in equation (2) are not accessible within most CFD codes, due to restricted access to the source code in commercial software or due to the used solution methodology for open-source codes[20]. Therefore, it is also not an option to apply a Galerkin projection on the matrices of the full order systems in the following way

$$\boldsymbol{B_r} = \Phi^T \boldsymbol{B} \Phi \tag{11}$$

$$\boldsymbol{C_r} = \Phi^T \boldsymbol{C} \Phi \tag{12}$$

where $\Phi = [\phi_1, \phi_2, ..., \phi_{N_r}]$. Thence, the POD-based Identification method aims at identifying these reduced matrices using a least-squares technique such as normal equations, QR-decomposition or SVD by minimizing the residual, $\boldsymbol{R}$,

$$\boldsymbol{R} = \dot{\boldsymbol{a}} + \boldsymbol{C_r a} - \mathcal{D}\boldsymbol{B_r a} \tag{13}$$

in the following way

$$\left[\hat{\boldsymbol{B}}_r\ \hat{\boldsymbol{C}}_r\right] = \min_{\boldsymbol{B}_r,\ \boldsymbol{C}_r} \|R\|\,. \tag{14}$$

In equation 13, the time-dependent coefficients are constructed via a projection of modes on the full order solution

$$\boldsymbol{a}(t) = \Phi^T \boldsymbol{T}(\boldsymbol{x}, t) \tag{15}$$

In addition, when the dynamical system is linear, no sources or sinks are present and the variables, for instance the diffusion coefficient $\mathcal{D}$, are not a function of a parameter $\mu \in$ parameter space $\mathcal{P}$, the ROM can be simplified by the assumption that the time-dependent coefficients are related via the linear mapping

$$\boldsymbol{A_r a}^{n+1} = \boldsymbol{a}^n \qquad \text{for } \ n = 0, \ldots, N_t - 1 \tag{16}$$

where $\boldsymbol{A_r}$ is an unknown matrix to be identified using a least-squares approach. In order to do that, two matrices, $\boldsymbol{X_0}$ and $\boldsymbol{X_1}$, are constructed that contain the known time-dependent coefficients at certain times in the following way

$$\boldsymbol{X_0} = [\boldsymbol{a}^0, \boldsymbol{a}^1, ..., \boldsymbol{a}^{N_t-1}] \tag{17}$$

$$\boldsymbol{X_1} = [\boldsymbol{a}^1, \boldsymbol{a}^2, ..., \boldsymbol{a}^{N_t}] \tag{18}$$

to satisfy equation 16 as good as possible for each time step in which the snapshot is collected, by minimizing the difference between $\boldsymbol{A_r X_1}$ and $\boldsymbol{X_0}$ Therefore, the reduced matrix, $\boldsymbol{A_r}$, is computed by minimizing the norm

$$\boldsymbol{A}_r = \arg\min_{\hat{\boldsymbol{A}}_r} \left\| \hat{\boldsymbol{A}}_r \boldsymbol{X_1} - \boldsymbol{X_0} \right\| \tag{19}$$

using a least-squares technique. A similar approach is applied for the Dynamic Mode Decomposition, where the snapshots are assumed to be related via a linear mapping and the discrete-time linear system is then fitted on the set of snapshots[27]. The POD-ID method differs in the sense that the mapping is done at the reduced level instead of the high-dimensional level at which the snapshots are obtained. The maximum number of modes to be considered, $N_r$, is of the order $\sqrt{N_t}$ as an overdetermined system is required in order to identify a reduced matrix $\boldsymbol{A_r}$ of size $N_r$ x $N_r$.

Finally, three types of scalar fields are considered: the full order field $T(\boldsymbol{x}, t)$, the projected field $T_r(\boldsymbol{x}, t)$ which is obtained by the projection of the FOM snapshots onto the POD basis and the prediction field, $T^{ROM}(\boldsymbol{x}, t)$. For every time instance, $t^n$ for n $= 0, \ldots, N_t$, the basis projection error, $\|e\|_{\ell_2}$, is given by

$$\|e\|_{\ell_2} = \sqrt{\frac{\langle (T - T_r), (T - T_r) \rangle_{\ell_2}}{\langle T, T \rangle_{\ell_2}}} \tag{20}$$

and the prediction error, $\|\hat{e}\|_{\ell_2}$, by

$$\|\hat{e}\|_{\ell_2} = \sqrt{\frac{\langle (T - T^{ROM}), (T - T^{ROM}) \rangle_{\ell_2}}{\langle T, T \rangle_{\ell_2}}}. \tag{21}$$

For both the $\ell_2$-error norm is considered, where $\langle \cdot, \cdot \rangle_{\ell_2(\Omega)}$ is the $\ell^2$ inner product of the fields over the domain $\Omega$.

## 3.3 | Boundary conditions

### 3.3.1 | (Non-)homogeneous boundary conditions

In order to control the boundary conditions at the reduced order level, the POD-ID method has to be extended. First of all, the boundary conditions need to be enforced in the ROM. However, an issue arises in the case of non-homogeneous Dirichlet BCs as any linear combination of snapshots used for the creation of the POD basis will, in general, not satisfy the same BCs and the same applies to the ROM. There is no problem in the case of homogeneous BCs as the linear combinations of snapshots with

homogeneous boundary conditions will naturally satisfy the same BCs. To solve these issues, the boundary points are added first to each snapshot before generating the POD basis functions, $\phi(\boldsymbol{x})$. As these basis functions are solely depending on the spatial coordinates and the arrangement can be chosen freely as long as it is consistent over the time steps, it is chosen to add the non-homogeneous boundary conditions to the end of each vector containing the snapshot data for the corresponding boundary points.

### 3.3.2 | Parametrized boundary conditions

In order to impose control on the boundary conditions at the reduced order level, the POD-ID method has to be extended. First of all, the boundary conditions need to be enforced in the ROM. However, an issue arises in the case of non-homogeneous Dirichlet BCs as any linear combination of snapshots used for the creation of the POD basis will, in general, not satisfy the same BCs and the same applies to the ROM. There is no problem in the case of homogeneous BCs as the linear combinations of snapshots with homogeneous boundary conditions will naturally satisfy the same BCs. Another way to deal with non-homogeneous BCs is to add an additional constraint to the transport equation in order to weakly enforce the boundary condition for the ROM with a penalty factor[3,23,25,28,29]. No modification of the snapshots is needed other than adding the boundary points in case these are not present, otherwise the ROM could become unstable. The constraint is added to the transport equation in the following way

$$\frac{\partial T}{\partial t} + \nabla \cdot (\langle \boldsymbol{u} \rangle T) - \nabla^2(\mathcal{D}T) + \tau\Gamma(T - T_{bc}(t)) = 0 \tag{22}$$

where $T_{bc}(t)$ is the (time-dependent) Dirichlet boundary condition, $\tau$ the penalty factor and $\Gamma$ is a null function except on the boundary where the condition is imposed[25]. In order to have an asymptotically stable solution, the penalty factor $\tau$ should be larger than 0. In case $\tau \to \infty$ a strong imposition would be approached and the ROM becomes ill-conditioned. This penalty factor can be found by numerical experimentation[23,25,30].

