SNAPSHOT LOCATION IN PROPER ORTHOGONAL DECOMPOSITION FOR LINEAR AND SEMI-LINEAR PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

A Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Zhiheng Liu

August 2013

SNAPSHOT LOCATION IN PROPER ORTHOGONAL DECOMPOSITION FOR LINEAR AND SEMI-LINEAR PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

Zhiheng Liu

APPROVED:

Prof. Ronald W. Hoppe,
Chairman

Prof. Tsorngwhay Pan

Prof. Ilya Timofeyev

Prof. Danny C. Sorensen

Dean, College of Natural Sciences and Mathematics

SNAPSHOT LOCATION IN PROPER ORTHOGONAL DECOMPOSITION FOR LINEAR AND SEMI-LINEAR PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

An Abstract of a Dissertation
Presented to
the Faculty of the Department of Mathematics
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

By
Zhiheng Liu
August 2013

Abstract

It is well-known that the performance of POD and POD-DEIM methods depends on the selection of the snapshot locations. In this work, we consider the selections of the locations for POD and POD-DEIM snapshots for spatially semi-discretized linear or semi-linear parabolic PDEs. We present an approach that for a fixed number of snapshots the optimal locations may be selected such that the global discretization error is approximately the same in each associated sub-interval. The global discretization error is assessed by a hierarchical-type a posteriori error estimator developed from automatic time-stepping for systems of ODEs. We compare the global discretization error of this snapshot selection on error equilibration for the full order model (**FOM**) with that for the reduced order model (**ROM**) to study its impact. This contribution also shows that the equilibration of the global discretization error for the **FOM** is preserved by its corresponding POD and POD-DEIM based **ROM**. The numerical examples illustrating the performance of this approach are provided.

Contents

Chapter 1

Introduction

Proper orthogonal decomposition (POD) is one of the most popular model order reduction methods that has been widely used in data analysis, pattern recognition, optimal control and inverse problem. The POD technique assumes the availability of the states (snapshots) $y(t_j)$ at times $\{t_j\}_{j=0}^{m}$ of the system:

$$\begin{cases} \dfrac{dy}{dt} = f(t, y(t)) \quad t \in (0, T], \\ y(0) = y_0, \end{cases}$$

where $f(t, y(t))$ is assumed to be linear or semi-linear. The snapshot POD consists in choosing an orthonormal basis such that the mean square error between the snapshots $y(t_j)'s$ and the corresponding $l - th$ partial sum is minimized on average, i.e., the POD-basis is the solution of the following minimization problem:

$$\begin{cases} \min J(\psi_, \cdots, \psi_l) = \displaystyle\sum_{j=1}^{m} ||y_j - \sum_{k=1}^{l} (y_j, \psi_k)\psi_k||^2, \\ \text{subject to: } (\psi_i, \psi_j) = \delta_{ij} \text{ for } 1 \leq i, j \leq l. \end{cases}$$

And it is well known that the solution of the above minimization problem is given by the eigenvectors of the following self-adjoint operator:

$$\mathcal{R}\psi = \sum_{j=0}^{m} < y(t_j), \psi > y(t_j), \tag{1.1}$$

Surprisingly good approximation properties are reported for POD based schemes in several articles, see [20, 21] for example. However, POD scheme is not able to really reduce the computational complexity for non-linear systems but on the contrary, it increases the computational complexity in some sense. As a result, several other methods are developed to overcome the difficulty of reducing the complexity of evaluating the non-linear term. For example, the missing point estimation (MPE) proposed in [22]; the trajectory piece-wise-linear (TPWL) approximation proposed in [23, 24]; and discrete empirical interpolation decomposition (DEIM) proposed in [4]. Among those approaches regarding methods to reduce the complexity of evaluating the non-linear term, DEIM is the most recent result to deal with the non-linear term. DEIM approach employs a small selected set of spatial grid points to avoid evaluation of the expensive L^2 inner products at every time step that are required to evaluate the non-linearities and focuses on approximating each non-linear function so that a certain coefficient matrix can be precomputed and, as a result, the complexity in evaluating the non-linear term is proportional to the small number of selected spatial indices. Hence, the reduced system from the procedure of DEIM considers both a POD basis for the the state variables and a POD basis related to each non-linear term.

On the other hand, one can deduce that both POD and POD-DEIM method require the availability of the states (snapshots) $y(t_j)$ at times $\{t_j\}_{j=0}^{m}$ of the discretized system and for fixed l, a solid upper bound (state error) of J is given by the sum of eigenvalues of $\mathcal{R}$ after $l-th$ index. This gives the rise to the interest of allocating the POD basis according to the time sub-intervals such that the global state error is minimized. [13] presents a method for allocating the optimal locations of the POD basis for linear PDE by adding new snapshots to the existing fixed snapshots, then proceed to solve the optimal locations of newly added snapshots by minimizing the integral form of the global state error:

$$J(y^h, \bar{t}, \phi) = \int_0^T \|y(t,x) - \sum_{i=1}^{l} y^h(t_i, x)\phi_i\| \; dt, \tag{1.2}$$

where $y^h(t,x)$ is the FE solution of the reduced system based on the fixed snapshots plus newly added snapshots. This approach is able to dramatically reduce the cost function $J(y^h, \bar{t}, \phi)$ but is also restricted to the linear systems.

In this thesis, we present the error equilibration in time approach to select the snapshot locations in a different point of view. We select the locations such that the global discretization error is approximately the same in each associated time sub-interval. This approach is developed from the ideal of the automatic time-stepping method (ATSM), which employs the a posteriori error estimation and an adaptive time step such that the change between two successive error estimations is bounded by the given tolerance. Unfortunately, the adaptive time step may become too small comparing to the mesh size when the solution of the system changes rapidly and hence a super large amount of steps are expected to occur. Different from ATSM, error equilibration in time deals with a fixed number of steps and requires to allocate the snapshots such that each error estimator is of the same order of magnitude in the associated sub-interval, i.e., it is the error estimator that is adaptive. This contribution is organized as follows: Chapter 2 is devoted to the Galerkin semi-discretization in space for evolution equations. The POD and POD-DEIM methods are reviewed in Chapter 3. Global discretization error for the full order model (**FOM**) and the reduced order model (**ROM**) are developed and studied in Chapter 4. In Chapter 5, some numerical examples are presented.

