A Greedy Non-Intrusive Reduced Order Model for Fluid Dynamics☆

Wang Chen[a], Jan S Hesthaven[b], Bai Junqiang[a,*], Zhang Yang[a], Yang Tihao[a]

[a]*Northwestern Polytechnical University, Xi'an 710072, P.R. China*
[b]*Swiss federal Institute of Technology in Lausanne(EPFL), CH-1015 Lausanne, Switzerland*

Abstract

We introduce a greedy non-intrusive reduced order method(ROM) for parameterized time-dependent problems with an emphasis on problems in fluid dynamics. The non-intrusive ROM is based on two-level proper orthogonal decomposition(POD) to extract temporal and spatial reduced basis from a set of candidates and a radial basis function(RBF) model to be used to approximate the coefficients of the reduced model. And instead of adopting uniform or random sampling strategies, the candidates are determined using an adaptive greedy approach to minimize the overall computational offline cost. Numerical studies are presented for a two-dimensional diffusion problem and a driven cavity problem governed by incompressible Navier-Stokes equations. The results demonstrate that the greedy non-intrusive ROM predicts the flow field accurately and efficiently.

Keywords: Greedy Approach, Two-level Proper Orthogonal Decomposition, Radial basis function, Non-intrusive Reduced order method, Parameterized time-dependent PDEs, Fluid dynamics

1. Introduction

Even though the ability to computationally model complex phenomena continues to increase, it remains challenging to many types of problems in optimization, design, control etc due to the need for a high number of computations for various parameters. To address this problem, the development of reduced-order method(ROM) [1, 2, 3, 4, 5, 6] that enables the approximated solutions with an acceptable loss of accuracy has gained a substantial interest as a compromise to solving the full problems. In reduced order models, the full problem is expressed by projecting the full system to a problem of specific subspace with much smaller degree of freedoms(DOFs). When properly chosen, this reduced linear space can adequately represent the dynamics of the full system under parameter variation at substantially reduced cost. Examples of some of these recent developments include proper orthogonal decomposition(POD) [7, 8], Harmonic Balance approach(HB)[9], and Volterra theory[10, 11].

Many ROMs extract a set of reduced basis from snapshots of the full solution and then project the full system onto the linear subspace determined by those basis modes through a projection approach

*Corresponding author
Email address: junqiang@nwpu.edu.cn (Bai Junqiang)

such as Galerkin projection[12, 13], called Projection-based ROMs. To approximate solutions, projection coefficients are determined through the projection process, and thus approximations are made. Although Projection-based ROMs are effective, it brings much inconvenience to compute projection coefficients, especially for complex nonlinear problems. Certainly this is not what we expect. To address this concern, there is a recent interest in non-intrusive ROMs[14] which seeks to develop ROMs based solely on access to snapshots. In other words, governing equations only play the role of providing snapshots and has nothing to do with the projection process. In the context of fluid dynamics, [15] presents a second order Taylor series method and an approach based on a Smolyak sparse grid collocation method, and applies these to simulate the flow past a cylinder with wind driven gyre. In [14], the authors explored the use of empirical interpolation method [16] to recover a reduced model based on a non-intrusive approach.

And development of ROMs for parametrized time-dependent problems is another issue as most researches are focused on parametrized steady problems. Compared to parametrized steady problems, ROMs for time-dependent parametrized partial differential equations(PDEs) have to approximate solutions as a function of time, spatial coordinates and a parameter vector, which turns out to be more challenging especially when the boundary conditions vary with time. To solve unsteady parametrized Navier-Stokes equations, [17] utilizes POD-Galerkin in patient-specific haemodynamics while [18] uses discrete empirical interpolation method. Some other work on this topic is dealing with simple parametrized time-dependent PDEs such as convection-diffusion problems [19] and boussinesq equations[20]. And in this paper we prefer to adopt a ROM of two-level POD plus RBF proposed in [21] for approximating parametrized time-dependent PDEs because of its non-intrusive and efficient feature.

Apart from the desire to develop ROMs of a non-intrusive nature, it is likewise important to ensure the accuracy of the reduced models by carefully selecting the snapshots. Due to the variation of possible solutions, it is not possible to determine the optimal candidates to generate snapshots. One approach is to use a random strategy such as Latin Hypercube sampling or minimum discrepancy sequences[22, 23], CVT[24], Monte Carlo methods[25], Sparse Grid[26] etc to generate snapshots and manipulate these to recover a suitable basis, e.g., by the use of a singular value decomposition(SVD). However, to guarantee the accuracy of the reduced model, an excessive number of snapshots may be required, leading to unacceptable computational cost. To address this concern, recent work has been devoted to strategies for guiding the selection of snapshots in a problem dependent manner, in spirit similar to adaptive sampling[27] used in surrogate models. In [28], an algorithm that iteratively applies surrogate-model optimization in the generation of snapshots is developed. A related approach is based on a greedy algorithm in which case the snapshots are adaptively determined by finding the location in the predetermined parameter space at which the error estimator in the reduced model is maximum, quantifying the quality of the ROM. In [29, 30], a POD-Greedy approach is introduced to recover reduced order models, albeit in an intrusive manner, with a discussion of convergence offered in [31, 32]. And Patera et al. has researched a lot on Greedy sampling for POD-Galerkin method when approximating parametrized problems[33, 34, 35, 36].

While Tan Bui-Thanh and Karen Willcox also proposed a model-constrained adaptive Greedy sampling methodology for large-scale systems with high-dimensional parametric input spaces[37]. However, not much work to develop greedy sampling has been done for non-intrusive methods.

In this work, our goal is to develop a greedy sampling for a non-intrusive ROM appropriate for time-dependent parametrized PDEs. Based on an approach set forward in [21], we first develop the non-intrusive ROM which employs a two-level approach for the construction of the spatial and temporal basis functions and estimates the undetermined coefficients of the reduced model by using a radial basis approximation(RBF). While the discussion has some degree of generality, we shall focus on problems and examples in the context of fluid dynamics. In contrast to previous work in [21], the reduced model adaptively generate snapshots based on a greedy approach to minimize the overall computational cost. We shall demonstrate the quality and flexibility on problems in heat conduction and driven cavity flows.

What remains of the paper is organized as follows. In Section 2, we develop the non-intrusive ROM including a two-level basis development and discuss the approximation based on the RBF. We also briefly discuss the use of the RBF-QR method [38] in this context to overcome ill-conditioning in RBF as the number of candidates increases. In Section 3, we develop a Greedy sampling approach to reduce the computational cost of the generation of the reduced basis. This sets the stage for Section 4 which contains several numerical verification cases, confirming the efficiency of the greedy non-intrusive ROM. Section 5 contains a few concluding remarks and points to additional developments.

2. Construction of the non-intrusive reduced order model

In the following, we focus on the construction of the non-intrusive ROM for parametrized time-dependent PDEs. The basic assumption is that the reduced basis solution can be expressed by the biorthogonal decomposition(BOD)[39] as

$$\hat{u}^{\theta}(x,t) = \hat{v}^{\theta}(x,t) + \sum_{k=1}^{K}\sum_{m=1}^{M} \alpha^{\theta}(k,m)\varphi^{k}(x)\xi^{m}(t). \tag{1}$$

Here $\hat{v}^{\theta}(x,t)$ denotes the auxiliary term created to properly account for boundary and initial condition and we return to the details of this shortly.

In (1), two types of basis functions are introduced: spatial basis functions $\varphi^{k}(x)$ and temporal basis functions $\xi^{m}(t)$. The coefficients $\alpha^{\theta}(k,m)(k = 1, 2, \cdots, K; m = 1, 2, \cdots, M)$ denote the undetermined coefficients in the approximation.

We begin the discussion by focusing on the approach proposed in [21] in which the basis is constructed using a two-level POD. This is achieved in two parts of the method, resulting in both spatial and temporal basis functions $\varphi^{k}(x), \xi^{m}(t)$. These are subsequently combined with a RBF to estimate undetermined coefficients $\alpha^{\theta}(k,m)$. In the following we outline this process in more detail.

2.1. Basis creation by proper orthogonal decomposition

The proper orthogonal decomposition(POD), first developed to correlate statistical data[40], has been widely used to recover eigenmodes for analysis of spatio-temporal problems in many fields, e.g., fluid dynamics[41], aeroelastic dynamics[42] and so on. The central idea is to determine an optimal linear basis from ensembles of statistical data with optimality measured in a least-square sense.

Let us consider a function $Y^{\theta}(x)(a \leq x \leq b)$ and assume that we have snapshots $y_i(x)(i = 1, 2, \cdots, n)$ where n denotes the number of snapshots. We now seek an optimal basis $\Phi_i(i = 1, 2, \cdots K)$ such that

$$\min_{v_1 \ldots v_k} \sum_{i=1}^{n} \|y_i(x) - \hat{y}_i(x)\|_2^2, \hat{y}_i(x) = y_0(x) + \sum_{i=1}^{K} a_i(\theta) \cdot \Phi_i(x),$$

under the condition that

$$(\Phi_i, \Phi_j) = \delta_{ij}, \quad (u, v) = \int_a^b uv dx,$$

and

$$y_0(x) = \frac{1}{n} \sum_{i=1}^{n} y_i(x).$$

The optimization problem to achieve $\Phi_i(x)$ can be transformed into a new one:

$$\max \sum_{i=1}^{n} ||y_i(x), \Phi_i(x)||^2 \tag{2}$$

If we define the correlation matrix $M \in R_{n \times n}$ with the elements defined as

$$M(i, j) = \int_a^b y_i(x) y_j(x) dx, \tag{3}$$

it is clearly seen that M is a symmetric positive semi-definite matrix and the eigensolutions

$$M \cdot v_i = \lambda_i v_i \qquad\qquad i = 1, 2, \cdots, n \tag{4}$$

yields the sought after basis, truncated at rank i. Then the reduced basis can be determined by

$$\Phi_i(x) = \sum_{j=1}^{n} v_i(j) \cdot y_j(x),$$

which are often referred to as the proper orthogonal modes. Defining the threshold $0 < \varepsilon < 1$, the first K eigenvalues $\lambda_k(k = 1, 2, \cdots K)$, arranged in an descending order, are determined through

$$K = min\{k \in [1, n] | \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \geq \varepsilon\} \tag{5}$$

which also quantifies the error of the truncation as

$$\sum_{j=1}^{n} \|y_i - \sum_{i=1}^{K} (y_i, v_i) v_i\|_2^2 \leq \frac{\sum_{j=K+1}^{n} \lambda_j}{\sum_{j=1}^{n} \lambda_j}.$$

The optimal basis can likewise be computed using an SVD of the matrix of snapshots. In this case, the sum of the singular values of the matrix offers the error estimator for the accuracy of the truncated linear space.

2.2. Problem statement

Let us now begin to consider the general parametrized time-dependent PDEs:

$$\begin{cases} \frac{\partial u^\theta}{\partial t} + F(u^\theta) = f^\theta & x \in \Omega \times t \in (0,T] \\ u^\theta_{|t=0} = u^\theta_0(x) & x \in \Omega \\ u^\theta(x,t) = g^\theta(x,t) & x \in \partial\Omega \times t \in (0,T] \end{cases} \tag{6}$$

where $t \in [0,T]$ denotes time, $x \in \Omega$ are the spatial coordinates, and Ω is the physical domain with a boundary $\partial\Omega$. Let also θ represents a vector of parameters with dimension p($p \geq 1$). We assume that the parameters belong to the range B. Finally, we refer to u^θ as the solution which depends on space, time and θ. Initial and boundary conditions are likewise presented in (6).

The general problem includes three spaces to take into consideration: the temporal space, the physics space, and the parameter space, respectively. To express this, we consider the ansatz:

$$\hat{u}^\theta(x,t) = \hat{v}^\theta(x,t) + \sum_{k=1}^{K}\sum_{m=1}^{M} \alpha^\theta(k,m)\varphi^k(x)\xi^m(t) \tag{7}$$

As already indicated, we consider the spatial basis functions $\varphi^k(x)$ and the temporal basis functions $\xi^m(t)$. The auxiliary function $\hat{v}^\theta(x,t)$, is required to satisfy

$$\begin{cases} \varphi^k_{|\partial\Omega} = 0 \\ \xi^m(0) = 0 \end{cases} \tag{8}$$

$$\begin{cases} \frac{\partial \hat{v}^\theta}{\partial t} - \triangle \hat{v}^\theta = 0 & x \in \Omega \times t \in (0,T] \\ \hat{v}^\theta_{|t=0} = u^\theta_0(x) & x \in \Omega \\ \hat{v}^\theta(x,t) = g^\theta(x,t) & x \in \partial\Omega \times t \in (0,T] \end{cases} \tag{9}$$

hence encoding the impact of the initial and boundary conditions.