At reduced order level, after a Galerkin projection, this translates to

$$\dot{\boldsymbol{a}} + \boldsymbol{C_r a} - \mathcal{D}\boldsymbol{B_r a} + \tau(\boldsymbol{E a} - \boldsymbol{\delta}(t)) = 0 \tag{23}$$

where $\boldsymbol{\delta}$ is the projection of the boundary values on the modes at the boundary and $\boldsymbol{E}$ the modes projected on the reduced basis at the boundary domain, $\partial\Omega$, given respectively, by

$$\delta_i(t) = \langle \phi_i(\boldsymbol{x}), T_{BC}(\boldsymbol{x}, t) \rangle_{\ell_2, \partial\Omega} \tag{24}$$

$$E_{ij} = \left\langle \phi_i(\boldsymbol{x}), \phi_j(\boldsymbol{x}) \right\rangle_{\ell_2, \partial\Omega} \tag{25}$$

For the POD-ID method, without the parametrization of the diffusion coefficient, the reduced system is given by

$$\left(\boldsymbol{A_r} + \tau \boldsymbol{E}\right) \boldsymbol{a}^{n+1} = \boldsymbol{a}^n + \tau \boldsymbol{\delta}(t^n) \tag{26}$$

that can be solved for $\boldsymbol{a}^{n+1}$, depending on the boundary, $T_{bc}(t)$, applied. The initial condition for the ROM is obtained by projecting the full order initial condition for the parametrized BC, $T_{bc}(t^0)$, onto the POD basis as follows

$$\boldsymbol{a}^0 = \Phi^T \boldsymbol{T}(t^0) \tag{27}$$

The overall algorithm for the POD-ID method including the penalty method is given below.

**Algorithm: POD-ID method including penalty method**

**Create ROM with POD-ID method:**

(1) Generate snapshots over a time period $[0, t^{N_t}]$ by solving the linear full order problem of Eq. 1;
(2) Retrieve the snapshots matrix $\boldsymbol{Y}$ from the solutions obtained as in Eq 4;
(3) Perform POD on $\boldsymbol{Y}$ to obtain the POD modes $\Phi$ using Eq. 7;
(4) Project the snapshots on the modes to obtain the corresponding coefficients $\boldsymbol{a}(t)$ using Eq. 15;
(5) Retrieve the matrices $\boldsymbol{X_0}$ and $\boldsymbol{X_1}$ from the coefficients with Eq. 17 and 18;
(6) Identify the reduced matrix $\boldsymbol{A_r}$ by minimizing the norm $\left\| \hat{\boldsymbol{A}}_r \boldsymbol{X_0} - \boldsymbol{X_1} \right\|$ with a least-squares technique as in Eq. 19;
(7) Project the initial field for the parametrized BC onto the POD basis to get the initial condition $\boldsymbol{a}^0$ for the ROM using Eq. 27;

**Impose BCs with penalty method:**

(8) Project the values for the parametrized non-homogeneous Dirichlet boundary on the modes to determine $\boldsymbol{\delta(t)}$ using Eq. 24;
(9) Project the modes on the reduced basis at the same boundary domain of previous step to determine $\boldsymbol{E}$ using Eq. 25;
(10) Set a value for the penalty factor $\tau$;

**Solve reduced order model:**

(11) Solve the reduced order problem of Eq. 26 for the time period $[0, t^{N_t}]$;
(12) **if** The boundary is not enforced in ROM solution **then**
go back to step (10).
**end if**
(13) Reconstruct the full order fields from the obtained coefficients using Eq. 3;
(14) Calculate the prediction error using Eq. 21.

## 4 | NUMERICAL EXPERIMENTS

The POD-ID method is tested for the classical numerical 2D lid-driven cavity benchmark problem[31,32] that consists of a 2D square cavity of length $L = 1$ m on which a (512 x 512) uniform mesh is constructed. The geometry is depicted in Figure 1. The boundary of the domain is divided in two parts: $\Gamma = \Gamma_0 \cup \Gamma_{lid}$, where $\boldsymbol{u} = (1,0)$ on $\Gamma_{lid}$ and $\boldsymbol{u} = (0,0)$ on $\Gamma_0$. A homogeneous Neumann boundary condition for the pressure is applied everywhere on $\Gamma$. The steady background velocity field is precomputed first in the offline phase for laminar flow with Re = 1000. The viscosity, $\nu$, is taken at $1 \cdot 10^{-4}$ m$^2$/s. The initial conditions are set to $\boldsymbol{u}(\boldsymbol{x},0) = 0$ m/s and $p(\boldsymbol{x},0) = 0$ Pa for the velocity and pressure, respectively. The reference value for the pressure is set to 0 Pa at coordinate (0,0). The calculation of the flow field is performed in the OpenFOAM environment[21], a finite volume open-source code[8,33,34], with the icoFoam solver for $t = 50$ s with time steps of 0.01s. The final time, $t = 50$ s, is taken as the steady velocity field. In fact, any steady flow field could be chosen.

The temperature is calculated everywhere in the heated lid-driven cavity[35] by solving the scalar transport equation (1) with $\mathcal{D}$ equal to $1 \cdot 10^{-5}$ m$^2$/s for the thermal diffusion constant and the flow field previously obtained. Furthermore, a homogeneous Dirichlet BC is applied on $\Gamma_0$ and a non-homogeneous BC with $T_{lid}^{FOM} = 1$ on $\Gamma_{lid}$. The initial condition is set to $T(\boldsymbol{x},0) = 0$. All temperatures are made non-dimensional by a reference temperature $T_{ref}$. The simulation is performed with the scalarTransportFoam solver of OpenFOAM with an implicit scheme for the time discretization. A constant time step of $\Delta t = 1 \cdot 10^{-3}$ s has been applied. Snapshots of the temperature are collected every 0.1 s, resulting in a total of 1000 snapshots. The POD modes and ROM are constructed according to Section 3.1 and 3.2, respectively. The time-independent BC for the ROM, $T_{lid}^{ROM}$, could be chosen freely, although it should be of the same order as $T_{lid}^{FOM}$ for which the POD has been performed to avoid discrepancies, as the statement "If it is not in the snapshots, it is not in the ROM" by Quarteroni et al.[3] still applies. The BCs are enforced with the penalty method according to Section 3.3.2 and the ROM is tested for $T_{lid}^{ROM} = 0.5$, 2 and 10. As the system is linear, parameterizing the time-independent Dirichlet BC on $\Gamma_{lid}$ is straightforward as the FOM solutions can be scaled according to a BC of interest and no ROM has to be constructed. Nevertheless, the ROMs are constructed in order to demonstrate the capability of the POD-ID method.