Chapter 2

Evolution Equations and their semi-discretization in space

We consider the semi-linear parabolic PDEs that can be written as the abstract evolution equations according to:

$$\begin{cases} y'(t) + \ Ay(t) - f(t, y(t)) = 0, \quad t \in (0, T], & (2.1a) \\ y(0) = \ y_0, & (2.1b) \end{cases}$$

where $y'(t) := dy/dt$ and A is a linear second order elliptic differential operator with $D(A) = H^2(\Omega) \cap H_0^1(\Omega)$, Ω is a bounded polyhedral domain in $\mathbb{R}^N$ $(N = 1, 2)$, with boundary $\Gamma := \partial\Omega$.

We further assume that:

$$T > 0, \quad y_0 \in L^2(\Omega),$$

and

$$f \in C([0, T], L^2(\Omega))$$

satisfies the Lipschitz condition in the second argument, i.e.,

$$\|f(t,y_1) - f(t,y_2)\|_{L^2(\Omega)} \leq L_f \|y_1 - y_2\|_{L^2(\Omega)}, \quad y_1, y_2 \in L^2(\Omega),\ t \in [0,T], \tag{2.2}$$

for some constant $L_f > 0$. In particular, we assume A to be of the form:

$$Ay = -\sum_{i,j=1}^{N} \frac{\partial}{\partial x_j}(a_{ij}\frac{\partial y}{\partial x_i}) - \sum_{i=1}^{N} b_i \frac{\partial y}{\partial x_i} - cy, \tag{2.3}$$

where $b_i \in L^\infty(\Omega),\ 1 \leq i \leq N,\ 0 \leq c(x) \in L^\infty(\Omega)$, and $a_{ij} \in L^\infty(\Omega),\ 1 \leq i,j \leq N$, such that for some $\alpha > 0$:

$$\sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geq \alpha \|\xi\|^2, \quad \xi \in \mathbb{R}^N,\ \forall x \in \Omega.$$

Moreover, we assume the coefficient functions to be such that $-A - \epsilon I$ is dissipative for some $\epsilon > 0$ with $D(A) \subset R(I - hA)$, where $0 < h < \frac{1}{\epsilon}$. Then, the solution operator of (2.1a), (2.1b) is a non-linear semi-group ([7, 8]). In case f does not depend on y, (2.1a) represents a linear evolution equation whose solution operator is a strongly continuous linear semi-group.

The computation of snapshots is done with respect to a spatially discretized system of equations (2.1a), (2.1b), for instance, the finite element method ([25, 26]). We define the function spaces:

$$W(0,T) = H^1((0,T), H^{-1}(\Omega)) \cap L^2((0,T), H_0^1(\Omega)),$$
$$\bar{W}(0,T) = \{y \in W(0,T) | f(t,y) \in L^2(\Omega), t \in (0,T)\},$$

The finite element method is based on the variational formulation of (2.1a), (2.1b) such that for almost all $t \in (0,T)$ and $v \in H_0^1(\Omega)$, it holds:

$$\begin{cases} < \dfrac{\partial y}{\partial t}, v >_{H^{-1}(\Omega),H_0^1(\Omega)} + a(y,v) - < f(t,y(t)), v >_{L^2(\Omega)} = 0, & \text{(2.4a)} \\ \qquad\qquad < y(0), v >_{L^2(\Omega)} = < y_0, v >_{L^2(\Omega)}, & \text{(2.4b)} \end{cases}$$

where the bi-linear form $a(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega)$ is given by:

$$a(y,v) = -\sum_{i,j=1}^{N} \int_{\Omega} a_{ij} \frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} \, dx + \sum_{i=1}^{N} \int_{\Omega} b_i v \frac{\partial y}{\partial x_i} \, dx + \int_{\Omega} cyv \, dx \tag{2.5}$$

A function satisfying (2.4a), (2.4b) is called a weak solution of (2.1a), (2.1b). For the existence of a weak solution, we refer to [5, 14]. In particular, under the previous assumptions on the operator A and the non-linear mapping f, the existence of a weak solution is guaranteed.

For the finite element approximation of (2.4a) and (2.4b), we let $\{\mathcal{T}_h(\Omega)\}$ be a shape regular family of geometrically conforming simplicial triangulation of Ω. We denote the set of nodal points of $\mathcal{T}_h(\Omega)$ by $\mathcal{N}_h(\Omega)$ and set $n = card(\mathcal{N}_h(\Omega))$. For $T_r \in \mathcal{T}_h(\Omega)$, we refer to h_{T_r} as the diameter of T_r and set $h = \max\{h_{T_r} | T_r \in \mathcal{T}_h(\Omega)\}$. We further refer to $V_h \subset H_0^1(\Omega)$ as the finite element space of continuous, piece-wise linear finite elements with respect to $\mathcal{T}_h(\Omega)$. Let $P_1(T_r)$ be the set of polynomials of degree no more than 1 on T_r, then:

$$V_h = \{v_h \in C(\bar{\Omega}) | v_h | T_r \in P_1(T_r), T_r \in \mathcal{T}_h(\Omega), \ v_h | \Gamma = 0\}. \tag{2.6}$$

We note that $dimV_h = n$ and $V_h = span\{\phi_h^1, \cdots, \phi_h^n\}$, where ϕ_h^i stands for the nodal basis function associated with the nodal point $a_{i,h} \in \mathcal{T}_h(\Omega)$. Then the finite element approximation of (2.4a) and (2.4b) reads: Find $y_h \in C^1([0,T], V_h)$ such that for all