In this paper we only consider problems with constant boundary conditions and setting $\hat{v}^\theta(x,t) = u_0(x)$ can be our choice in this case which satisfies (9). To see this, consider the initial condition,

$$\hat{u}^\theta(x,0) = \hat{v}^\theta(x,0) + \sum_{k=1}^{K}\sum_{m=1}^{M} \alpha^\theta(k,m)\varphi^k(x)\xi^m(0) \tag{10}$$

The second term on the right hand side in (10) vanishes because of the assumption that $\xi^m(0)_{|m=1,2,\cdots,M} = 0$. By introducing $\hat{v}^\theta(x,t) = u_0(x)$, we have

$$\hat{u}^\theta(x,0) = u_0(x) \tag{11}$$

For the boundary condition, the term $\sum_{k=1}^{K}\sum_{m=1}^{M}\alpha^{\theta}(k,m)\varphi^{k}(x)\xi^{m}(0)$ can also be ignored with $\varphi^{k}_{|\partial\Omega}=0$, thus

$$\hat{u}^{\theta}_{|\partial\Omega}(t)=\hat{u}^{\theta}_{|\partial\Omega}(t=t_0)=g^{\theta} \tag{12}$$

Therefore $\hat{v}^{\theta}(x,t)=u_0(x)$ satisfies (9) and the boundary and initial conditions are satisfied.

2.3. Two-level basis development

To develop a reduced order model for the parametrized time-dependent PDEs (6), two-level approach, based on POD, was adopted in [21] to construct both the temporal and the spatial basis. As the basis generation for spatial and temporal variation is similar, we discuss just one.

We first define a coarse spatial grid $\chi=\chi_j\in\Omega_h(j=1,2,\cdots,N_x)$ and a coarse temporal one with N_T time instants as $\Upsilon=\{t_n|0=t_1<t_2<\cdots t_{N_T}=T\}$ to generate snapshots. Furthermore, we define a set of candidates $\theta_i(i=1,2,\cdots,N)$ from the given parameter space.

By executing the accurate solver we recover shifted spatial snapshots $\Psi=\{u^{\theta_i}(\cdot,t_j)-v^{\theta_i}(\cdot,t_j),j=1,2,\cdots N_T\}$. For each parameter point θ_i, we adopt the POD onto the temporal snapshots to recover K^i eigenvectors $r^{k,i}(k=1,2,\cdots K^i)$. The corresponding spatial modes are finally recovered as

$$\varphi^{k,i}(x)=\sum_{n=1}^{N_T}(r^{k,i})_n(u^{\theta_i}(x,t_n)-v^{\theta_i}(x,t_n)) \tag{13}$$

From the efforts above, we get a total of $\sum_{i=1}^{N}K^i$ spatial basis. And by applying an SVD, we finally compress those eigenfunctions and retain K Spatial modes $\varphi^{k}(k=1,2,\cdots K)$.

Temporal snapshots are given by $\{u^{\theta}_i(\chi_j,\cdot)-v^{\theta}_i(\chi_j,\cdot),j=1,2,\cdots N_x\}$. Using the similar approach, we recover M temporal modes $\xi^{m}(m=1,2,\cdots M)$.

2.4. RBF approximation

The final challenge is to estimate the coefficients α^{θ} at any parameter point in (7) while retaining a good approximation accuracy. For this, we employ a RBF method to compute α^{θ}.

At first, we express (7) as

$$\hat{u}^{\theta}=\hat{v}^{\theta}+\varphi\alpha^{\theta}\xi^{T} \tag{14}$$

A least square approach[43] is adopted to determine coefficients for candidates $\theta_i(i=1,2,\cdots N)$, namely $\alpha^{\theta_i}(i=1,2,\cdots N)$

$$\min_{\alpha\in\mu_{KM}}(||\hat{u}^{\theta}-\hat{v}^{\theta}-\varphi\alpha^{\theta}\xi^{T}||_2^2+\mu||\alpha^{\theta}||_2^2) \tag{15}$$

where $\mu\geq 0$ is a penalized parameter to avoid overfitting and $||\cdot||_2$ represents the 2-norm for a matrix with the definition $||A||_2=tr(A^TA)$.

By substituting shifted snapshots u^i-v^i, coefficients α^{θ_i} corresponding to θ_i can be computed by

$$(\varphi^T\varphi)\alpha^{\theta_i}(\xi^T\xi)+\mu\alpha^{\theta_i}=\varphi^T(u^i-v^i)\xi \tag{16}$$

For fixed values of k, m, we expand $\alpha^{\theta}(k,m)$ for θ as follows:

$$\alpha^{\theta}(k,m) = \sum_{i=1}^{N} \gamma_i^{km} \phi(||\theta - \theta_i||_2) \qquad (k = 1, 2, \cdots K; m = 1, 2, \cdots M) \tag{17}$$

where $\phi(r)$ is the chosen RBF kernal function and γ_i^{km} denotes the undetermined coefficients, which are independent on θ.

By substituting θ_j into (17) and relating α^{θ_j} in (16), $\gamma_i^{km}(i = 1, 2, \cdots N; k = 1, 2, \cdots K; m = 1, 2, \cdots M)$ are computed as

$$\alpha^{\theta_j}(k,m) = \sum_{i=1}^{N} \gamma_i^{km} \phi(||\theta_j - \theta_i||_2) \qquad (i = 1, 2, \cdots N; k = 1, 2, \cdots K; m = 1, 2, \cdots M) \tag{18}$$

Finally from $\gamma_i^{km}(i = 1, 2, \cdots N; k = 1, 2, \cdots K; m = 1, 2, \cdots M)$, we can recover α for arbitrary θ value in parameter space though (17).

In this work we use a Gaussian radial basis function[44], defined as :

$$\phi(r) = \exp\left(-\frac{r^2}{2\sigma}\right),$$

with a parameter $\sigma > 0$. We define a matrix

$$\Phi_{i,j} = \phi(||\theta_i - \theta_j||_2) \quad (i, j = 1, 2, \cdots N) \tag{19}$$

When σ is very small, Φ defined in (19), used in RBF, will be ill-conditioning. As a result, the approximation will break down. On the contrary, if σ is very large, the approximation accuracy gets affected.

To address this problem, which becomes particularly problematic when the number of candidates increases or an unstructured sampling pattern in parameter space is used, we adopt the RBF-QR, proposed in [38]. In this approach, one eliminates the effect of σ by QR decomposition and restored high accuracy even for very small values of σ.

At first, the standard RBF is transformed into

$$\phi(r) = C \cdot D \cdot T(r) \tag{20}$$

where $T(r)$ is a matrix consisting of Chebyshev splines, D is a diagonal matrix of σ's increasing powers, $o(\sigma^{2j})(j = 1, 2, \cdots \infty)$, and C is a matrix with elements $o(1)$, only dependent on design variables r. And the ill-conditioning problems mainly stem from the diagonal matrix D. [38] proposed to adopt QR decomposition on C, and then multiply $\phi(r)$ with $D_N^{-1} R_N^{-1} Q^T$, forming a new RBF:

$$\psi(r) = D_N^{-1} \cdot R_N^{-1} \cdot Q^T \cdot \phi(r) = D_N^{-1} \cdot R_N^{-1} \cdot Q^T \cdot Q \cdot R \cdot D \cdot T(r) = D_N^{-1} \cdot R_N^{-1} \cdot R \cdot D \cdot T(r) \tag{21}$$

D_N^{-1} and R_N^{-1} is the first $N \times N$ submatrix of D and R respectively. It is predicted in [38] that the matrix $D_N^{-1} \cdot R_N^{-1} \cdot R \cdot D$ is well conditioned in most cases, leading to $\psi(r)$ more stable than $\phi(r)$. And we shall discuss RBF-QR's advantage in our scheme further in Section 4.

2.5. Overview

Once the spatial and temporal basis functions are determined and the coefficients α^θ are expressed through the expansion of RBF, the solution can be approximated at any θ in parameter space through (7). Algorithm 1 outlines the steps of the non-intrusive ROM.

Algorithm 1 Reduced-order method approach

1: Generate snapshots with an appropriate numerical solver.
2: **for** each $\theta_i (i = 1, 2, \cdots, N)$ **do**
3: set a threshold ε_1 and apply POD to extract a set of temporal modes $\varphi^{k,i}$ and spatial modes $\xi^{m,i}$;
4: **end for**
5: Collect all the temporal and spatial basis functions, and then apply SVD to determine the final required reduced basis $\varphi^k (k = 1, 2, \cdots K)$ and $\xi^m (m = 1, 2, \cdots M)$ with another threshold ε_2.
6: Utilize the reduced basis, snapshots and appropriate RBF $\phi(r)$ to compute γ_i^{km}.
7: Detemine coefficients α^θ at any condition using (17) and then approximate the solution with (7).

3. A greedy approach for model development

The main shortcoming with the algorithm discussed so far is the selection of the samples in parameter space. In [21] this was done in a simple manner, having the potential for computing a large number of samples when only few are needed to ensure an approximation of the required accuracy. While in this paper we pursue the development of a greedy algorithm to enable an optimal sampling in parameter space.

3.1. Greedy Sampling

The core idea of the greedy algorithm is to adaptively selecting samples iteratively by locating the new sample at each iteration where the estimated error in the reduced model reaches the maximal in parameter space. The sampling initiates from N^0 candidates. Directly from the ROM presented above, we propose the error estimator at the iter-th iterative step as

$$\Delta(iter) = ||u^\theta(\cdot) - \hat{u}^\theta(\cdot)||_2 = ||u^\theta(\cdot) - v^\theta(\cdot) - \varphi^e \alpha^\theta \xi^{eT}||_2 \tag{22}$$

where φ^e and ξ^e represents basis functions at current greedy iterative step. Here $||\cdot||_2$ still denotes the squared norm for a matrix with $||A||_2 = tr(A^T A)$. And θ in parameter space which satisfies $||(u^\theta - v^\theta - \varphi^e \alpha^\theta \xi^{eT})||_2$ corresponds to the new candidate to be added, denoted by θ_s. However, as u^θ and v^θ is undetermined except when the numerical solver is used, we cannot determine θ_s directly with this error estimator (22),otherwise it is very expensive.

To develop an efficient alternative, let us recall the process of computing α^θ using the least square as well as RBF approximation in Section 2. To recover the undetermined matrix γ_i^{km}, we substitute α^{θ_l} computed by (16) for the corresponding θ_l into (17) as

$$\alpha^{\theta_l} = \sum_{i=1}^{N^e} \gamma_i^{km} \left(\phi \left(\frac{|\theta_l - \theta_i|}{\sigma} \right) \right) \qquad l = 1, 2, \cdots N^e \tag{23}$$

Where $N^e = iter + N^0$ denotes the candidate number at current step. From this, it is easy to solve γ_i^{km} and α^θ can be computed for any θ using (17). Finally, we recover the approximation solution $\hat{u}^\theta$ by using (7) and α^θ. It is clear that the approximated solution for any θ is recovered similarly to the interpolation of snapshots in parameter space. The difference is that it interpolates the coefficients $\alpha^\theta(k, m)$ of each candidate, instead of interpolating the snapshots directly. Therefore we suggest to take the point where the mean distance between α^θ and α^{θ_l} is the biggest as the new candidate.

The distance between α^{θ_l} and α^θ can be expressed by

$$D_l = \alpha^\theta - \alpha^{\theta_l} = \sum_{i=1}^{N^e} \gamma_i^{km} (\phi(\frac{|\theta - \theta_i|}{\sigma}) - \phi(\frac{|\theta_l - \theta_i|}{\sigma})) \qquad l = 1, 2, \cdots, N^e \tag{24}$$

.