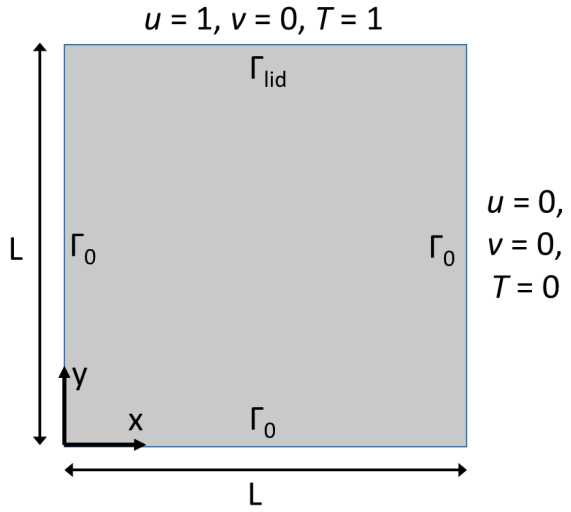


**FIGURE 1** A sketch of the geometry of the heated lid-driven cavity problem including boundary conditions.

Finally, the ROM is tested for a time-dependent boundary condition given by

$$T(x, 1, t) = Asin(2\pi f t) + T_0 \tag{28}$$

where $f$ is the frequency of the wave, $A$ the amplitude and $T_0$ the off-set. Full order solutions, with this time-dependent BC, are calculated for $f = 0.01$ s$^{-1}$, $A = 1$ and $T_0 = 1$, which will be referred to as the base case (set 1 in Table 1). The ROM is tested for five sets of parameters that are defining the Dirichlet BC, summarized as set 2-5 in Table 1. To evaluate the accuracy of the POD-ID method for all ROMs, the $\ell_2$-error norm is considered according to equation (21).

**TABLE 1** Parameter sets for time-dependent Dirichlet BCs defined by equation (28)

| | $f$ | $A$ | $T_0$ |
|---|---|---|---|
| **Set 1** | 0.01 | 1 | 1 |
| **Set 2** | 0.01 | 2 | 1 |
| **Set 3** | 0.01 | 1 | 3 |
| **Set 4** | 0.01 | 2 | 3 |
| **Set 5** | 0.011 | 1 | 1 |
| **Set 6** | 0.02 | 1 | 1 |

In the offline phase, the snapshots are created with OpenFOAM, while in the second part of the offline phase, creating the POD modes and constructing the reduced system of equations (8), is performed with MATLAB [36]. Also the online phase, solving the reduced order systems for different Dirichlet boundary conditions is done with MATLAB. The whole offline-online procedure is carried out with a single processor on an Intel core i5. The ROM's online computational time depends on the number of modes and is no longer dependent on the number of degrees of freedom of the FOM.

## 5 | RESULTS

In this section the accuracy of the ROM is tested for both time-independent and time-dependent Dirichlet BC. Before these tests, the background velocity is precomputed. The flow field is shown together with the corresponding pressure field in Figure 2. Then, the full order simulation for the time-independent Dirichlet BC $T_{lid}^{FOM} = 1$ is performed until $t = 100$ s. The evolution of the temperature field in time is shown in Figure 3 for $t = 1$, 10, 50 and 100 s.

During the offline phase the reduced basis for the temperature was calculated following the procedure described in Section 3. The decay of the normalized POD eigenvalues is plotted in Figure 4 in order to determine the number of basis functions needed to create the reduced subspace. The figure shows that 18 basis functions are required to have a truncation error less than $10^{-6}$. The cumulative eigenvalues (CV) can be found in Table 2 and 36 modes are sufficient to retain more than 99.9% of the energy for temperature. However, as pointed out in Section 3.2, the maximum number of modes to be considered is of the order $\sqrt{N_t}$, which is 31 modes for a snapshot matrix containing 1000 snapshots. 31 modes correspond to a truncation error less than $10^{-7}$ and more than 99.8% of the energy for temperature is retained, so this number is used for the reduced basis.

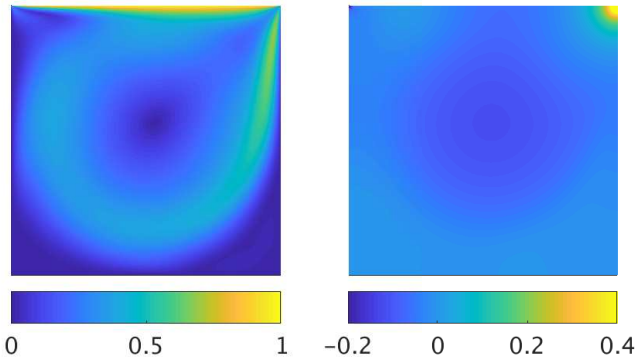


**FIGURE 2** (Left) the background velocity field in m/s and (right) the corresponding pressure field in Pa for the lid-driven cavity simulation at $t = 50$ s.

One reduced matrix, containing all linear terms, is determined as last in the offline phase with the least-squares technique (QR-decomposition) in order to describe the reduced system. The ROM is in the same form as equation (8). In the online phase the ROM is solved for the same initial and boundary conditions as for the FOM and a numerical experimentation is performed on the penalty method on a couple of values for $\tau$, namely 1, 10, 100, 1000 and 10.000. It is found that a penalty factor of 100 enforces the BCs without afflicting the ROM with ill-conditioning problems. Also a factor of 1000 and 10.000 did not lead to unstable solutions. Therefore a penalty factor $\tau = 1000$ is chosen for all ROM simulations. The accuracy of the ROM is checked by calculating the $\ell_2$-error of the temperature field with equation (21) and comparing it with the projection error according to equation (20), here referred to as the basis projection. Both are plotted in Figure 5 on the right. The ROM is describing the system as accurate as projecting the first 31 modes. For the full order simulation a computational time of 165 minutes is required to collect 1000 snapshots. Generating the POD modes and determining the reduced matrix, $\boldsymbol{A_r}$, by solving the least-squares problem in the offline step requires about 6 s and 0.6 s, respectively.