$t \in [0, T]$ and $v_h \in V_h$ it holds:

$$\begin{cases} < \dfrac{\partial y_h}{\partial t}, v_h >_{L^2(\Omega)} + a(y_h, v_h) + < f(y_h), v_h >_{L^2(\Omega)} = 0, & (2.7a) \\ \qquad < y_h(0), v_h >_{L^2(\Omega)} = < y_{h,0}, v_h >_{L^2(\Omega)}, & (2.7b) \end{cases}$$

We may identify the finite element function $y_h(\cdot, t)$ with a vector $\mathbf{y}(t)$ and denote by $\mathbf{M},\ \mathbf{A} \in \mathbb{R}^{n\times n}$ the mass and the stiffness matrix, respectively, such that:

$$\mathbf{y}(t) = \begin{pmatrix} y_{1,h}(t) \\ \vdots \\ y_{n,h}(t) \end{pmatrix},\ y_{i,h}(t) = y_h(a_{i,h}, t),\ a_{i,h} \in \mathcal{T}_h(\Omega),$$

$$\mathbf{M} = (\mathbf{M}_{i,j}) = (< \phi_h^i, \phi_h^j >_{L^2(\Omega)}),\ 1 \le i, j \le n,$$

$$\mathbf{A} = (\mathbf{A}_{i,j}) = (a(\phi_h^i, \phi_h^j)),\ 1 \le i, j \le n.$$

We further refer to $\mathbf{y}_0$ as the vectors with components:

$$\mathbf{y}_0 = \begin{pmatrix} < y_{h,0}, \phi_h^1 >_{L^2(\Omega)} \\ \vdots \\ < y_{h,0}, \phi_h^n >_{L^2(\Omega)} \end{pmatrix},$$

and denote by $\mathbf{f}(t, \mathbf{y}) : \mathbb{R}^n \to \mathbb{R}^n,\ t \in [0, T]$ the non-linear map:

$$\mathbf{f}(t, \mathbf{y}) = \begin{pmatrix} < f(t, \sum_{j=1}^{n} y_{j,h}\phi_h^j), \phi_h^1 >_{L^2(\Omega)} \\ \vdots \\ < f(t, \sum_{j=1}^{n} y_{j,h}\phi_h^M), \phi_h^n >_{L^2(\Omega)} \end{pmatrix}.$$

Then, the finite element approximation (2.7a) and (2.7b) can be equivalently written as the following initial-value problem for a system of non-linear first order ODEs:

$$\begin{cases} \mathbf{M}\mathbf{y}'(t) + \mathbf{A}\mathbf{y}(t) - \mathbf{f}(t, \mathbf{y}(t)) = 0, \quad t \in [0, T], & \text{(2.8a)} \\ \mathbf{M}\mathbf{y}(0) = \mathbf{y}_0. & \text{(2.8b)} \end{cases}$$

Since $\mathbf{f}$ is continuous and satisfies the Lipschitz condition in the second argument, due to the Theorem of Picard-Lindelöf ([6]), the initial-value problem (2.8a) and (2.8b) admits a unique solution.

We consider (2.8a) and (2.8b) as the Full Order Model (**FOM**) for which we will describe the application of the POD and POD-DEIM in the following Chapter 3.

Chapter 3

POD and DEIM

3.1 Galerkin POD for Evolution Equations

3.1.1 POD and SVD

In this chapter, we consider the singular value decomposition for a linear map from finite dimensional Hilbert space $\mathbf{V}$ into another finite dimensional Hilbert space $\mathbf{W}$.

Proposition 3.1. *Let $F : \mathbf{V} \to \mathbf{W}$ be a linear operator and $dim(\mathbf{V}) = m,\ dim(\mathbf{W}) = n,\ m \leq n$. Then there exist non-zero real numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$; orthogonal basis $\{\upsilon_k\}_{k=1}^m$ of V and orthogonal basis $\{\omega\}_{k=1}^m$ of W such that:*

$$F(\upsilon_k) = \sigma_k \omega_k, \textit{and } F^*(\omega_k) = \sigma_k \upsilon_k \quad \textit{for } k = 1, \cdots, m.$$

Proof. For a proof we refer to Proposition 1 in [9]. □

Proposition 3.1 implies the following corollary.

Corollary 3.2. *Under the hypothesis of Proposition 3.1 we have:*

$$F^*F(\omega_k) = \sigma_k^2\omega_k, \text{ and } FF^*(\upsilon_k) = \sigma_k^2\upsilon_k \text{ for } k = 1, \cdots, m.$$

The Proper Orthogonal Decomposition (**POD**) can be formulated as a constrained minimization problem. For that purpose, we consider the initial value problem (2.8a), (2.8b) and assume the availability of the snapshots $\{\mathbf{y}_0, \mathbf{y}_1, \cdots, \mathbf{y}_M\} \subset \mathbb{R}^n$, where $\mathbf{y}_i = \mathbf{y}(t_i)$ is the finite element solution of (2.8a), (2.8b) at time instance t_i, where $0 \leq t_i \leq T,\ 0 \leq i \leq M$.

For any $l \in \mathbb{N},\ l < n$, we want to find the orthonormal basis $\{\omega_k\}_{k=1}^l$ such that the mean square error between the members of the ensembles and their $l - th$ partial sum is minimized:

$$(\mathbf{P}^l) \begin{cases} \min J(\omega_, \cdots, \omega_l) = \sum_{j=1}^{M} ||\mathbf{y}_j - \sum_{k=1}^{l} < \mathbf{y}_j, \omega_k > \omega_k||^2 \\ \text{subject to: } (\omega_i, \omega_j) = \delta_{ij} \text{ for } 1 \leq i, j \leq l \end{cases}$$

The solution to problem $\mathbf{P}^l$ is given in [9] as stated in the following corollary.