Thus we propose an alternative for error estimator on the form in 2-norm sense based on α distance.

$$\begin{aligned} \Delta_1(iter) &= \frac{\sum_{l=1}^{N^e} || \sum_{k=1}^{K} \sum_{m=1}^{M} [(\alpha^\theta(k,m) - \alpha^{\theta_l}(k,m)) \varphi^k(x) \xi^m(t)]||_2}{N^e} \\ &= \frac{\sum_{l=1}^{N^e} || \sum_{k=1}^{K} \sum_{m=1}^{M} [(\sum_{i=1}^{N^e} \gamma_i^{km} (\phi(\frac{|\theta-\theta_i|}{\sigma}) - \phi(\frac{|\theta_l-\theta_i|}{\sigma}))) \varphi^k(x) \xi^m(t)]||_2}{N^e} \\ &= \frac{\sum_{l=1}^{N^e} ||\sum_{i=1}^{N^e} (\phi(\frac{|\theta-\theta_i|}{\sigma}) - \phi(\frac{|\theta_l-\theta_i|}{\sigma}))(\sum_{k=1}^{K} \sum_{m=1}^{M} (\gamma_i^{km}) \varphi^k(x) \xi^m(t))||_2}{N^e} \end{aligned} \tag{25}$$

where $\sum_{k=1}^{K} \sum_{m=1}^{M} (\gamma_i^{km} \varphi^k(x) \xi^m(t))$ is independent on θ and considered as coefficients.

The cost of computing (25) directly remains too high to be practical in the Greedy method if θ is considered as a continuous variable vector. Thus we select a series of $\theta_{ii}^*(ii = 1, 2, \cdots N^*)$ discrete points in parameter space, and identify the new candidate from those points. Subsequently, we compute Δ_1 for each sample point, and search for the one with the maximum value, which is regarded as the new candidate.

Without further simplifications, the computational cost of the greedy approach still remains high since the cost of evaluating φ and ξ is not small and N^* is very big. To overcome this, we propose a simpler one which ignores basis function:

$$\Delta_2(iter) = \sum_{i=1}^{N^e} || \sum_{i=1}^{N^e} \gamma_i^{km} (\phi(\frac{|\theta - \theta_i|}{\sigma}) - \phi(\frac{|\theta_l - \theta_i|}{\sigma}))||_2 \tag{26}$$

Strictly speaking, Δ_1 and Δ_2 doesn't belong to error estimators but could be an estimator for us to find points with large error. In concrete application we find that compared to (25), (26) performs very similar but is much less time-consuming. Going forward, we use (26) to drive the greedy selection in θ.

For problems with more than one state variable $u(q)(q = 1, 2, \cdots Q)$, we use (26) to compute Δ_2 for each variable, denoted by $\Delta_2(q, iter)$, and define the error estimator for this problem denoted by Δ_{mul} as

$$\Delta_{mul}(iter) = \sqrt{\sum_{q=1}^{Q} w(q) * \Delta_2(q, n)^2} \tag{27}$$

where $w(q)$ is a weight that represents the importance of each variable. Naturally, often this is subjective and problem dependent.

To terminate the growth of the reduced basis, we can employ different strategies. For example, we can just set the number of iterative steps to a constant, or base it on the decrease of the approximation error below a certain threshold. As we seek to reduce the overall approximation error, our criterion includes the following three steps:

At first, we set an appropriate step number N_{pre} to run the Greedy ROM's iteration. Our experience is 15 for 1-parameter problem and 35 for 2-parameter problems.

Secondly, at every iterative step, we have to make an judgement: once an approximation error between the solver and Greedy ROM is less than a preset value, the Greedy ROM system terminates. As to which kind of error to focus on, we would prefer to use the 2-norm of errors on a set of test points as the error indicator. However, in concrete application, this often results in substantial computational cost. As an alternative, the mean error on the determined new candidates is chosen as the indicator.

Finally, if the iteration terminates because of the indicator, there is no need to add additional candidates. If not, we have to make a decision based on the error. If the error is below a threshold or has remained constant for a few iterations, we terminate the process. Otherwise, the iteration continues.

3.2. Summery of Greedy ROM Approach

As presented above, we propose two alternatives for error indicators to decide sampling, however both are heuristic and it is difficult for us to provide a rigorous bound for them. But for numerical cases in section 4, the ROM performs well with candidates determined by the indicator, demonstrating that those may not be optimal but at least very good samples for building reduced basis.

By adopting the error indicator as well as the criterion for termination, we can express the whole Greedy ROM scheme, summarized in Algorithm 2.

Algorithm 2 Greedy ROM Scheme

1: Choose few snapshots as a start.
2: **for** $iter = 1; iter \leq N_{pre}; iter++$ **do**
3: Apply two-level POD to extract reduced basis ξ^e and φ^e, and adopt RBF to compute γ_i^{km}.
4: Use the error indicator to determine new candidate in the whole parameter space.
5: Add the determined candidate to the existing snapshots.
6: Once the 2-norm of errors on a set of test points is less than a preset value, terminate the Greedy sampling iteration.
7: **end for**
8: Make a judgement whether to add additional steps or not based on the error variation trend.
9: Use the ROM in section 2 to extract basis from the snapshots determined above and approximate solutions we need.

4. Numerical test-case

In the following, we consider two different cases to demonstrate the efficiency and accuracy of the proposed method. It is, however, worth emphasizing that the proposed algorithm is generic and applicable to general problems.

4.1. Two-dimensional heat diffusion

We consider the time-dependent diffusion equation.

$$\frac{\partial u}{\partial t} = \upsilon \triangle u \qquad in \qquad \Omega = [0,1]^2, t \in [0,T] \tag{28}$$

where T=5.0s, and viscosity coefficient $\upsilon = 1$. And the time-step $dt = 0.004$ seconds thus $N_t = 1251$. The problem is solved using a second order finite difference method with a uniform spatial grid of 41×41 points. For the generation of snapshots, we select $N_T = 51$ points uniformly in [0,T] to generate spatial snapshots, and a coarse spatial mesh of 11×11 points is used to generate temporal snapshots.

A. Single parameter case

Let us first consider the case with just one parameter introduced in the boundary condition

$$g^\theta(x,t) = sin(\frac{\pi t}{2T})(1 - \theta sin(2\pi(x - 0.5))) + \frac{\theta t}{T} \tag{29}$$

where $\theta \in [0.5, 1.5]$.

Initial solution variables are set 0 everywhere in $\Omega = [0,1]^2$,

$$u_0^\theta(x) = 0 \tag{30}$$

The threshold ε_1 to recover temporal and spatial modes are both set 10^{-9}. When applying the SVD on these different sets of modes, thresholds ε_2 for temporal modes are set to 10^{-3}, and ε_2 for spatial modes equals to 10^{-4}. And Gaussian radial basis functions with $\sigma = 0.005$ is adopted.

The greedy approach is initiated from two candidates $\theta = 0.5, 1.5$ and at each iteration it provides a new one. In Table 1 we list the samples determined in the iterative approach, using the standard RBF approximation as well as the RBF-QR technique and we note that the results are different, in particular as the iteration count increases. This is explained by the increasing dominance of the ill-conditioning of the standard RBF approximation as the number of samples increases.

Table 1: θ of chosen candidates at each iteration

Iteration	1	2	3	4	5	6	7	8	9	10
RBF_OLD	1	1.25	0.75	1.44	1.19	0.94	0.7	0.55	1.47	1.41
RBF_QR	0.89	1.19	0.69	1.34	1.04	0.79	0.6	0.96	1.11	1.26
Iteration	11	12	13	14	15					
RBF_OLD	1.49	1.43	1.39	0.36	1.34					
RBF_QR	1.41	1.47	0.63	1.49	0.61					

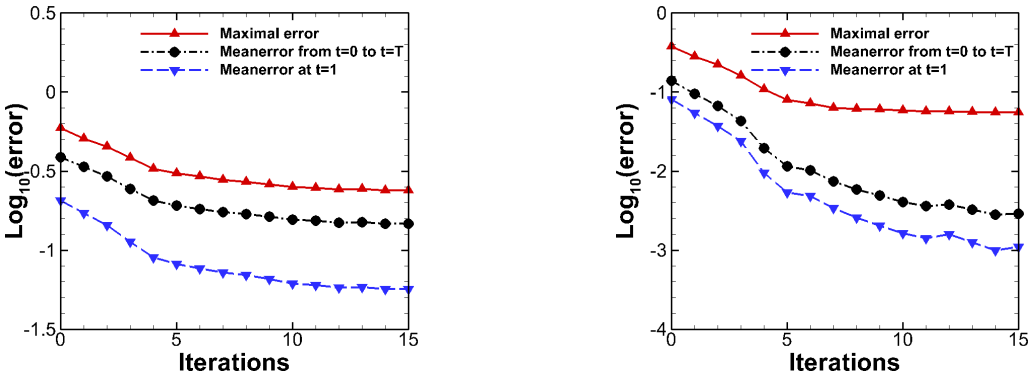
To evaluate the accuracy, we choose $N^* = 51$ points θ^* uniformly in the parameter space as test samples and measure the maximum error, the mean error over time and the mean error at $t = 1$s, defined as

$$\text{Maximal error}: \quad Error1 = \max_{t_n \in [0,T]} ||\hat{u}^\theta(x, t_n) - u^\theta(x, t_n)|| \tag{31}$$

$$\text{Mean errors at } T \geq t \geq 0: \quad Error2 = \frac{1}{N_t} \sum_{n=1}^{N_t} ||\hat{u}^\theta(\cdot, t_n) - u^\theta(\cdot, t_n)|| \tag{32}$$

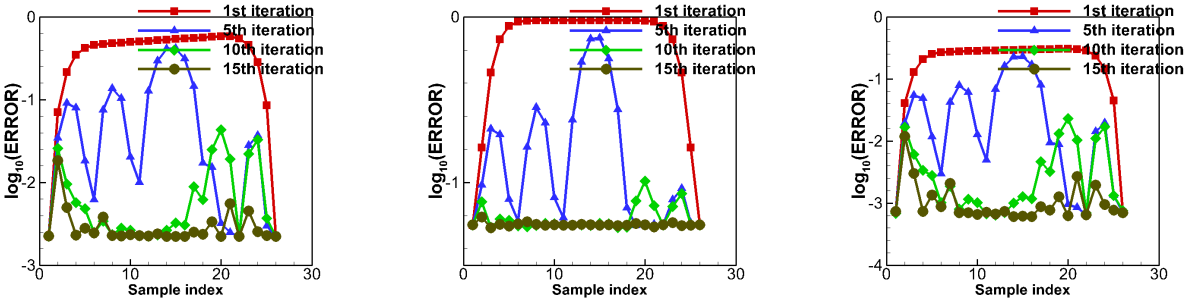
$$\text{Mean error at t=1s}: \quad Error3 = ||\hat{u}^\theta(\cdot, t = 1) - u^\theta(\cdot, t = 1)|| \tag{33}$$

Figure 1 shows the trend of the errors defined above for the test samples in greedy system scheme using both the standard RBF approximation and the improved RBF-QR. The results clearly confirm the importance of using improved technique to ensure the accuracy. Going forward, we will only consider solutions with RBF-QR. In Figure 2, the errors for each test sample all drop when iterative step n grows, demonstrating a global error reduction achieved.



(a) standard RBF (b) RBF-QR

Figure 1: Errors between solution approximated by reduced model and solutions calculated directly by the solver. In (a) we use the standard RBF in the reconstruction while in (b) the RBF-QR is used.