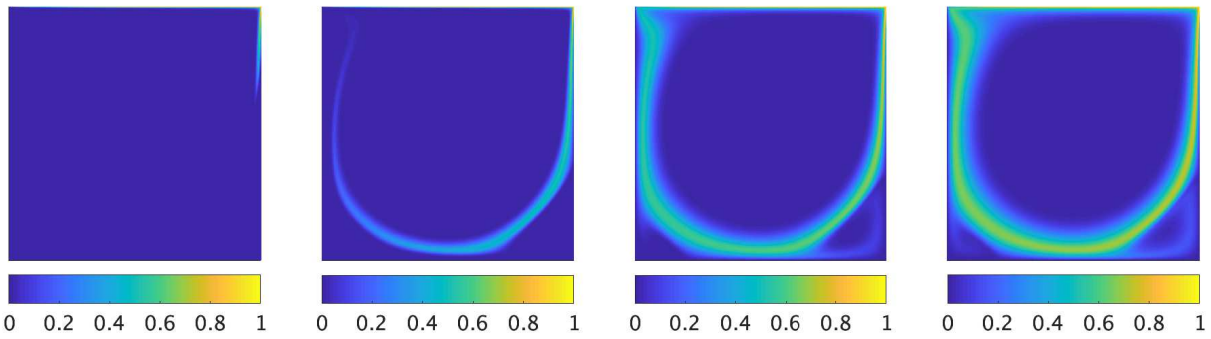


**FIGURE 3** Evolution of the temperature with $T_{lid}^{FOM} = 1$ (base case) inside the cavity for time instances $t = 1$, 10, 50 and 100 s.

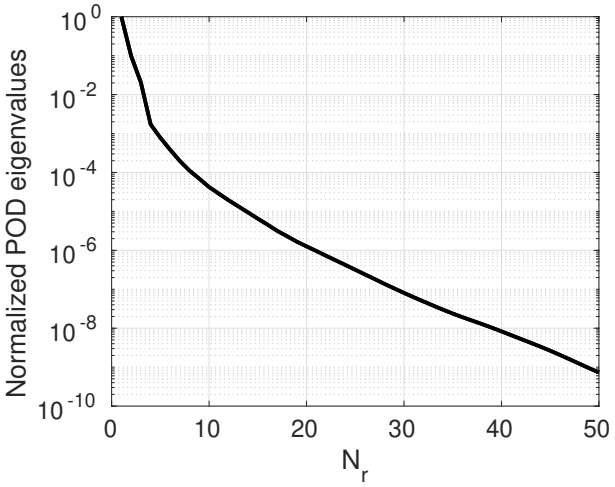


**FIGURE 4** Normalized POD eigenvalues of the base case with time-independent BC.

**TABLE 2** Cumulative eigenvalues (CV) of the base case for temperature

| # of modes | CV of T |
|---|---|
| **1** | 0.71253 |
| **2** | 0.82443 |
| **3** | 0.87342 |
| **4** | 0.90353 |
| **5** | 0.92473 |
| **10** | 0.97132 |
| **20** | 0.99333 |
| **31** | 0.99806 |
| **36** | 0.99901 |

## 5.1 | Imposing time-independent BCs

The constructed ROM is tested for $T_{lid}^{ROM} = 0.5$, 1, 2 and 10. The ROM simulations are performed until $t = 100$ s for a constant time step of 0.1 s. Full order simulations have been performed for comparison. A cross-section of the temperature field for $x = 0.5$ m at $t = 100$ s is plotted in Figure 5 on the left. The accuracy of the ROM is analyzed by calculating the prediction $\ell_2$-error of the temperature field and comparing it with the basis projection error of the FOM snapshots onto the POD basis in Figure 5 on the right. The ROM is describing the system with the same accuracy for any of the tested BCs.

## 5.2 | Imposing time-dependent BCs

In order to construct a ROM for time-dependent BCs, snapshots have to be computed for a FOM with a similar BC as for which the ROM has to be constructed. The time-dependent Dirichlet BC of the FOM is therefore given by equation (28) with amplitude $A = 1$, frequency $f = 0.01$ Hz and offset $T_0 = 1$. The time-dependent BC is enforced in the ROM with the penalty method according to the methodology described in Section 3. Snapshots are collected every 0.1s for the temperature, resulting in a total number of 1000 snapshots. As done previously, 31 modes are used for the ROM construction. No other parametrization is considered and thus only one reduced matrix, containing all linear terms, is determined with the QR-decomposition technique. ROM simulations are carried out for the parameter set 2, 3 and 4 (listed in Table 1). For each of the BCs a full order simulation is performed to compare the ROM solution. Figure 10 shows the comparison of the temperature fields for $t = 100$ s for the FOM, the corresponding ROM and the relative error between the two. Cross-sections of the temperature field at $x = 0.5$ m and at $y = 0.5$ m for $t = 80$ s are plotted in Figure 6 on the left and on the right, respectively.

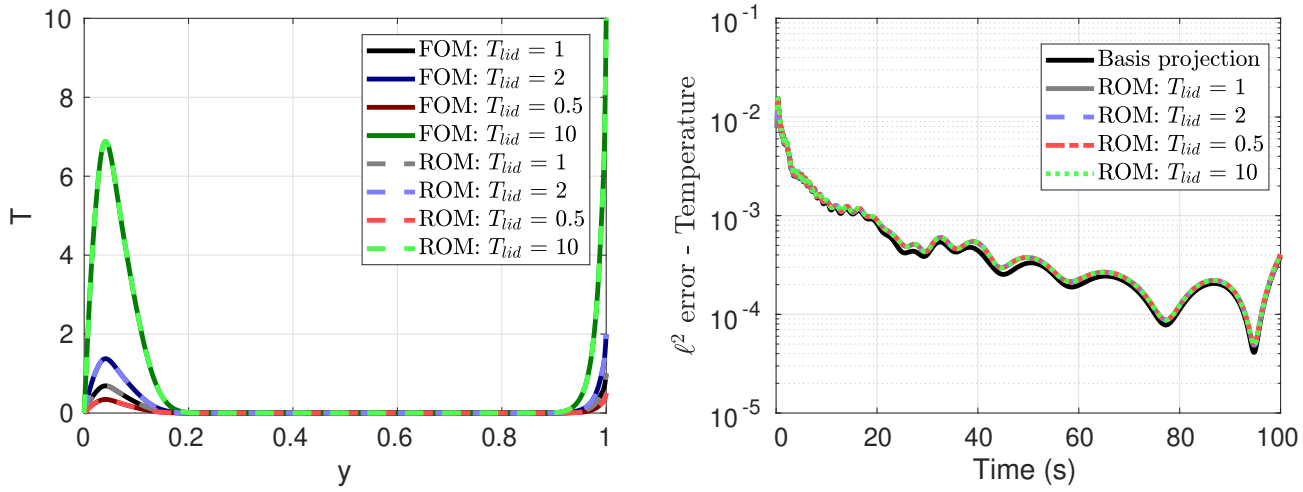


**FIGURE 5** (Left) cross-section of the temperature field for $x = 0.5$ m at $t = 100$ s. $T_{lid}^{FOM} = 1$ is the base case and the ROM is tested for $T_{lid}^{ROM} = 0.5$, 1, 2 and 10 with the penalty method. (Right) the corresponding $\ell_2$-error analysis over time.