Corollary 3.3. *Let* $Y : \mathbb{R}^n \to \mathbb{R}^n$*, such that* $Y(\upsilon) = \sum_{k=1}^{M} < \upsilon, e_k > y_k$*, where* e_k*'s are the canonical basis in* $\mathbb{R}^n$*. Let* $K = (k_{ij}),\ 1 \leq i, j \leq n,\ k_{ij} =< \mathbf{y}_j, y_i >$*, then the POD-basis* $\{\omega_k\}_{k=1}^l$ *are given by* $\omega_k = \frac{1}{\sqrt{\lambda_k}} Y(\upsilon_k)$*, where the pair* $\{\lambda_k, \upsilon_k\}$ *solves the following eigenvalue problem:*

$$K\upsilon_k = \lambda_k\upsilon_k, \text{ for } k = 1, \cdots, l. \tag{3.2}$$

The argumented minimal takes the form:

$$argmin(\boldsymbol{P}^l) = \sum_{k=l+1}^{M} \lambda_k \tag{3.3}$$

It is clear that the POD-basis $\{\omega_k\}_{k=1}^{l}$ are the eigenvectors of matrix K and because that the error between the ensembles and their corresponding $l-th$ partial sum is bounded below by the sum of the suqared singular values after $l-th$, the major advantage of POD is that we can choose sufficiently small lower bound with the appropriately chosen dimension l.

The matrix $V = [\upsilon_1, \upsilon_2, \cdots, \upsilon_l] \in \mathbb{R}^{n \times l}$ is the so-called POD basis matrix, which is used to reduce the order of the model by replacing $\mathbf{y}(t)$ by $V\mathbf{y}^l(t)$, where $\mathbf{y}^l(t) \in \mathbb{R}^l$. We multiply V^T to the equations and get:

$$\begin{cases} V^T\mathbf{M}V\dfrac{d\mathbf{y}^l(t)}{dt} + V^T\mathbf{A}V\mathbf{y}^l(t) - \ V^T\mathbf{f}(t, V\mathbf{y}^l(t)) = 0, & (3.4a) \\ \qquad V^T\mathbf{M}V\mathbf{y}^l(0) = V^T\mathbf{y}_0 & (3.4b) \end{cases}$$

(3.4a), (3.4b) are usually refered to as the POD based **ROM**. The vector $V\mathbf{y}^l(t)$ is the POD approximation solution to the **FOM** (2.8a), (2.8b). The following estimation for the difference between the **FOM** solution and the **ROM** approximation solution holds true:

$$||\mathbf{y}_j - V\mathbf{y}_j^l|| \leq ||I_M - VV^T|| = (\sum_{k=l+1}^{M} \lambda_k)^{\frac{1}{2}} = \tau_{pod} \tag{3.5}$$

The constant τ_{pod} is the so-called POD state error.

3.2 POD Discrete Empirical Interpolation (POD-DEIM)

We let $\widetilde{N}(u) = V^T\mathbf{f}(t, Vu)$ and assume that the non-linear function $f(t, y(t))$ is continuous differentiable in y, then the Jacobian of the non-linear term of equation 3.4a is of the form:

$$\hat{\mathbf{J}}_f(u) = -V^T\mathbf{f}_y(t, Vu)V. \tag{3.6}$$

In most circumstances, the evaluation of $\widetilde{N}(u)$ and the inverse of Jacobian $\hat{\mathbf{J}}_f^{-1}(u)$ at certain $u = \mathbf{y}^l(t_j)$ are required to compute the solutions of the initial value problem (3.4a), (3.4b). But the non-linear term $\widetilde{N}(u)$ has a computational complexity that depends on n, the dimension of the original **FOM**. In [4], page 2743, the authors show that if there are q components in the non-linear function f, then the computational complexity for evaluating $\widetilde{N}(u)$ is roughly $O(\alpha(q) + 4nl)$, where α is some function of q, i.e., the computational complexity of POD based **ROM** is not reduced at all for non-linear terms.

The same inefficiency occurs in the evaluation of Jacobian $\hat{\mathbf{J}}_f(u)$ when a solvers of Newton-type is used because at each Newton iteration, besides the non-linear term $\widetilde{N}(u)$, the Jacobian must also be computed with a computational cost which still depends on the dimension of the **FOM**.

$$\hat{\mathbf{J}}_f(u) = -\underbrace{V^T}_{l\times n}\underbrace{\mathbf{f}_y(t, Vu)}_{n\times n}\underbrace{V}_{n\times l} \tag{3.7}$$

The computational complexity for evaluating 3.7 is roughly $O(\alpha(n) + 2n^2l + 2nl^2 + 2nl)$ if $\hat{\mathbf{J}}_f$ is dense, $O(\alpha(n) + nl + 2nl^2 + 2nl)$ if $\hat{\mathbf{J}}_f$ is sparse or diagonal (From [4], page 2743).

In order to reduce the computational complexity of $\widetilde{N}(u)$, an efficient way is to approximate the non-linear function $f(u) = \mathbf{f}(t, Vu)$ by its projection onto the subspace spanned by a set of linearly independent vectors $\{N_1, N_2, \cdots, N_{l_d}\}$, where $N_i \in \mathbb{R}^{n\times 1}$, $1 \le i \le l_d$, such that for matrix $N = [N_1, N_2, \cdots, N_{l_d}]$ and coefficient vector $c(u)$ the following approximation holds true:

$$f(u) \approx Nc(u).$$

We use DEIM method to deal with the approximation of non-linear terms and choose matrix P such that:

$$P = [e_{p_1}, e_{p_2}, \cdots, e_{p_{l_d}}] \subset \mathbb{R}^{n \times l_d}, \text{ where } e_{p_i} = [0, \cdots, 0, 1, 0, \cdots, 0]^T \in \mathbb{R}^{n \times 1},$$

and

$$P^T f(u) = (P^T N) c(u). \tag{3.8}$$

Then, we have:

$$f(u) \approx N c(u) = N(P^T N)^{-1} P^T f(u) = N(P^T N)^{-1} f(P^T u). \tag{3.9}$$

One can combine POD with DEIM, that is to say, POD can be used to reduce the linear terms while DEIM is used to deal with the non-linear term. As for the projection basis matrix N, one can apply the POD method to the matrix whose columns are the non-linear evaluations at different time instances, i.e., the non-linear snapshot matrix, then choose the first l_d eigenvectors as the columns of matrix N. In [4], an algorithm has been provided by the authors to compute the interpolation indices $\{p_i\}_{i=1}^{l_d}$ for DEIM. The POD-DEIM algorithm is listed in Algorithm1.