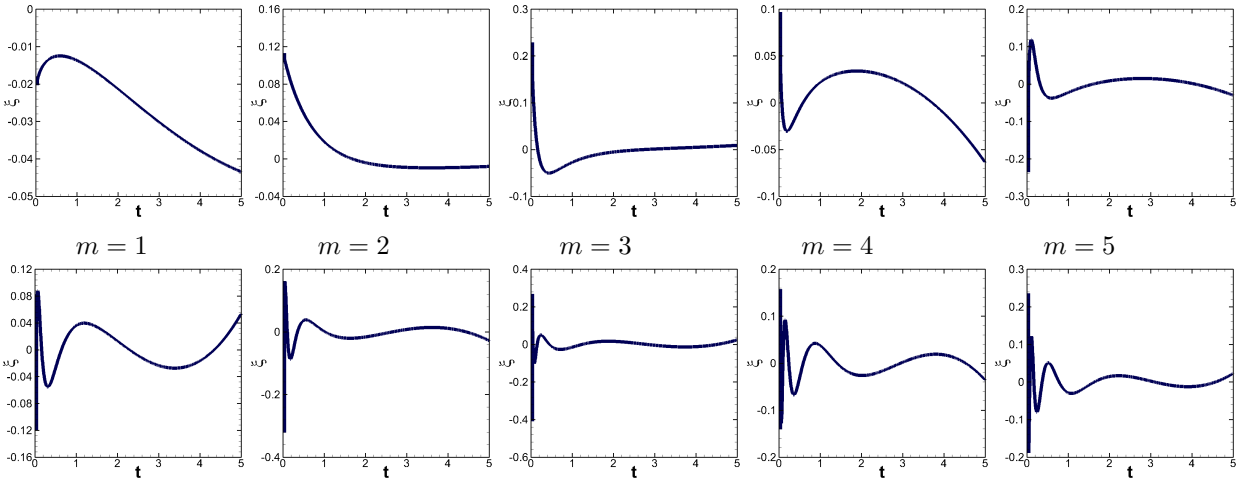
(a) Maximal error (b) Mean error at $T \geq t \geq 0$ (c) Mean error at t=1s

Figure 2: Errors(Eq. (31-33)) in log10 scale for test points θ^*

To further evaluate efficiency, approximations of the intrusive reduced model presented in Section 2 with 17 samples uniformly distributed in the parameter space is considered as a comparison. We recall that this is the same number with candidates provided by Greedy sampling. The errors of both approximation are compared in Table (2). From the table, the errors of the Greedy ROM are almost identical with that of the ROM with uniform points, illustrating that in this case Greedy ROM does not present much advantage over the uniformly sampled case. The reason is partly due to the simplicity of the boundary condition and the single parameter in this case.

Table 2: Comparison of errors between Greedy ROM and the ROM with uniform snapshots

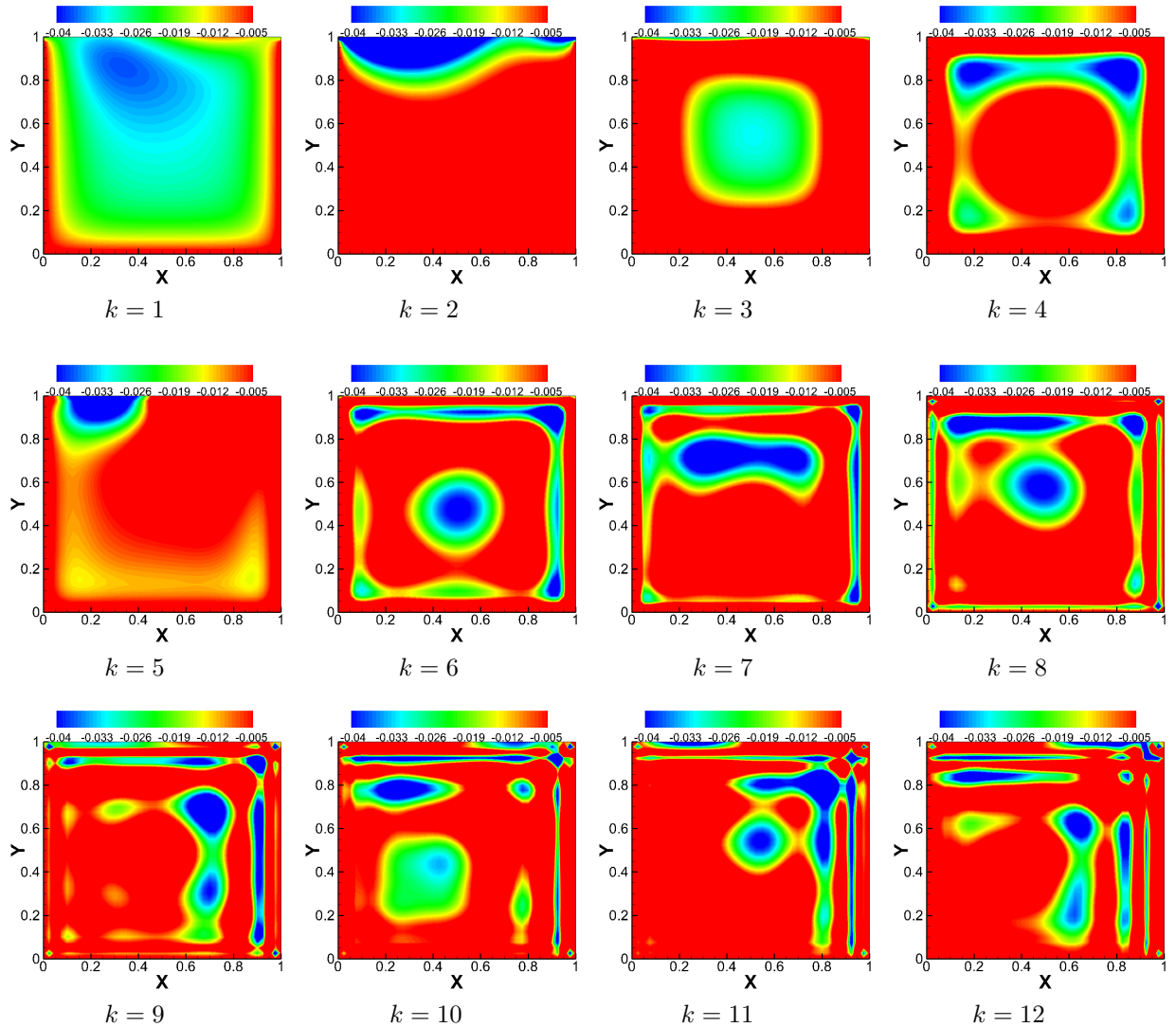
	Maximal error	Mean error at $T \geq t \geq 0$	Mean error at t=1s
Greedy ROM	-1.257	-2.537	-2.96
ROM with uniform snapshots	-1.257	-2.552	-3.021



$m = 1$ $m = 2$ $m = 3$ $m = 4$ $m = 5$

$m = 6$ $m = 7$ $m = 8$ $m = 9$ $m = 10$

Figure 3: The first 10 temporal modes



$k = 1$ $k = 2$ $k = 3$ $k = 4$

$k = 5$ $k = 6$ $k = 7$ $k = 8$

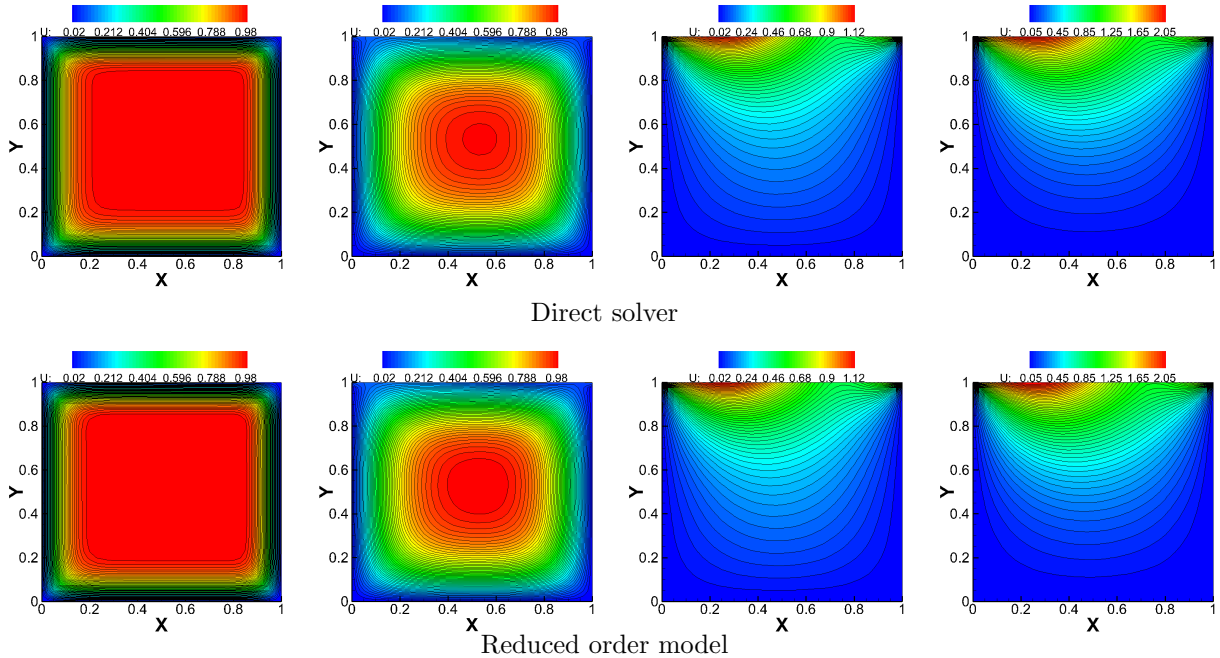
$k = 9$ $k = 10$ $k = 11$ $k = 12$

Figure 4: The first 12 spatial modes

With those 17 candidates, the Greedy ROM generates a total of $\sum_{i=1}^{I} M^i = 284$ temporal basis and $\sum_{i=1}^{I} K^i = 216$ spatial basis. After the SVD compression 12 temporal basis $\xi^m (m = 1, 2, \cdots M)$ - see Figure 3 for an illustration of these, and 10 spatial ones $\varphi^k (k = 1, 2, \cdots K)$, illustrated in Figure 4, are retained. We observe that all temporal modes in Figure 3 change rapidly close to 0, consistent with the rapid changes of the solution during the initial temporal phase.

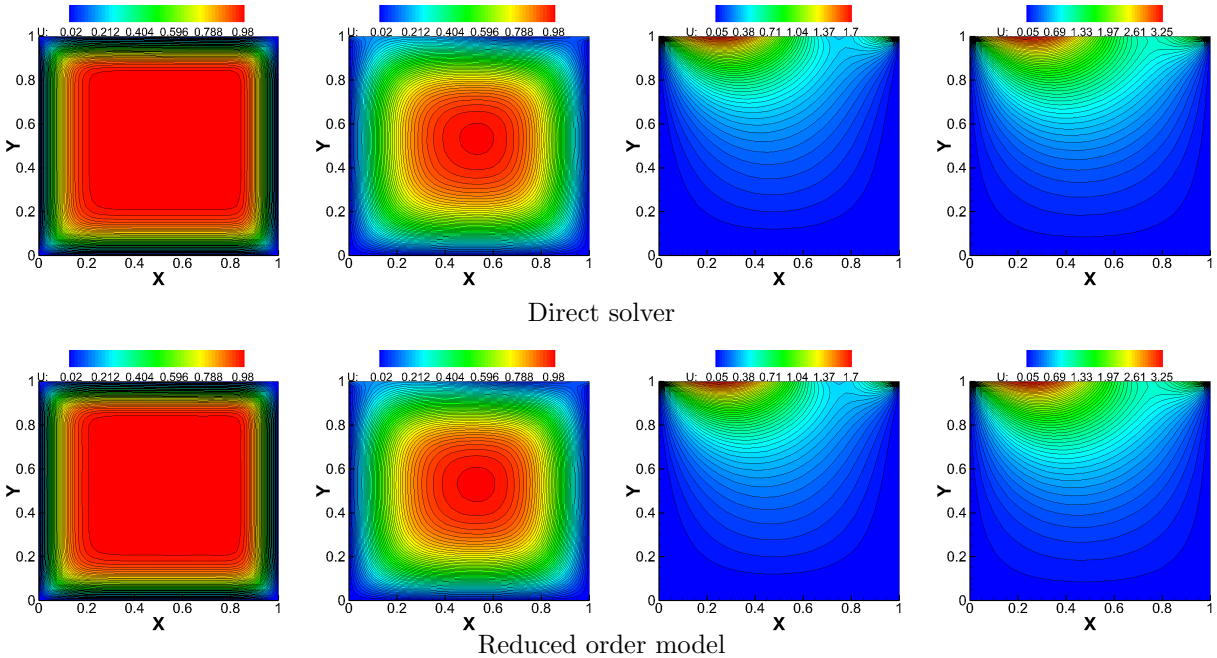
After the reduced basis built, it takes 0.12 seconds to approximate one point with those reduced basis while the numerical solver have to spend 251.97 seconds, illustrating that a great online benefit is achieved in computational cost with the proposed approach.