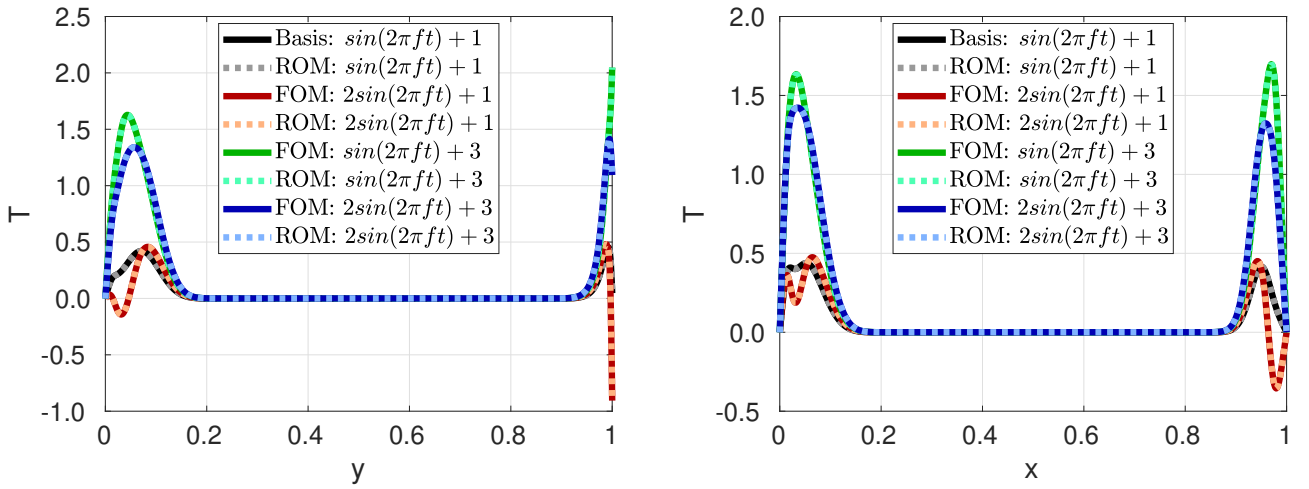


**FIGURE 6** Cross-section of the temperature field at $x = 0.5$ m (left) and $y = 0.5$ m (right) at $t = 80$ s. The time-dependent BCs are listed in the legend with $f = 0.01$ Hz. The BCs are enforced in the ROM with a penalty method.

These figures show that the full order solution cannot simply be scaled as done previously for the time-independent BCs. Nevertheless, the ROMs are capable of approximating the FOM solutions. For each ROM the prediction $\ell_2$-error is compared with the $\ell_2$-error of the temperature reconstruction from the time-dependent coefficient given by projection of the snapshots onto the POD functions in Figure 7. The $\ell_2$-errors for the ROMs are of the same order as the basis projection and decrease, on average, up to about 80 s, meaning that the reduced model is stable in that time interval. The prediction errors slightly increase near the end of the ROM simulation, following the trend of the basis projection error. The error can be reduced by adding more modes to the basis. However, as the POD-ID method requires the system to be overdetermined, the maximum number of modes that can be used for identifying the ROM is set by the square root of the number of snapshots, meaning that 31 modes is near the limit when only 1000 snapshots are used for the POD, instead of only one set as in this work. In order to add more modes to the POD basis, one needs to increase the number of snapshots. Each ROM simulation required a computational time of about 2 s. The speed-up is thus of the order $10^3$ compared with simulation the FOM.

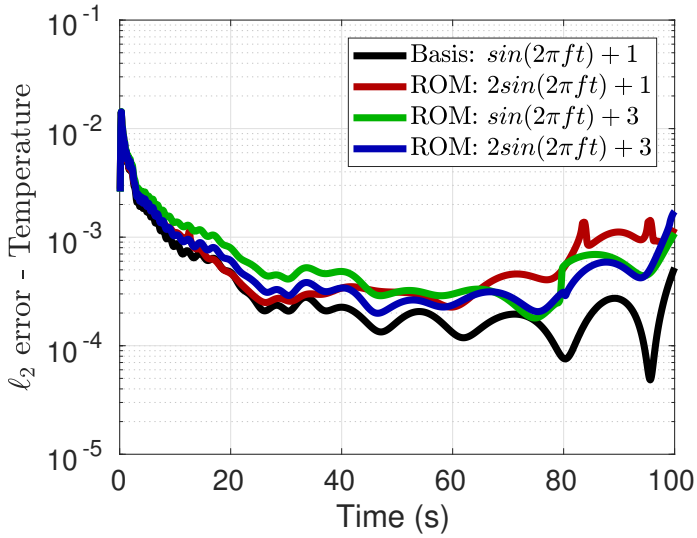


**FIGURE 7** $\ell_2$-error analysis plotted over time for the FOM and ROMs with $f = 0.01$ Hz using 1000 snapshots for the POD basis creation.

Finally, the frequency of the time-dependent BC is parametrized and ROM simulations are carried out for the parameter set 5 and 6 (listed in Table 1) using the same basis (set 1 in Table 1) as for previous cases. Cross-sections of the temperature field at $x = 0.5$ m and $y = 0.5$ m are plotted in Figure 8 on the left and right, respectively. The corresponding $\ell_2$-errors are plotted in Figure 9. In case the frequency is increased by 10%, the $\ell_2$-error is still of the same order of the basis projection. However, doubling the frequency results in an increase of the $\ell_2$-error by one order after about 10 s of simulation time.