We let $\widetilde{f}(u) = N(P^T N)^{-1} P^T f(u)$. It is clear that $\widetilde{f}(u)$ is indeed an interpolation approximation of the non-linear function $f(u)$ in the sense that

$$P^T \widetilde{f}(u) = P^T N(P^T N)^{-1} P^T f(u) = (P^T N)(P^T N)^{-1} P^T f(u) = P^T f(u)$$

The following lemma states the error between the $f(u)$ and its approximation $\widetilde{f}(u)$.

Algorithm 1 POD-DEIM

INPUT: Non-linear snapshots $\{n_1, \cdots, n_m\} \subset \mathbb{R}^n$.

OUTPUT: Interpolation indices $p = [p_1, \cdots, p_{l_d}]^T \in \mathbb{R}^n$.

Apply POD to matrix $[n_1, \cdots, n_m]$ and choose the first l_d POD basis as $\mathbf{U} \in \mathbb{R}^{n \times l_d}$.

DEIM projection basis matrix$\mathbf{U} = [u_1, \cdots, u_{l_d}]$.

$[|\rho|, p_1] = max\{|U_1|\}$,

$N = [u_1], \; P = [e_{p_1}], \; p = [p_1]$,

for $i = 2 : l_d$ **do**

Solve $(P^T N)c = P^T u_i$ for c,

$r = u_i - Nc$,

$[|\rho|, p_i] = max\{|r|\}$,

$N \leftarrow [N, u_i], \; P \leftarrow [P, e_{p_i}], \; p \leftarrow [p, p_i]$.

end for

Lemma 3.4. *Let $f \in \mathbb{R}^n$ be any vector function of dimension n, $N \in \mathbb{R}^{n \times l_d}$ be a given DEIM basis matrix. Then the DEIM approximation of order $l_d \leq n$ for f in the space spanned by the columns of N is given by:*

$$\widetilde{f} = N(P^T N)^{-1} P^T f,$$

where $P = [e_{p_1}, e_{p_2}, \cdots, e_{p_{l_d}}]$ with $p = [p_1, \cdots, p_{l_d}]$ be the output of Algorithm 1. Moreover, an error bound for $\widetilde{f}$ is given by:

$$||f - \widetilde{f}|| \leq C_D ||(\mathbf{I}_{n \times n} - NN^T) f||, \ C_D = ||(P^T N)^{-1}||. \tag{3.10}$$

Proof. We refer to the Lemma 3.2 in [4]. □

Chapter 4

Snapshot Location in POD

4.1 Error estimator for systems of ODEs

This section is devoted to the numerical solution of an initial-value problem for a system of non-linear first order ODEs by the implicit Euler scheme and the estimation of the global discretization error by a hierarchical type error estimator based on the implicit trapezoidal rule. We first focus on general non-linear systems in $\mathbb{R}^n$ and deal with some modifications for linear and semi-linear systems, then apply the results to the **FOM** (2.8a), (2.8b).

4.1.1 General non-linear systems.

Given a function $f : \mathbb{R}_+ \times D \to \mathbb{R}^n$, $D \subset \mathbb{R}^n$, $\mathbf{y}_0 \in \mathbb{R}^n$, and $T > 0$, we consider the initial-value problem

$$y'(t) = f(t, y(t)), \quad 0 \le t \le T, \tag{4.1a}$$

$$y(0) = \mathbf{y}_0, \tag{4.1b}$$

where $y'(t) := dy(t)/dt$. If f is continuous in both arguments and satisfies a Lipschitz condition in the second argument, the theorem of Picard-Lindelöf [6] guarantees the existence and uniqueness of a solution.

For the numerical solution we assume

$$0 =: t_0 < t_1 < \cdots < t_M := T, \ \ M \in \mathbb{N}, \tag{4.2}$$

to be a partitioning of the time interval $[0, T]$ with step sizes

$$\tau_m := t_m - t_{m-1}, \quad 1 \leq m \leq M, \tag{4.3}$$

and we set $\tau := \max \{\tau_m \mid 1 \leq m \leq M\}$. We denote by y^m the approximation of $y(t_m), 0 \leq m \leq M$, by the implicit Euler scheme

$$\begin{cases} y^m - \tau_m \ f(t_m, y^m) = y^{m-1}, \quad 1 \leq m \leq M, & \text{(4.4a)} \\ y^0 = y_0. & \text{(4.4b)} \end{cases}$$

We note that at each time step (4.4a) represents a non-linear algebraic system which can be solved by a Newton type method.

If the solution y of (4.1a),(4.1b) satisfies $y \in C^2([0, T])$, the implicit Euler method (4.4a),(4.4b) is convergent of order 1, i.e., there exists a constant $C > 0$ such that

$$\|y(t_m) - y^m\| \leq C \ \tau, \quad 1 \leq m \leq M.$$

We are interested in a computable quantity e^m, called error estimator, that provides an upper and a lower bound for the global discretization error $y(t_m) - y^m$ in the sense that there exist constants $0 < C_E \leq C_R$ satisfying:

$$\|y(t_m) - y^m\| \leq C_R \ e^m, \tag{4.5a}$$

$$\|y(t_m) - y^m\| \geq C_E \ e^m. \tag{4.5b}$$

An estimator for which (4.5a) holds true is called reliable, since for a given tolerance $TOL > 0$ the criterion

$$e^m < TOL \tag{4.6}$$

assures that the global discretization error is of the same order of magnitude as TOL. It is said to be efficient, if (4.5b) is satisfied which avoids overestimation and thus prevents waste of computational time.