We observe in Figure 3 that the errors become very small after greedy sampling. However, this does not allow us to conclude that the Greedy ROM approximation is accurate. To evaluate this, we need to consider the direct error in flow field's prediction across variations in θ. We consider the flow field at two points, neither of which belong to the candidates. One is $\theta = 0.55$ and the other is $\theta = 1.15$. Figure 5 and 6 display the solution contours of both points at different time instants, obtained by the reduced model as well as the direct solver and Figure 7 gives the error distributions $|u^\theta(x,t) - \hat{u}^\theta(x,t)|$. We observe good agreements of fluid fields at both points and small errors distributed in the whole field.



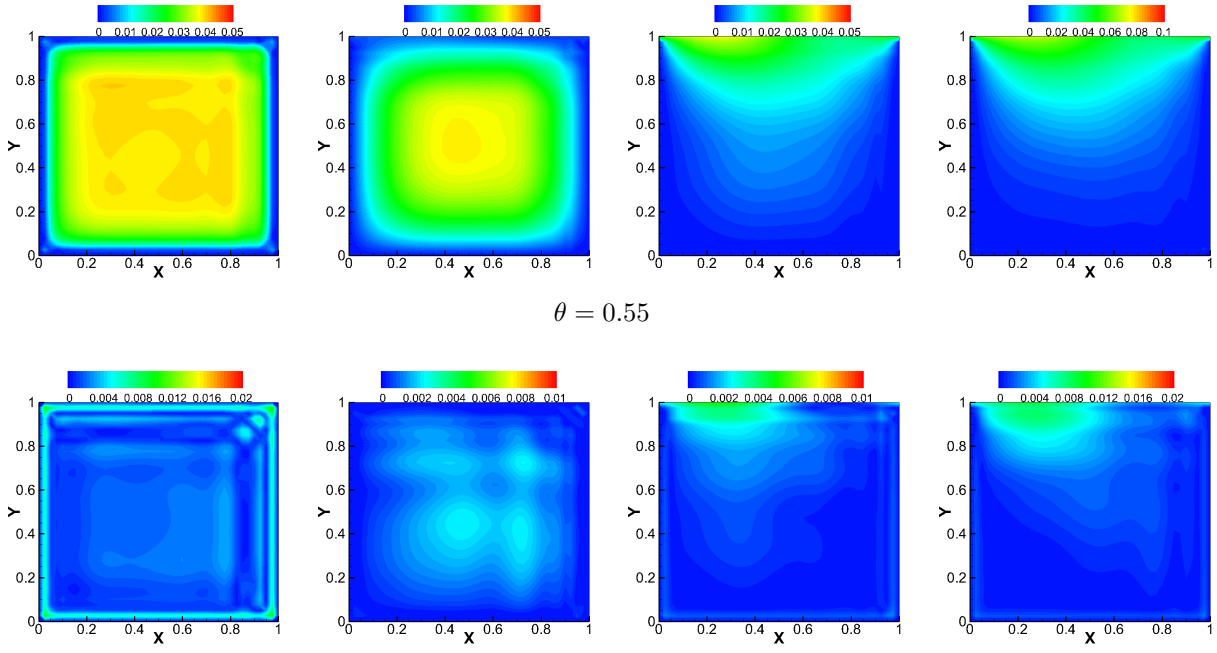
(a)t=0.04s (b)t=0.2s (c)t=2s (d)t=5s

Figure 5: Contour of solution at different times for $\theta = 0.55$



(a)t=0.04s (b)t=0.2s (c)t=2s (d)t=5s

Figure 6: Contour of solution at different times for $\theta = 1.15$



(a)t=0.04s (b)t=0.2s (c)t=2s (d)t=5s

Figure 7: $\Delta u(x,t) = |u^{\theta}(x,t) - \hat{u}^{\theta}(x,t)|$ for $\theta = 0.55, 1.15$

B. Dual parameters' case

For the single parameter problem discussed above, the Greedy ROM has no obvious advantage over direct ROM(ROM with uniform sampling). Let us turn to a two-parameter problem and consider the two-dimensional diffusion problem with two parameters θ_1, θ_2, associated with the initial condition and the boundary condition, respectively.

The initial condition is given by

$$u_0^{\theta}(x) = \theta_1 \times c + \bar{u} \tag{34}$$

where c is a function of x and y, $c(x, y) = \lambda(x + 0.5)(x - 0.5)(y + 0.5)(y - 0.5)$ and $\bar{u}$ is the solution of the Poisson problem

$$\begin{cases} -\triangle \bar{u} = 1 & in \quad \Omega \\ \bar{u} = 0 & on \quad \partial\Omega \end{cases} \tag{35}$$

The boundary condition is time-independent but depends on θ_2 as

$$g^{\theta}(x, t) = \theta_2 \qquad x \in \partial\Omega \tag{36}$$

We assume a interval $[0.5, 1.5]$ for θ_1, θ_2 and set the constant parameters in the Greedy ROM system as: $\varepsilon_1 = 10^{-9}$; ε_2 for temporal and spatial basis equals to 0.005 and 10^{-3} respectively; $\sigma = 0.25$.

The initial candidates consist of four corners in the parameter space, {(0.5,0.5), (0.5,1.5), (1.5,0.5), (1.5,1.5)}. After 35 iterations, we obtain a candidate ensemble consisting of a total of 39 parameter points (together with initial points), see round spots in Figure 8. Figure 9 shows the trend of the three errors with increasing sampling, measured on a 15×15 uniform grid. After 35 iterations, the mean error at $T \geq t \geq 0$ as well as mean error at $t = 1$s is below than 10^{-3}, which ensures a good accuracy of the Greedy ROM in this case.

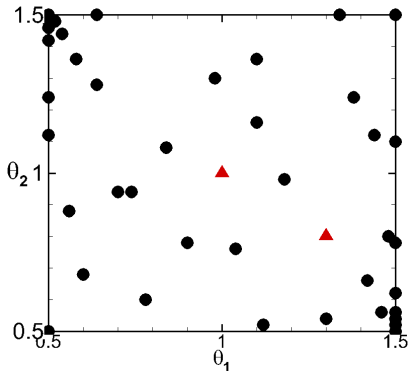


Figure 8: Final parameter points after 35 iterations

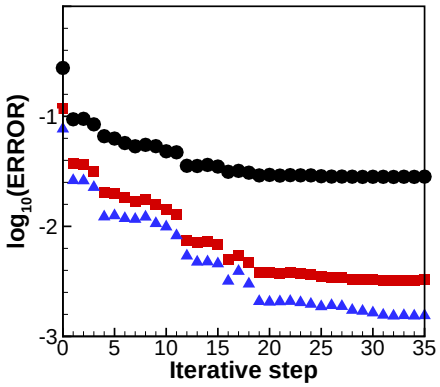
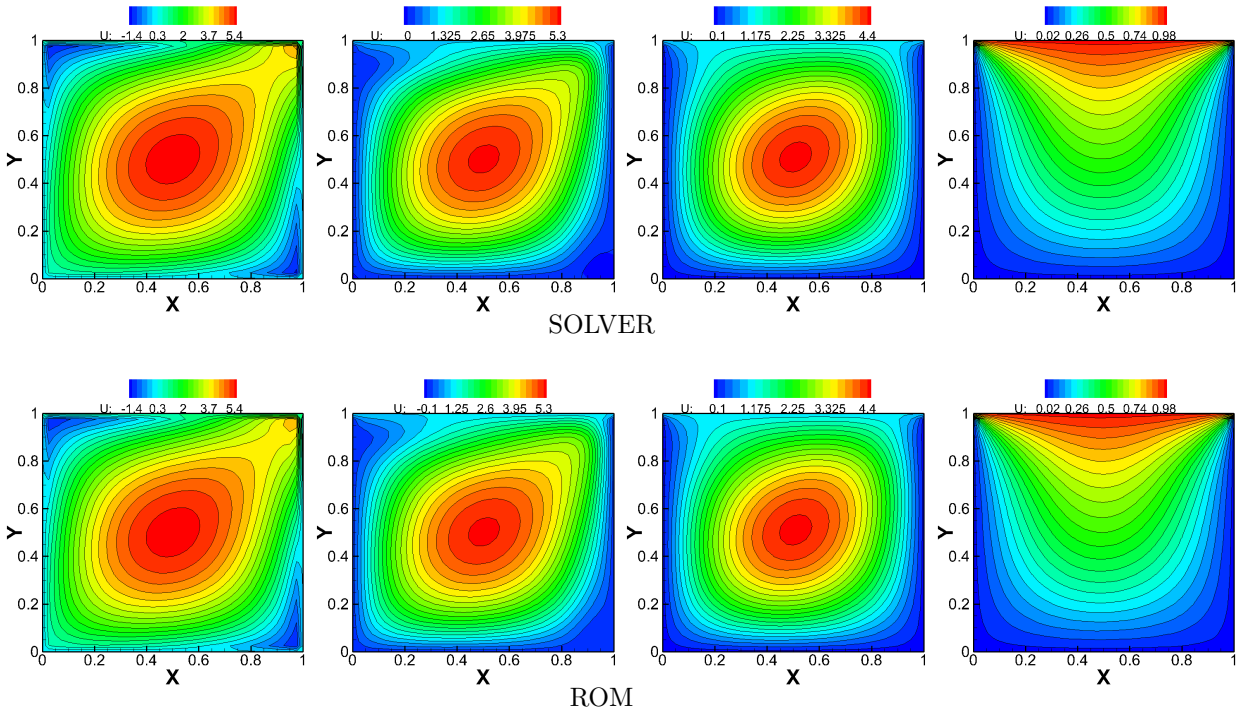


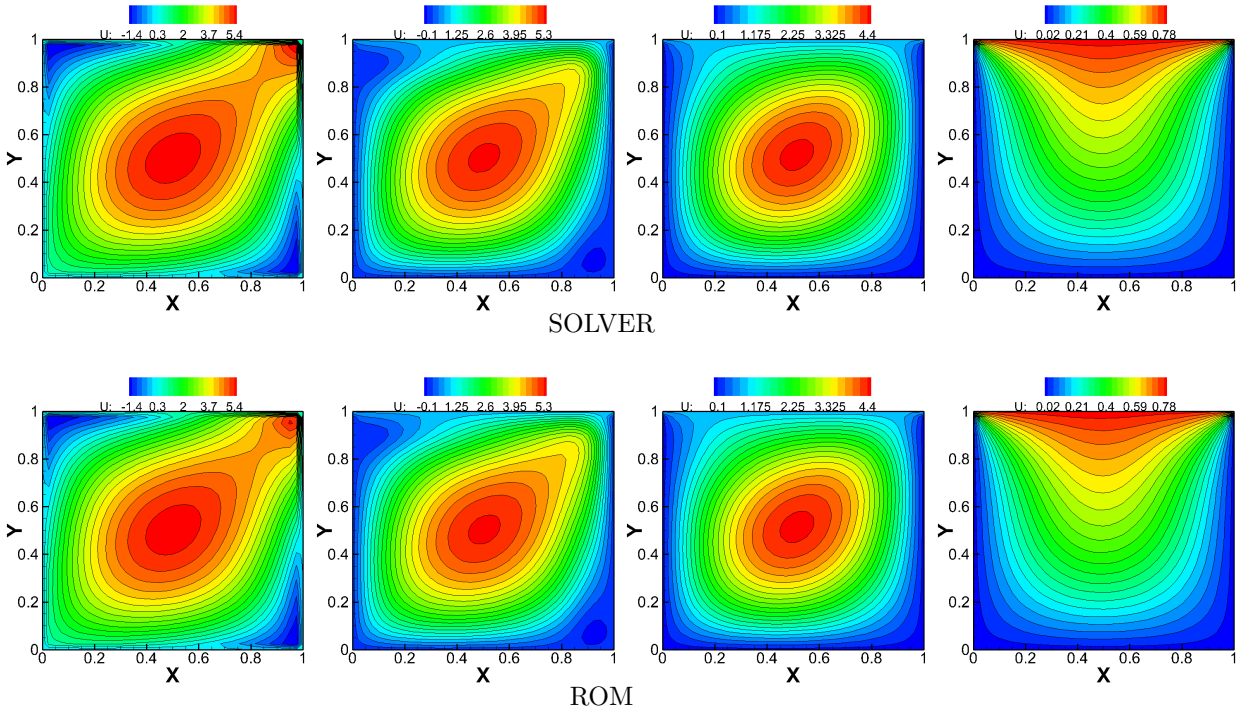
Figure 9: Errors between solution approximated by the reduced model and solutions calculated directly by the solver across 15×15 test points

We again pick two parameter points (1,1) and (1.3,0.8), marked by delta spots in Figure 8, neither of which belong to the candidates or close to the candidates, to gauge the accuracy of the flow field solution. u contours are presented as a comparison, see Figure 10 and 11. As expected, the reduced model provides a very accurate approximation of the solution at any time instant in the interval $[0, T]$.