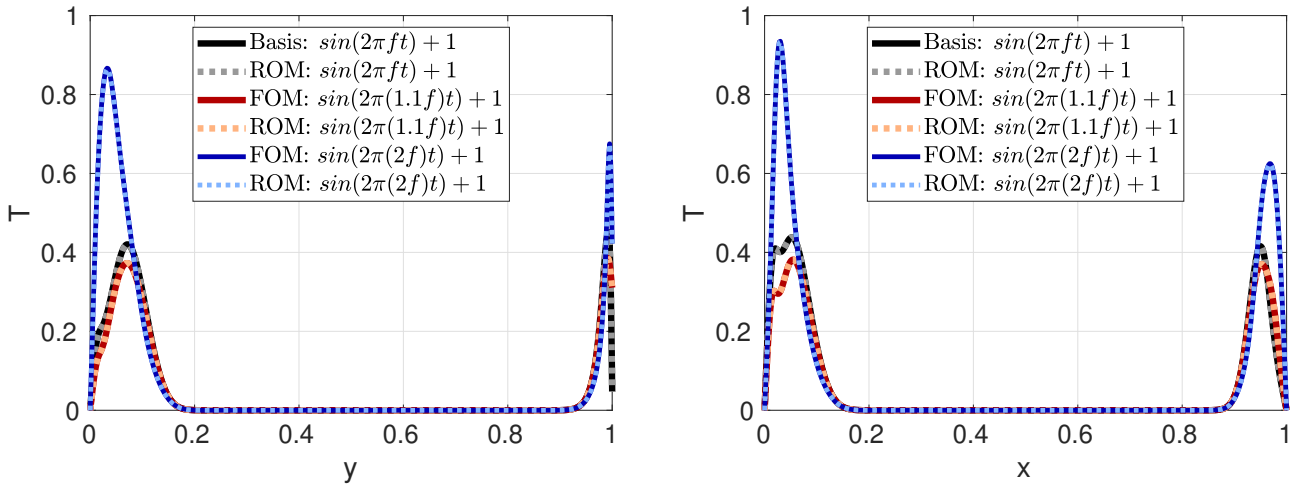


**FIGURE 8** Cross-section of the temperature field at $x = 0.5$ m (left) and y = 0.5 m (right) at $t = 80$ s with time-dependent BCs listed in the legend with $f = 0.01$ Hz.

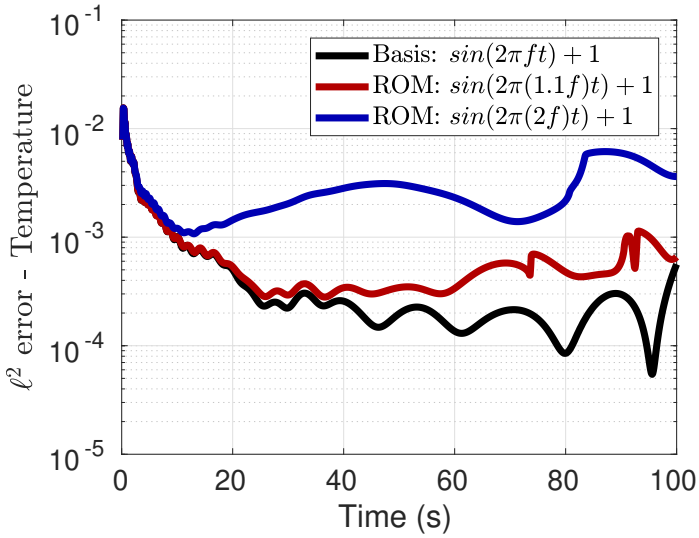


**FIGURE 9** The $\ell_2$-error plotted over time for the FOM and ROMs. The time-dependent BCs are listed in the legend with $f = 0.01$ Hz.

## 6 | DISCUSSION

The ROMs constructed with the presented POD-ID method for controlling the non-homogeneous time-independent Dirichlet BC are capable of approximating the full order solutions for the linear unsteady convection-diffusion equation as the relative error of the main variable, namely the temperature, is of the same order as the projection error of the snapshots onto the POD basis. In this study the penalty factor of 1000 was determined via numerical experimentation. The main advantage of the penalty method is that is can be applied non-intrusively. However, that the factor cannot be determined a priori is a drawback of the method[23]. Also, although it has not been observed here, it is possible that a penalty factor needs to be chosen above a certain threshold to enforce the BC in the ROM, which then will lead to an unstable ROM solution[29].The range of the factor for which the solution is stable can for instance be determined using Poincaré maps[29]. Also the bounds on the factor that ensure asymptotic stability of the ROM can be derived[28]. The prediction error increases when the frequency of the time-dependent BC is doubled. There are two ways to increase the number of snapshots in order to reduce the error and to enhance stability. First of all, more full order sets can be used for the POD. For example, adding full order snapshots for $f = 0.03$ Hz to the snapshot matrix in order to parametrize the frequency in the interval [0.01, 0.03] Hz in the ROM. This can be combined with the second method to increase the snapshot matrix, namely by sampling the full order solution more frequently. Then, more modes can be used to identify the reduced matrix with the least-squares technique as, in order to keep the system overdetermined, the size of the reduced matrix is limited by the square root of the number of snapshots. Even more snapshots would be required in the case of non-linear systems, because then at least as many reduced matrices are to be identified as there are modes to be stored in the offline phase[24,25,26]. For example, the non-linear convective term of the Navier-Stokes equations can be approximated by $\boldsymbol{a}^T \boldsymbol{C}_r \boldsymbol{a}$, where $\boldsymbol{C}_r$ is a third order tensor. Then, the reduced problem grows with the cube of the number of modes in order to maintain an offline-online decomposition. Consequently, more snapshots are required to keep the system overdetermined in order to identify all these matrices. In that case, the required number of snapshots scales with the cube of the number of modes required. Otherwise, many matrices will be empty in case the system is not overdetermined and the problem becomes ill-conditioned. Therefore, it is not fully feasible to use the POD-ID method for fully non-linear problems. In theory, one can solve the ROM for a different time step, $dt$, than used in the FOM, simply rewriting the problem of equation 16 in the following way: $\boldsymbol{a}^{n+1}(1 + \boldsymbol{A}_r dt)^{-1} = \boldsymbol{a}^n$. However, the ROM becomes unstable when approximating the solution for time instances at which no snapshots are collected for the POD, because the POD-ID method identifies a reduced matrix with the least-squares technique that fits the full order snapshot data, like a black-box system, and is not capable of approximating the solution at intermediate time instances. Besides, it is redundant to construct a ROM in order to impose the time-independent BC of the linear problem investigated in this paper. One can simply obtain the results for the parametrized BCs at the same time instances for which snapshots were collected, as parameterizing the Dirichlet BC condition only changes the solution with respect to a reference point. This does, however, not

apply to the control of time-dependent BCs as then the solutions do not scale linearly with respect to a reference point. Then, when intermediate results are required, interpolation techniques could be used to approximate the solution at intermediate times. The same applies also for the long time stability of the ROM.