Let us assume that $\hat{y}^m, 0 \leq m \leq M$, with $\hat{y}^0 = \mathbf{y}_0$ are approximations of $y(t_m)$ which improve on y^m in the sense that for some constant $0 \leq q < 1$ it holds

$$\|y(t_m) - \hat{y}^m\| \leq q \ \|y(t_m) - y^m\|, \quad 1 \leq m \leq M. \tag{4.7}$$

We define $\eta^m = \hat{y}^m - y^m$ and choose $e^m := \|\eta^m\|$ as an estimator. The triangle inequality

$$\|y(t_m) - y^m\| \leq \|y(t_m) - \hat{y}^m\| + e^m \leq q \ \|y(t_m) - y^m\| + e^m$$

infers reliability of the estimator according to

$$\|y(t_m) - y^m\| \leq \frac{1}{1-q} \ e^m, \tag{4.8}$$

whereas the other direction of the triangle inequality

$$\|y(t_m) - y^m\| \geq e^m - \|y(t_m) - \hat{y}^m\| + e^m \geq e^m - q \ \|y(t_m) - y^m\|$$

implies efficiency of the estimator by means of

$$\frac{1}{1+q} \ e^m \leq \|y(t_m) - y^m\|. \tag{4.9}$$

A natural candidate for providing $\hat{y}^m$ is the implicit trapezoidal rule

$$\begin{cases} \hat{y}^m - \frac{\tau_m}{2} \ f(t_m, \hat{y}^m) = \hat{y}^{m-1} + \frac{\tau_m}{2} \ f(t_{m-1}, \hat{y}^{m-1}), \quad 1 \leq m \leq M, & \text{(4.10a)} \\ \hat{y}^0 = \mathbf{y}_0, & \text{(4.10b)} \end{cases}$$

which is known to be convergent of order 2 provided $y \in C^3([0,T])$, i.e.,

$$\|y(t_m) - y^m\| \le C\ \tau^2, \quad 1 \le m \le M.$$

Hence, for sufficiently smooth y and sufficiently small τ the so-called saturation assumption (4.7) can be expected to hold true.

4.1.2 Linear and semi-linear systems.

In the linear case, the right-hand side in (4.1a) is given by $f(t,y) := -Ay + f(t)$, where we assume $A \in \mathbb{R}^{n\times n}$ to be regular and $f \in C([0,T])$. Then, (4.1a),(4.1b) reads

$$\begin{cases} y'(t) + Ay(t) = f(t), \quad 0 \le t \le T, & (4.11a) \\ \qquad\qquad y(0) = y_0. & (4.11b) \end{cases}$$

At each time step, the implicit Euler scheme requires the solution of the linear system

$$(I + \tau_m A)y^m = y^{m-1} + \tau_m f(t_m), \quad 1 \le m \le M \tag{4.12}$$

whereas the trapezoidal rule takes the form

$$\begin{aligned} (I + \frac{\tau_m}{2}A)\hat{y}^m &= (I - \frac{\tau_m}{2}A)\hat{y}^{m-1}, \\ &\quad + \frac{\tau_m}{2}(f(t_{m-1}) + f(t_m)), \quad 1 \le m \le M. \end{aligned} \tag{4.13}$$

We rearrange terms in (4.13):

$$(I + \tau_m A)\hat{y}^m = \hat{y}^{m-1} + \frac{\tau_m}{2}A(\hat{y}^m - \hat{y}^{m-1}) + \frac{\tau_m}{2}(f(t_{m-1}) + f(t_m)). \tag{4.14}$$

If we replace $\hat{y}^m$, $\hat{y}^{m-1}$ on the right-hand side of (4.14) by y^m, y^{m-1} respectively, then subtract (4.12) from (4.14), we obtain

$$(I + \tau_m A)\eta^m = \frac{\tau_m}{2}\Big(A(y^m - y^{m-1}) - f(t_m) + f(t_{m-1})\Big).$$

Hence, the error estimator can be computed as follows

$$e^m = \frac{\tau_m}{2}\|(I+\tau_m A)^{-1}\Big(A(y^m - y^{m-1}) - f(t_m) + f(t_{m-1})\Big)\|. \tag{4.15}$$

In the semi-linear case, the right-hand side in (4.1a) reads $f(t,y) := -Ay + f(y)$, where $A \in \mathbb{R}^{n\times n}$ is regular and $f \in C(D)$. The initial-value problem (4.1a),(4.1b) takes the form

$$y'(t) + Ay(t) = f(t, y(t)), \quad 0 \le t \le T, \tag{4.16a}$$

$$y(0) = y_0, \tag{4.16b}$$

and the implicit Euler scheme reads

$$(I + \tau_m A)y^m - \tau_m f(t_m, y^m) = y^{m-1}, \quad 1 \le m \le M, \tag{4.17}$$

whereas the trapezoidal rule takes the form

$$\begin{aligned}(I + \frac{\tau_m}{2}A)\hat{y}^m - \frac{\tau_m}{2}(f(t_{m-1}, \hat{y}^{m-1}) + f(t_m, \hat{y}^m))\\ = (I - \frac{\tau_m}{2}A)\hat{y}^{m-1}, \quad 1 \le m \le M.\end{aligned} \tag{4.18}$$