(a)t=0 (b)t=0.1s (c)t=0.5s (d)t=5s

Figure 10: u contour at different time instants for $\theta_1 = 1$ and $\theta_2 = 1$



(a)t=0 (b)t=0.1s (c)t=0.5s (d)t=5s

Figure 11: u contour at different time instants for $\theta_1 = 1.3$ and $\theta_2 = 0.8$

4.2. Lid-driven cavity flow

As a more complex problem in fluid dynamics, a two-dimensional driven cavity flow with two parameters θ_1 and θ_2 is considered as the second verification case. For this, we consider the incompressible Navier-Stokes equations(excluding body forces):

$$\begin{cases} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \\ \frac{\partial u}{\partial t} + u \cdot \frac{\partial u}{\partial x} + v \cdot \frac{\partial u}{\partial y} = -\frac{1}{\rho_0} \cdot \frac{\partial p}{\partial x} + \nu \Delta u \\ \frac{\partial v}{\partial t} + u \cdot \frac{\partial v}{\partial x} + v \cdot \frac{\partial v}{\partial y} = -\frac{1}{\rho_0} \cdot \frac{\partial p}{\partial y} + \nu \Delta v \end{cases} \tag{37}$$

where ρ_0 is a constant. Unlike the diffusion problems, there are now three state variables: the velocity in x direction(u), the velocity in y direction(v), and the pressure (p). And the Reynolds number equals to 100, thus the viscosity ν is determined.

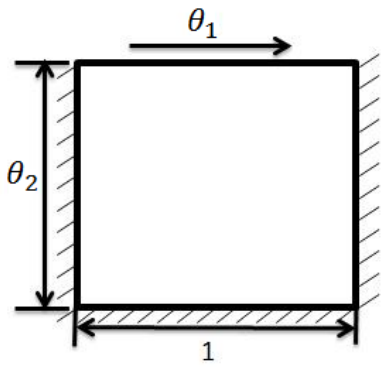


Figure 12: Schematic diagram of computational domain

The computational domain is shown in Figure 12. θ_1 denotes the velocity of the top boundary, and θ_2 refers to the height of the cavity such that the computational domain is $[0,1] \times [0,\theta_2]$. The initial velocities are 0 everywhere at the boundary except at the top boundary. In numerical simulation, T=5.0s and time-step $dt = 0.01s$. We use a second order finite difference scheme with a uniform spatial grid made of 41×41 points. In addition, a coarser spatial mesh of 11×11 points is generated to extract temporal snapshots and a uniform $N_t = 51$ point set in [0,T] is used to generate spatial snapshots.

We assume that $\theta_1, \theta_2 \in [1,2]$. It is worth noting that as the flow field characteristics and accuracy requirements of the different variables are different, different constants should be given in our approximation system for each state variable: ε_1 to compute temporal and spatial modes are set $\{10^{-5}, 10^{-5}, 10^{-9}\}$ for u, v, p; ε_2 for the temporal modes are given by$\{0.03, 0.05, 10^{-3}\}$, while ε_2 for the spatial modes are set $\{10^{-4}, 10^{-4}, 10^{-5}\}$ respectively; and in RBF we use $\sigma = \{0.02, 0.02, 1\}$ for the 3 variables.

A. Adaptive sampling

The sampling is initiated from four points $(\theta_1, \theta_2) = \{(1,1),(1,2),(2,1),(2,2)\}$. After 35 iterations the candidates are located at round spots in Figure 13.

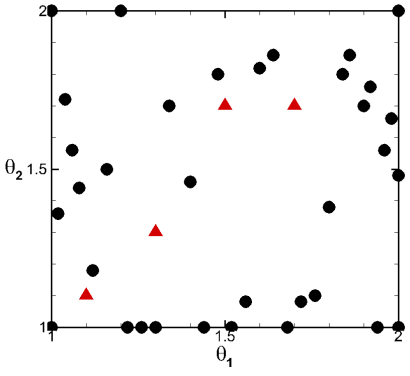
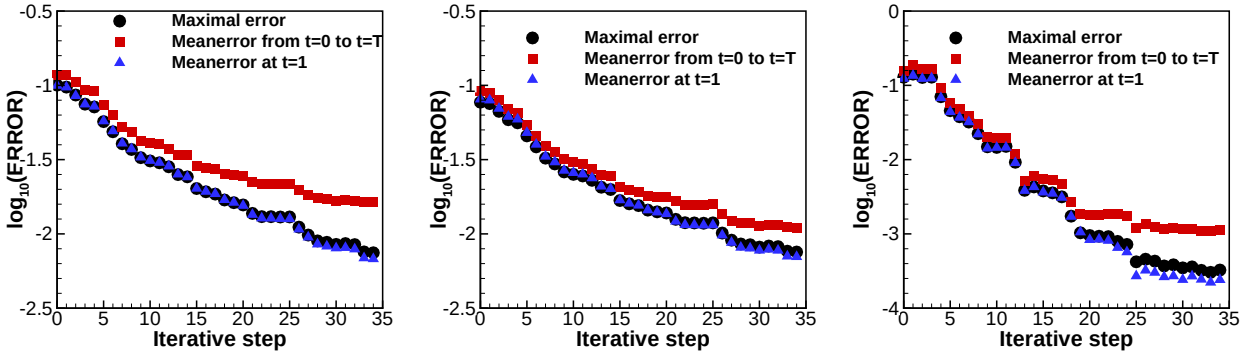


Figure 13: Final parameter points after 35 iterations

We select 11×11 points uniformly distributed in the parameter space $[1,2] \times [1,2]$ as test points.

Figure 14 shows the trend of the three state variables' errors defined in (31)-(33) for the test points between the reduced order solution and by the numerical solver. With the iteration increasing, errors of the three state variables all gradually decrease and the variation turn flat after 35 iterations, after which the errors are less than 10^{-2} for the velocity components, while p errors are less than 10^{-3}.



(a) u (b) v (c) p

Figure 14: Errors the reduced order model and the full solutions at uniformly distributed test points

Table 3: Errors between the Greedy ROM and uniform sampling ROM

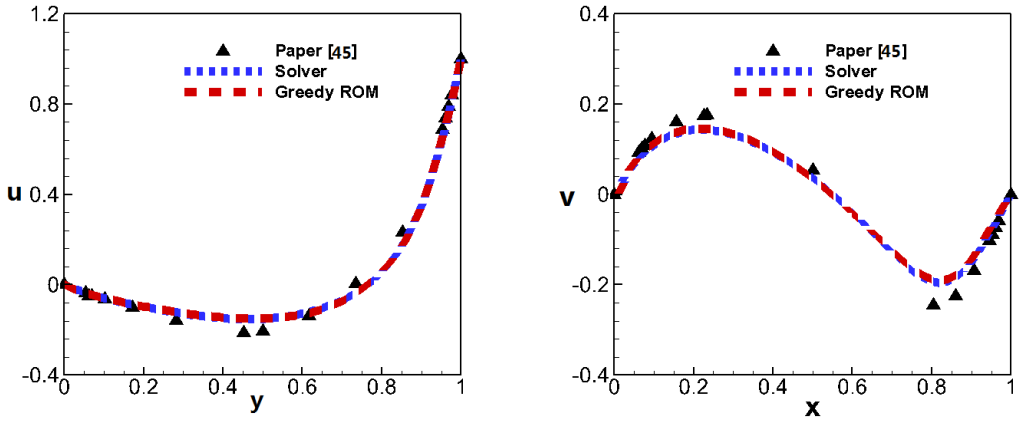
	Maximal error			Mean error at $T \geq t \geq 0$			Mean error at t=1s		
	u	v	p	u	v	p	u	v	p
Greedy ROM	-1.848	-2.068	-2.975	-2.310	-2.2289	-3.557	-2.359	-2.258	-3.671
ROM with uniform candidates	-1.246	-1.370	-2.251	-1.264	-1.434	-2.521	-1.241	-1.394	-2.586

Considering the ROM in Section 3 with 37 candidates spread uniformly in the parameter space we compare in Table (3) the errors defined in (31)-(33) at test points for the two different approaches. For the same variable, the errors of Greedy ROM system are substantially smaller than those computed by the ROM with 37 candidates, showing that in this case the candidates determined by Greedy ROM system has a notable advantage over the direct uniform sampling approach.

B. Fluid field approximation

After the sampling process, the scheme provides 11, 10 and 8 temporal modes as well as 9, 11 and 7 spatial modes for u, v and p, respectively. For each test sample, the reduced basis only cope with $11 \times 9 + 10 \times 11 + 8 \times 7 = 265$ unknowns. While in this test-case with 3 state variables, the full model deals with a mesh of 1681 cells and 501 time intervals, meaning that it has to solve 2.52 million unknowns for each sample. Thus the DOFs of problem solving has been greatly reduced. Figure 15 compares velocity profiles of Greedy ROM system with those by the direct incompressible Navier-stokes solver when $\theta_1 = 1$ and $\theta_2 = 1$. As the figures shown, the approximated velocity profiles coincide very well with those of the solver. Solutions in [45] are also presented for reference. The example also highlights that

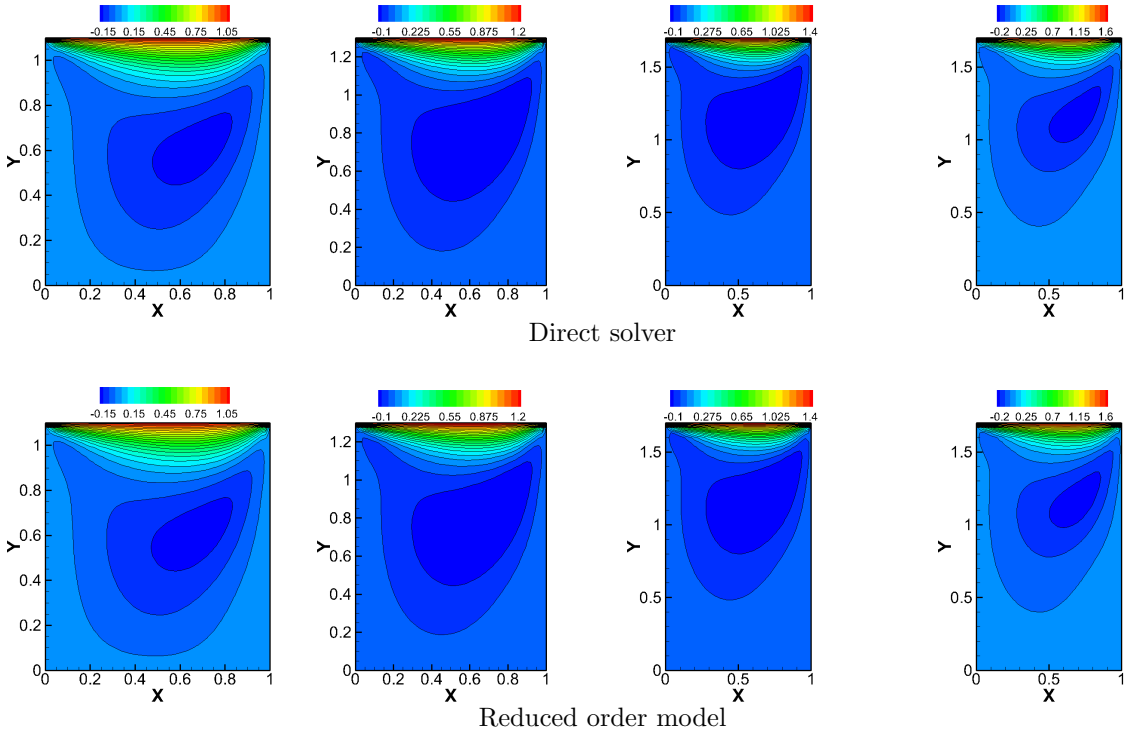
the accuracy of the reduced model is consistent with the numerical solver, i.e., the differences observed in Figure 15 can be attributed to the quality of the solver.