Finally, this parameterizing the BCs could be done for any given background velocity field, including those of turbulent flows. As long as there is no two-way coupling between the fluid flow and the heat transfer, the problem stays linear and could be described by a single reduced matrix. Finally, the speed-up is of the order of $10^3$. Less modes could be used to speed up the calculation even more, but the error will then increase and the ROM could become unstable.

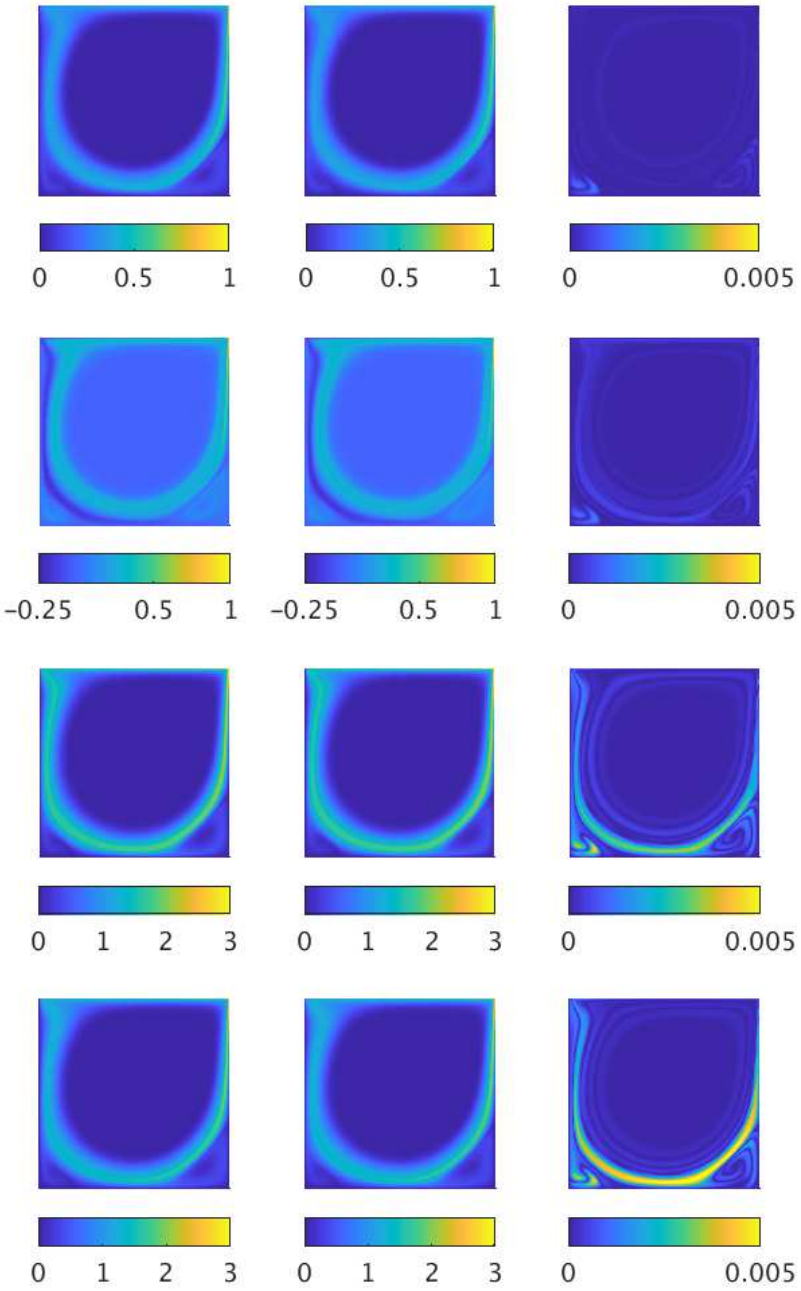


**FIGURE 10** Comparison of the temperature field for the FOM (left) and ROM solutions (middle) at $t = 100$ s for parameter set 1 - 4 (from top to bottom). The absolute difference between the ROM and corresponding FOM is plotted on the right. All ROMs are obtained with 31 modes.

## 7 | CONCLUSIONS AND PERSPECTIVES

In this paper it has been demonstrated that the proposed POD-based Identification method is capable of constructing a ROM that could be used for controlling the non-homogeneous time-(in)dependent Dirichlet BCs of the scalar transport convection - diffusion equations by enforcing the BCs in the ROM with a penalty method rather than having to perform a high fidelity simulation with the Finite Volume approximation for every BC of interest. However, the ROM could only approximate the solution at the same time instances the snapshots are collected for determining the POD basis. Nevertheless, the POD-ID method together with the penalty method could be applied for linear problem that require boundary control. For instance, determining the heat transport by airflow in a room for transient thermal analysis for buildings where the temperature at a wall is the parameter of interest or for pollution dispersion modeling where the concentration is controlled. However, the main shortcoming of the POD-ID method, at this stage, is that it is not feasible to use the method for fully non-linear problems as the required number of snapshots scales with the cube of the number of modes and at least as many reduced matrices are to be identified as the number of modes used. Therefore, further research on non-linear problems is required. Furthermore, in case of a larger parameter space to be investigated, one has to perform the POD on snapshots collected for more parameter values and/or one has to sample more frequently in order to capture the full dynamics of the system. Instead of sampling the full order solution with a higher constant frequency in order to extend the snapshot matrix, the sampling method could be optimized for example with a greedy method[3]. Finally, other types of parametrization could be applied, for instance, parameterizing the diffusion coefficient. However in that case, not just one, but two reduced matrices have to be identified, one for the diffusive and one for the convective term.

## References


1. Hesthaven J, Rozza G, Stamm B. *Certified reduced basis methods for parametrized partial differential equations*. Springer International Publishing. 1st ed. 2016.

2. Chinesta F, Huerta A, Rozza G, Willcox K. *Model Order Reduction*. Wiley Encyclopedia of Computational Mechanics. 2nd ed. 2016.

3. Quarteroni A, Rozza G. *Reduced order methods for modeling and computational reduction*. 9 of *MS&A*. Springer. 1st ed. 2014.

4. Benner P, Gugercin S, Willcox K. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM review* 2015; 57(4): 483–531. doi: 10.1137/130932715

5. ANSYS Inc. Fluent 18.2 Getting Started Guide. 2017.

6. Jasak H, Jemcov A, Tuković Ž. OpenFOAM: A C++ library for complex physics simulations. In: International workshop on coupled methods in numerical dynamics, IUC, Dubrovnik, Croatia. ; 2007.