We use a slightly modified trapezoidal rule by replacing $f(t_{m-1}, \hat{y}^{m-1})$ with $f(t_m, y^m)$, i.e., we use y^m from the implicit Euler scheme as a predictor and the trapezoidal scheme as a corrector:

$$(I + \frac{\tau_m}{2}A)\hat{y}^m - \frac{\tau_m}{2}(f(t_m, y^m) + f(t_m, \hat{y}^m)) = (I - \frac{\tau_m}{2}A)\hat{y}^{m-1} \tag{4.19}$$

and rearrange (4.19) as:

$$(I + \tau_m A)\hat{y}^m - \frac{\tau_m}{2}(f(t_m, y^m) + f(t_m, \hat{y}^m)) = \hat{y}^{m-1} + \frac{\tau_m}{2}A(\hat{y}^m - \hat{y}^{m-1}). \tag{4.20}$$

Subtracting (4.17) from (4.20), we obtain:

$$(I + \tau_m A)\eta^m - \frac{\tau_m}{2}(-f(t_m, y^m) + f(t_m, \hat{y}^m) = \hat{y}^{m-1} - y^{m-1} + \frac{\tau_m}{2}A(\hat{y}^m - y^{m-1}). \tag{4.21}$$

We rewrite (4.21) as:

$$(I + \tau_m A)\eta^m - \tau_m f(t_m, \hat{y}^m) = \hat{y}^{m-1} - y^{m-1} + \frac{\tau_m}{2} A(\hat{y}^m - y^{m-1})$$
$$- \frac{\tau_m}{2}(f(t_m, y^m) + f(t_m, \hat{y}^m)). \tag{4.22}$$

If we replace $\hat{y}^{m-1}$, $\hat{y}^m$ by y^{m-1}, y^m on the right hand side of (4.22) and note that $\hat{y}^m = \eta^m + y^m$, we obtain the equation for η^m:

$$(I + \tau_m A)\eta^m - \tau_m f(t_m, y^m + \eta^m) = \frac{\tau_m}{2} A(y^m - y^{m-1})$$
$$- \tau_m f(t_m, y^m), \quad 1 \leq m \leq M. \tag{4.23}$$

The error estimator e^m then can be computed from the norm of η^m.

Remark 4.1. By replacing $\hat{y}^i$ with y^i, for $i = m-1$ and m in (4.14) and (4.22), we actually assume that the initial values for the implicit Euler scheme and the modified trapezoidal scheme are the same. This can be done by taking $\mathbf{y}_{avg}^{m-1} = \frac{1}{2}(\hat{y}^{m-1} + y^{m-1})$ as the initial value for both schemes.

Remark 4.2. Instead of the implicit Euler scheme (4.17) we may also use the semi-implicit Euler scheme:

$$(I + \tau_m A)y^m = y^{m-1} + \tau_m f(y^{m-1}), \quad 1 \leq m \leq M, \tag{4.24}$$

which has the advantage that no non-linear system has to be solved. In this case, the estimator is given as follows

$$e^m = \frac{\tau_m}{2} \|(I + \tau_m A)^{-1} \Big(A(y^m - y^{m-1}) + f(y^m) - f(y^{m-1}) \Big) \|. \tag{4.25}$$

4.2 Equilibration of the error in time

The a posteriori error estimation described in section 4.1 provides the possibility of computing an adaptive selection of time instants $t_m \in (0,T), 1 \leq m \leq M-1$, such that in each time sub-interval $[t_{m-1}, t_m]$ the global discretization error e^m is bounded by a certain portion TOL of $\|y^m\|$. In order to enforce $\|\frac{e^m}{y^m}\| < TOL$, the length of each time step must be adaptive and as a result, the number of time steps M is not fixed but dependent on the behaviour of the solution.

A different point of view is to prescribe a fixed number of time steps and choose the time instants $t_m, 1 \leq m \leq M-1$, such that in each time interval $[t_{m-1}, t_m]$ the global discretization error $\|y(t_m) - y^m\|$ is of the same order of magnitude. Since we do not know the exact solution, as a substitute for $\|y(t_m) - y^m\|$ we use the error estimator e^m. In other words, the goal is to find $t_m, 1 \leq m \leq M-1$, such that

$$e^m \approx e_{m'}, \quad 1 \leq m \neq m' \leq M. \tag{4.26}$$

The vectors $\{\eta^m\}_{m=1}^{M}$ can be computed by applying Newton's method to (4.23) for a semi-linear system. In this section, we assume that $f(t, y(t))$ is semi-linear, continuous in t and satisfies the following Lipschitz condition in the second argument:

$$||f(\cdot, x_1) - f(\cdot, x_2)|| \leq L_f ||x_1 - x_2|| \tag{4.27}$$

We consider the semi-discretized systems (2.8a), (2.8b):

$$\mathbf{M}\mathbf{y}'(t) + \mathbf{A}\mathbf{y}(t) - \mathbf{f}(t, \mathbf{y}(t)) = 0, \quad t \in [0,T],$$

$$\mathbf{M}\mathbf{y}(0) = \mathbf{y}_0.$$

We discretize the above initial value problem in time by the implicit Euler scheme and the modified trapezoidal rule with an implicit Euler predictor, respectively and solve for $\mathbf{y}^m$ and η^m by Newton's method.