(a) u-velocity distribution along vertical vine through the geometric center of the cavity

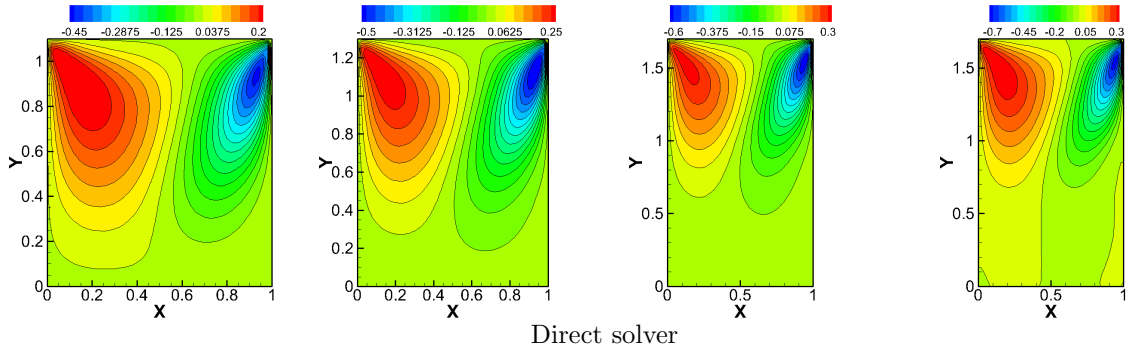
(b) v-velocity distribution along horizontal line through the geometric center of the cavity

Figure 15: Velocity profile

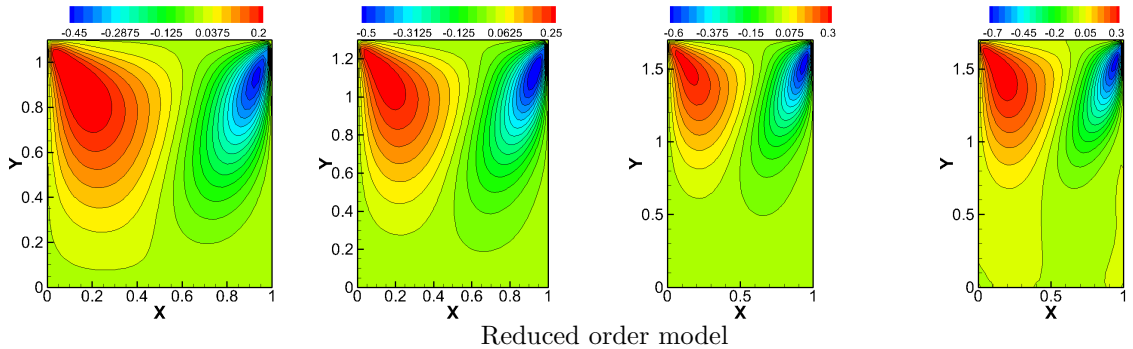


(a)$(1.1, 1.1)$ (b)$(1.3, 1.3)$ (c)$(1.5, 1.7)$ (d)$(1.7, 1.7)$

Figure 16: X-velocity contour at the four points



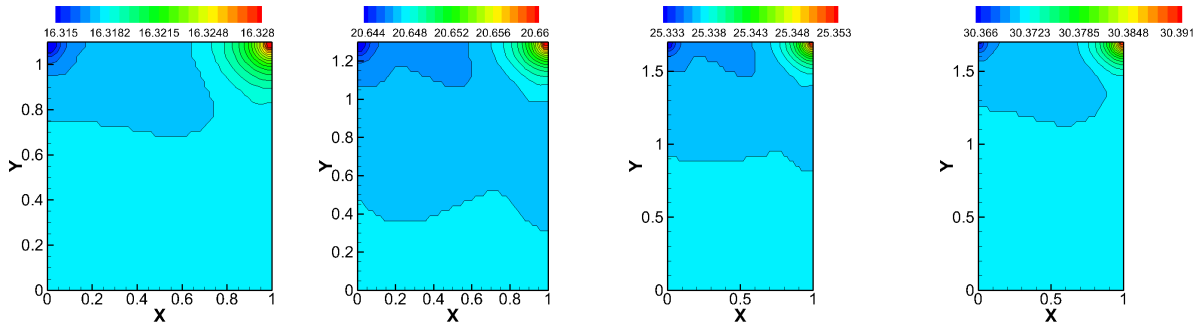
Direct solver



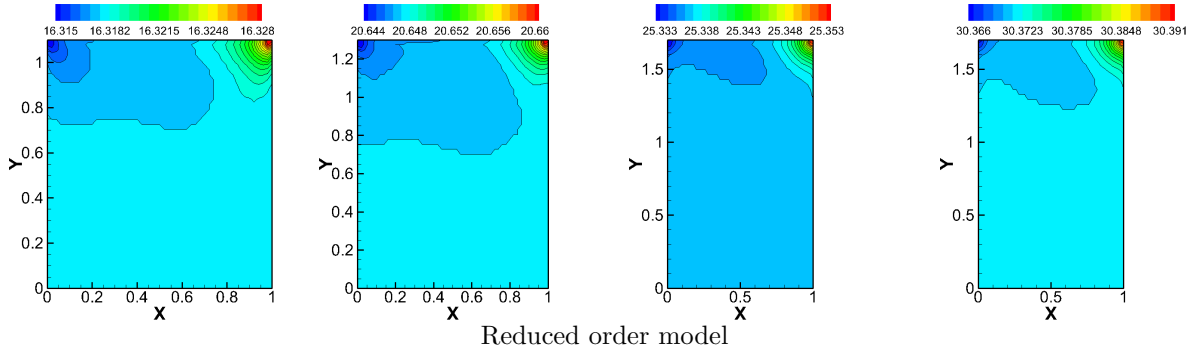
Reduced order model

(a)$(1.1, 1.1)$ (b)$(1.3, 1.3)$ (c)$(1.5, 1.7)$ (d)$(1.7, 1.7)$

Figure 17: Y-velocity contour at the four points



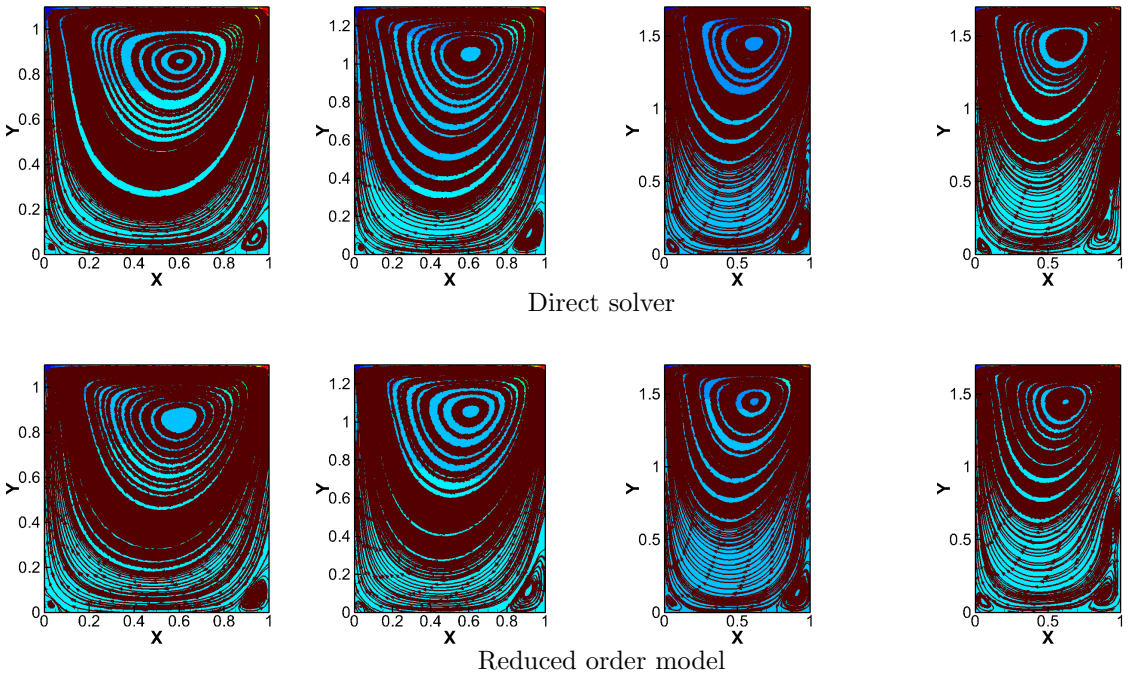
Direct solver



Reduced order model

(a)$(1.1, 1.1)$ (b)$(1.3, 1.3)$ (c)$(1.5, 1.7)$ (d)$(1.7, 1.7)$

Figure 18: Pressure contour at the four points



(a)$(1.1, 1.1)$ (b)$(1.3, 1.3)$ (c)$(1.5, 1.7)$ (d)$(1.7, 1.7)$

Figure 19: Streamlines at the four points

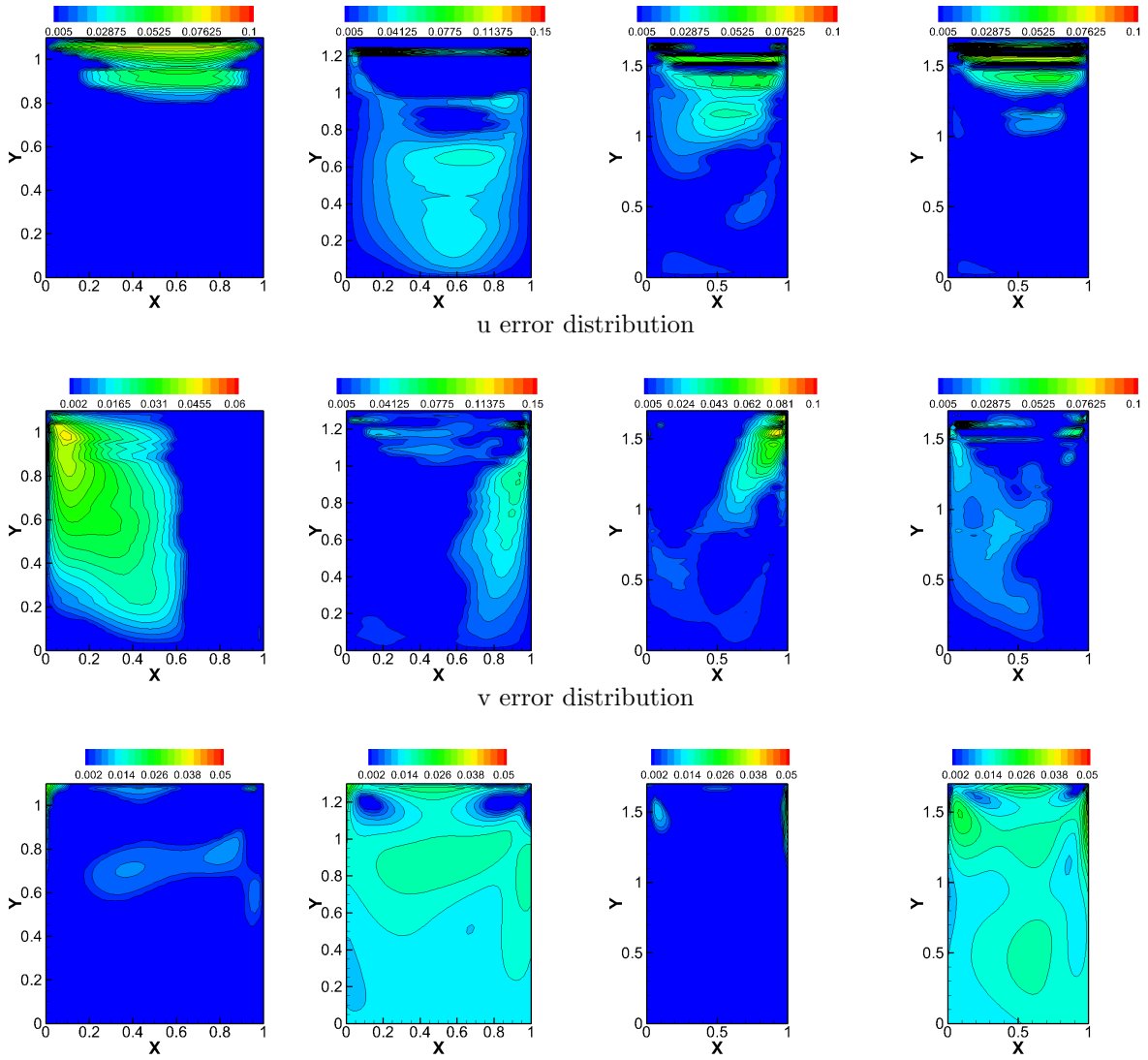


Figure 20: Error distributions at the four points

To further confirm the accuracy of the Greedy ROM approach, flow fields, contours and streamlines are given for four points marked by delta spots in Figure 13 . From the contours of the three state variables in Figure 16-18, we observe an excellent agreement between the contours approximated by Greedy ROM and computed by the direct solver. Error distributions in Figure 20 also demonstrate the approximation accuracy of our proposed scheme. Although streamlines are extremely sensitive and minor variation in velocity contours may lead to difference in the streamlines, we also observe a good agreement in streamlines, see Figure 19. Even when the height of cavity varies a lot, the approach is able to simulate streamlines well, including capturing the corner vortices accurately.