7. Eymard R, Gallouët T, Herbin R. Finite volume methods. *Handbook of numerical analysis* 2000; 7: 713–1018.

8. Versteeg HK, Malalasekera W. *An introduction to computational fluid dynamics: the finite volume method*. 2nd ed. 2007.

9. Fletcher C. *Computational Techniques for Fluid Dynamics Vol. I. Fundamental and General Techniques*. Springer-Verlag. 2nd ed. 1991.

10. Weller J, Lombardi E, Bergmann M, Iollo A. Numerical methods for low-order modeling of fluid flows based on POD. *International Journal for Numerical Methods in Fluids* 2010; 63(2): 249–268. doi: 10.1002/fld.2025

11. Kunisch K, Volkwein S. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM Journal on Numerical analysis* 2002; 40(2): 492–515. doi: 10.1137/S0036142900382612

12. Xiao D, Fang F, AG B, C P, Navon I, Muggeridge A. Non-intrusive reduced-order modelling of the Navier-Stokes equations based on RBF interpolation. *International Journal for Numerical Methods in Fluid* 2015; 79: 580–595. doi: 10.1002/fld.4066

13. Xiao D, Yang P, Fang F, et al. A non-intrusive reduced-order model for compressible fluid and fractured solid coupling and its application to blasting. *Journal of Computational Physics* 2016; 330: 221–224. doi: 10.1016/j.jcp.2016.10.068

14. Xiao D, Fang F, Pain C, IM N. A parameterized non-intrusive reduced order model and error analysis for general time-dependent nonlinear partial differential equations and its applications. *COMPUTER METHODS IN APPLIED MECHANICS AND ENGINEERING* 2017; 317: 868–889. doi: 10.1016/j.cma.2016.12.033

15. Xiao D, Heaney C, Mottet L, et al. A reduced order model for turbulent flows in the urban environment using machine learning. *Building and Environment* 2019; 148: 323–337. doi: 10.1016/j.buildenv.2018.10.035

16. Walton S, Hassan O, Morgan K. Reduced order modelling for unsteady fluid flow using proper orthogonal decomposition and radial basis functions. *Applied Mathematical Modelling* 2013; 37(20): 8930–8945. doi: 10.1016/j.apm.2013.04.025

17. Audouze C, De Vuyst F, Nair P. Nonintrusive reduced-order modeling of parametrized time-dependent partial differential equations. *Numerical Methods for Partial Differential Equations* 2013; 29(5): 1587–1628. doi: 10.1002/num.21768

18. Viberg M. Subspace-based methods for the identification of linear time-invariant systems. *Automatica* 1995; 31(12): 1835–1851. doi: 10.1016/0005-1098(95)00107-5

19. Schmid PJ. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics* 2010; 656: 5–28. doi: 10.1017/S0022112010001217

20. Vervecken L, Camps J, Meyers J. Stable reduced-order models for pollutant dispersion in the built environment. *Building and Environment* 2015; 92: 360–367. doi: 10.1016/j.buildenv.2015.05.008

21. Weller HG, Tabor G, Jasak H, Fureby C. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computers in physics* 1998; 12(6): 620–631. doi: 10.1063/1.168744

22. Annoni J, Seiler P. A method to construct reduced-order parameter-varying models. *International Journal of Robust and Nonlinear Control* 2017; 27(4): 582–597. doi: 10.1002/rnc.3586

23. Graham W, Peraire J, Tang K. Optimal control of vortex shedding using low-order models. Part I - open-loop model development. *International Journal for Numerical Methods in Engineering* 1999; 44(7): 945–972. doi: 10.1002/(SICI)1097-0207(19990310)44:7<945::AID-NME537>3.0.CO;2-F

24. Stabile G, Rozza G. Finite volume POD-Galerkin stabilised reduced order methods for the parametrised incompressible Navier–Stokes equations. *Computers & Fluids* 2018; 173: 273–284. doi: 10.1016/j.compfluid.2018.01.035

25. Lorenzi S, Cammi A, Luzzi L, Rozza G. POD-Galerkin method for finite volume approximation of Navier–Stokes and RANS equations. *Computer Methods in Applied Mechanics and Engineering* 2016; 311: 151–179. doi: 10.1016/j.cma.2016.08.00

26. Stabile G, Hijazi S, Mola A, Lorenzi S, Rozza G. POD-Galerkin reduced order methods for CFD using Finite Volume Discretisation: vortex shedding around a circular cylinder. *Communications in Applied and Industrial Mathematics* 2017; 8(1): 210–236. doi: 10.1515/caim-2017-0011

27. Annoni J, Gebraad P, Seiler P. Wind farm flow modeling using an input-output reduced-order model. In: IEEE. ; 2016: 506–512

28. Kalashnikova I, Barone M. Efficient non-linear proper orthogonal decomposition/Galerkin reduced order models with stable penalty enforcement of boundary conditions. *International Journal for Numerical Methods in Engineering* 2012; 90(11): 1337–1362. doi: 10.1002/nme.3366

29. Sirisup S, Karniadakis G. Stability and accuracy of periodic flow solutions obtained by a POD-penalty method. *Physica D: Nonlinear Phenomena* 2005; 202(3-4): 218–237. doi: 10.1016/j.physd.2005.02.006

30. Bizon K, Continillo G. Reduced order modelling of chemical reactors with recycle by means of POD-penalty method. *Computers & Chemical Engineering* 2012; 39: 22–32. doi: 10.1016/j.compchemeng.2011.10.001

31. Botella O, Peyret R. Benchmark spectral results on the lid-driven cavity flow. *Computers & Fluids* 1998; 27(4): 421–433. doi: 10.1515/caim-2017-0011

32. Schreiber R, Keller HB. Driven cavity flows by efficient numerical techniques. *Journal of Computational Physics* 1983; 49(2): 310–333. doi: 10.1016/0021-9991(83)90129-8

33. Jasak H. *Error analysis and estimation in the Finite Volume method with applications to fluid flows*. PhD thesis. Imperical College, University of London, 1996.

34. Moukalled F, Mangani L, Darwish M, others . The finite volume method in computational fluid dynamics. *An Advanced Introduction with OpenFOAM and Matlab* 2016: 3–8.

35. Taher M, Saha S, Lee Y, Kim H. Numerical study of lid-driven square cavity with heat generation using LBM. *American Journal of Fluid Dynamics* 2013; 3(2): 40–47. doi: 10.5923/j.ajfd.20130302.04

36. The MathWorks Inc. MATLAB Release 2017a. 2017.