C. Time-consuming evaluation

Computational efficiency is a key factor to judge a ROM and in this section computational cost of the proposed greedy ROM is presented to evaluate its efficiency, which is divided into two parts:

In the sampling part, with n grows, the CPU timing also increases, meaning that time-consuming at the last step is the largest, thus only time-consuming at the last cycle is presented, as Table 4. At this moment, it takes 28.81 seconds to extract reduced basis and 2.7355 seconds to solve γ_i^{km} with 39 snapshots, and the process of determining the new candidate from $51\times51 = 2601$ points with the estimator only cost 3.3041 seconds. Finally the numerical solver is called up to generate the new snapshot, which costs 73.2s. The total time-consuming at 35th step is $28.81 + 2.7355 + 3.3041 + 73.2 = 108.05$ seconds.

The second part is adopting ROM to approximate solutions with existing snapshots determined by sampling approach. From Table 5, in the offline stage it takes $28.81 + 2.7355$ seconds and in the online stage, it takes 0.0468 s for the ROM approximating one test sample with the snapshots determined by sampling approach, while our numerical solver cost 73.2s, demonstrating that the CPU timing's speed-up is very significant and inspiring.

Combining with the high approximation accuracy presented, we conclude that our proposed approach performs excellently and efficiently in this case.

Table 4: Time-consuming (in seconds) required by Greedy sampling at 35th cycle.

Phase	Time-consuming
Sampling:Extracting reduced basis	28.81
Sampling:Computing $\gamma_i^{km}(i = 1, 2, \cdots N)$	2.7355
Sampling:Determining the new candidate	3.3041
Numerical Solver	73.2
Total	108.05

Table 5: Time-consuming (in seconds) required by the numerical solver and by the proposed scheme(offline and online computations).

Phase	Time-consuming
ROM: extracting reduced basis (offline)	28.81
ROM: computing $\gamma_i^{km}(i = 1, 2, \cdots N)$(offline)	2.7355
ROM: approximation for one sample (online)	0.0468
Numerical Solver(for one sample)	73.2

5. Concluding remarks

We have presented the development of a non-intrusive reduced order model based on a greedy approach. While the focus has been on problems in fluid dynamics, the approach is applicable to general parametrized time-dependent problems. To enable a robust approximation for many snapshots, we introduced the use of RBF-QR to overcome ill-conditioning. The efficacy and accuracy of the method has been demonstrated by two case applications, including heat conduction and a driven cavity flow.

While the results are encouraging and the approach have been demonstrated on non-trivial examples, future work should seek to improve the effectiveness of the Greedy ROM algorithm, possibly through the use of a local reduced basis, and apply this approach to problems of a more complex nature, including examples where dimension of parameter space is large and boundary conditions vary with time.

References

[1] Jan S Hesthaven, Gianluigi Rozza, and Benjamin Stamm. Certified reduced basis methods for parametrized partial differential equations. *SpringerBriefs in Mathematics. Springer*, 2016.

[2] Alfio Quarteroni and Gianluigi Rozza. *Reduced order methods for modeling and computational reduction*, volume 9. Springer, 2014.

[3] David J Lucia, Philip S Beran, and Walter A Silva. Reduced-order modeling: new approaches for computational physics. *Progress in Aerospace Sciences*, 40(1):51–117, 2004.

[4] M Frangos, Y Marzouk, K Willcox, and B van Bloemen Waanders. Surrogate and reduced-order modeling: A comparison of approaches for large-scale statistical inverse problems. *Large-Scale Inverse Problems and Quantification of Uncertainty*, 123149, 2010.

[5] Tan Bui-Thanh, Karen Willcox, and Omar Ghattas. Model reduction for large-scale systems with high-dimensional parametric input space. *SIAM Journal on Scientific Computing*, 30(6):3270–3288, 2008.

[6] Kazufumi Ito and SS Ravindran. A reduced-order method for simulation and control of fluid flows. *Journal of computational physics*, 143(2):403–425, 1998.

[7] René Pinnau. Model reduction via proper orthogonal decomposition. In *Model Order Reduction: Theory, Research Aspects and Applications*, pages 95–109. Springer, 2008.

[8] Stefan Volkwein. Model reduction using proper orthogonal decomposition. *Lecture Notes, Institute of Mathematics and Scientific Computing, University of Graz. see http://www. uni-graz. at/imawww/volkwein/POD. pdf*, 2011.

[9] Kenneth C Hall, Jeffrey P Thomas, and William S Clark. Computation of unsteady nonlinear flows in cascades using a harmonic balance technique. *AIAA journal*, 40(5):879–886, 2002.

[10] Walter A Silva. *Discrete-time linear and nonlinear aerodynamic impulse responses for efficient CFD analyses*. 1997.

[11] Walter A Silva. Reduced-order models based on linear and nonlinear aerodynamic impulse responses. In *NASA CONFERENCE PUBLICATION*, pages 369–380. NASA, 1999.

[12] María-Luisa Rapún and José M Vega. Reduced order models based on local pod plus galerkin projection. *Journal of Computational Physics*, 229(8):3046–3063, 2010.

[13] RG Duran. Galerkin approximations and finite element methods. *Lecture notes (available at the authors website)*, 1996.

[14] Fabien Casenave, Alexandre Ern, and Tony Lelièvre. A nonintrusive reduced basis method applied to aeroacoustic simulations. *Advances in Computational Mathematics*, 41(5):961–986, 2015.

[15] D Xiao, F Fang, AG Buchan, CC Pain, IM Navon, and A Muggeridge. Non-intrusive reduced order modelling of the navier–stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 293:522–541, 2015.

[16] Maxime Barrault, Yvon Maday, Ngoc Cuong Nguyen, and Anthony T Patera. An empirical interpolationmethod: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique*, 339(9):667–672, 2004.

[17] Francesco Ballarin, Elena Faggiano, Sonia Ippolito, Andrea Manzoni, Alfio Quarteroni, Gianluigi Rozza, and Roberto Scrofani. Fast simulations of patient-specific haemodynamics of coronary artery bypass grafts based on a pod–galerkin method and a vascular shape parametrization. *Journal of Computational Physics*, 315:609–628, 2016.

[18] Saifon Chaturantabut and Danny C Sorensen. Discrete empirical interpolation for nonlinear model reduction. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 4316–4321. IEEE, 2009.

[19] NC Nguyen and J Peraire. An efficient reduced-order modeling approach for non-linear parametrized partial differential equations. *International Journal for Numerical Methods in Engineering*, 76(1):27–55, 2008.

[20] David J Knezevic, Ngoc-Cuong Nguyen, and Anthony T Patera. Reduced basis approximation and a posteriori error estimation for the parametrized unsteady boussinesq equations. *Mathematical Models and Methods in Applied Sciences*, 21(07):1415–1442, 2011.

[21] Christophe Audouze, Florian De Vuyst, and Prasanth B Nair. Nonintrusive reduced-order modeling of parametrized time-dependent partial differential equations. *Numerical Methods for Partial Differential Equations*, 29(5):1587–1628, 2013.

[22] IM Sobol'. On the systematic search in a hypercube. *SIAM Journal on Numerical Analysis*, 16(5):790–793, 1979.

[23] Thomas J Santner, Brian J Williams, and William I Notz. *The design and analysis of computer experiments*. Springer Science & Business Media, 2013.

[24] Qiang Du, Vance Faber, and Max Gunzburger. Centroidal voronoi tessellations: Applications and algorithms. *SIAM review*, 41(4):637–676, 1999.

[25] Vicente J Romero, John V Burkardt, Max D Gunzburger, and Janet S Peterson. Initial evaluation of pure and latinized centroidal voronoi tessellation for non-uniform statistical sampling. 2005.

[26] Fabio Nobile, Raúl Tempone, and Clayton G Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.

[27] Michael Sasena, Matthew Parkinson, Pierre Goovaerts, Panos Papalambros, and Matthew Reed. Adaptive experimental design applied to ergonomics testing procedure. In *ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages 529–537. American Society of Mechanical Engineers, 2002.

[28] Mohammad Ahmadpoor, Bahram Notghi, and John C Brigham. A generalized iterative approach to improve reduced-order model accuracy for inverse problem applications. *Journal of Engineering Mechanics*, 142(5):04016020, 2016.

[29] Bernard Haasdonk and Mario Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 42(2):277–302, 2008.

[30] Bernard Haasdonk. Convergence rates of the pod–greedy method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(3):859–873, 2013.

[31] Annalisa Buffa, Yvon Maday, Anthony T Patera, Christophe Prudhomme, and Gabriel Turinici. A priori convergence of the greedy algorithm for the parametrized reduced basis method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(3):595–603, 2012.

[32] Peter Binev, Albert Cohen, Wolfgang Dahmen, Ronald DeVore, Guergana Petrova, and Przemyslaw Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. *SIAM journal on mathematical analysis*, 43(3):1457–1472, 2011.

[33] K Veroy and AT Patera. Certified real-time solution of the parametrized steady incompressible navier–stokes equations: rigorous reduced-basis a posteriori error bounds. *International Journal for Numerical Methods in Fluids*, 47(8-9):773–788, 2005.

[34] Karen Veroy, Christophe Prud'Homme, Dimitrios Rovas, and Anthony Patera. A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations. In *16th AIAA Computational Fluid Dynamics Conference*, page 3847, 2003.

[35] Gianluigi Rozza, Dinh Bao Phuong Huynh, and Anthony T Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15(3):229, 2008.

[36] Martin A Grepl. Certified reduced basis methods for nonaffine linear time-varying and nonlinear parabolic partial differential equations. *Mathematical Models and Methods in Applied Sciences*, 22(03):1150015, 2012.

[37] Tan Bui-Thanh, Karen Willcox, and Omar Ghattas. Parametric reduced-order models for probabilistic analysis of unsteady aerodynamic applications. *AIAA journal*, 46(10):2520–2529, 2008.

[38] Bengt Fornberg, Elisabeth Larsson, and Natasha Flyer. Stable computations with gaussian radial basis functions. *SIAM Journal on Scientific Computing*, 33(2):869–892, 2011.

[39] T Dudok de Wit, A-L Pecquet, J-C Vallet, and R Lima. The biorthogonal decomposition as a tool for investigating fluctuations in plasmas. *Physics of plasmas*, 1(10):3288–3300, 1994.

[40] Kari Karhunen. Zur spektraltheorie stochastischer prozesse. 1946.

[41] John Leask Lumley. The structure of inhomogeneous turbulent flows. *Atmospheric turbulence and radio wave propagation*, pages 166–178, 1967.

[42] Michael C Romanowski and Earl H Dowell. Reduced order euler equations for unsteady aerodynamic flows: numerical techniques. *AIAA paper*, 96, 1996.

[43] Linda Kaufman. Maximum likelihood, least squares, and penalized least squares for pet. *IEEE Transactions on Medical Imaging*, 12(2):200–214, 1993.

[44] Steven J Nowlan. Maximum likelihood competitive learning. In *NIPS*, pages 574–582, 1989.

[45] UKNG Ghia, Kirti N Ghia, and CT Shin. High-re solutions for incompressible flow using the navier-stokes equations and a multigrid method. *Journal of computational physics*, 48(3):387–411, 1982.