For $\mathbf{y}^m$, we solve the following equation:

$$(\mathbf{M} + \tau_m \mathbf{A})\mathbf{y}^m - \tau_m \mathbf{f}(t_m, \mathbf{y}^m) - \mathbf{y}^{m-1} = 0. \tag{4.29}$$

For $\eta^m = \hat{\mathbf{y}}^m - \mathbf{y}^m$, we solve:

$$(\mathbf{M} + \tau_m \mathbf{A})\eta^m - \tau_m \mathbf{f}(t_m, \mathbf{y}^m + \eta^m) - \frac{\tau_m}{2}\mathbf{A}(\mathbf{y}^m - \mathbf{y}^{m-1}) + \tau_m \mathbf{f}(t_m, \mathbf{y}^m) = 0. \tag{4.30}$$

We define the relative error $e_{rel} = [e^1_{rel}, \cdots, e^M_{rel}]$:

$$e^m_{rel} = \frac{e^m}{max\{e^m : 1 \leq m \leq M\}}, \quad 1 \leq m \leq M. \tag{4.31}$$

Then, the goal of the error equilibration in time is to find the distribution of time instances $\{0 = t_0, t_1, \cdots, t_{M-1}, t_M = T\}$ such that the error estimators $\{e^m\}^M_{m=1}$ are approximately the same, i.e., the difference between the error estimators are bounded by the given tolerance TOL:

$$1 - TOL < e^m_{rel} < 1.$$

As $\frac{e^m}{max\{e^m\}} \leq \frac{e^m}{mean\{e^m\}}$, so the error equilibration in time can be formulated as a constrained minimization problem of variance as follows:

$$\min J(\overrightarrow{t}, \overrightarrow{\mathbf{y}}, \overrightarrow{\hat{\mathbf{y}}}) = \frac{1}{M}\sum_{i=1}^{M} |e^i - \frac{1}{M}\sum_{i=1}^{M} e^i|^2, \tag{4.32}$$

$$\tag{4.33}$$

subject to:

$$
\begin{cases}
e^m = \|\eta^m\|, \quad 1 \le m \le M, & (4.34a) \\
\eta^m = \|\mathbf{y}^m - \hat{\mathbf{y}}^m\|, & (4.34b) \\
(\mathbf{M} + \tau_m \mathbf{A})\mathbf{y}^m - \tau_m \mathbf{f}(t_m, \mathbf{y}^m) - \mathbf{y}^{m-1} = 0, \quad 1 \le m \le M, & (4.34c) \\
(\mathbf{M} + \tau_m \mathbf{A})\eta^m - \tau_m \mathbf{f}(t_m, \mathbf{y}^m + \eta^m) - \frac{\tau_m}{2}\mathbf{A}(\mathbf{y}^m - \mathbf{y}^{m-1}) & \\
+ \tau_m \mathbf{f}(t_m, \mathbf{y}^m) = 0, \quad 1 \le m \le M, & (4.34d)
\end{cases}
$$

where the vectors $\overrightarrow{t}$, $\overrightarrow{\mathbf{y}}$, $\overrightarrow{\hat{\mathbf{y}}}$ are given by:

$$
\overrightarrow{t} = \begin{pmatrix} t_1 \\ \vdots \\ t_M \end{pmatrix}, \quad \overrightarrow{\mathbf{y}} = \begin{pmatrix} \mathbf{y}^1 \\ \vdots \\ \mathbf{y}^M \end{pmatrix}, \quad \overrightarrow{\hat{\mathbf{y}}} = \begin{pmatrix} \hat{\mathbf{y}}^1 \\ \vdots \\ \hat{\mathbf{y}}^M \end{pmatrix}.
$$

The TOL can be used as a stopping criteria for the evaluation of cost function J. In this work, we provide an algorithm based on the bi-section method. For any given time partitioning, we compare each e^m with the mean value of the error estimators, then extend or shrink the length of the time sub-interval $[t_{m-1}, t_m]$ accordingly. In order to update several time sub-intervals simultaneously, we collect the total length that we cut off from the sub-intervals which need to be shrunk, then distribute it to the sub-intervals which need to be extended accordingly to the corresponding ratio of $\frac{e^m}{mean}$. The detailed algorithm is listed in Algorithm 2:

4.2.1 Newton's method for full order model (FOM)

In this section, we apply Newton's Method to equations (4.29), (4.30) and consider the existence and uniqueness of the Newton iterates.

Algorithm 2 Equilibration of error in time

INPUT Tolerance TOL, $P^0 = \{0 = t_0, t_1, \cdots, t_{M-1}, t_M = T\}$

OUTPUT Time instance $P^* = \{0 = t_0^*, t_1^*, \cdots, t_{M-1}^*, t_M^* = T\}$ for equilibration of error

Initialization: $e_{rel} = [1, 1, \cdots, 1]$, TOL, $flag = 1$.

while $flag > TOL$ **do**

0. Compute time step length $\Delta t_m = t_m - t_{m-1},\ 1 \le m \le M$.

1. Compute $\{\mathbf{y}^m\}_{m=1}^M$ from (4.29).

2. Input $\{\mathbf{y}^m\}_{m=1}^M$ to (4.30) and compute $\{\eta^m\}_{m=1}^M$.

3. Compute e_{rel} from $\{\eta^m\}_{m=1}^M$ according to (4.31), $flag = max\{e_{rel}\}$.

4. Compute the average of the error estimator $\hat{e} = \frac{1}{M}\sum_{i=1}^{M} e^m$.

5. Compute $\alpha_m = 1 - \frac{\hat{e}}{e^m},\ 1 \le m \le M$.

6. Compute $\Delta t_+ = \sum_{i \in S_+} \alpha_i \Delta t_i$, $\Delta t_- = -\sum_{j \in S_+^c} \alpha_j \Delta t_j$, where $S_+ = \{m : \alpha_m \ge 0\}$.

7. $\Delta\hat{t} = min\{\Delta t_+, \Delta t_-\}$.

while $\Delta\hat{t} > 0$ **do**

Loop 1.1 $j = 1$ where $j \in S_+^c$;

Loop 1.2 $\Delta t_j = \Delta t_j - \min\{-\alpha_j \Delta t_j, \Delta\hat{t}\}$;

Loop 1.3 $\Delta\hat{t} = \Delta\hat{t} - \min\{-\alpha_j \Delta t_j, \Delta\hat{t}\}$; $j = j + 1$;

end while

while $\Delta\hat{t} > 0$ **do**

Loop 2.1 $i = 1$ where $i \in S_+$;

Loop 2.2 $\Delta t_i = \Delta t_i + \min\{\alpha_j \Delta t_i, \Delta\hat{t}\}$;

Loop 2.3 $\Delta\hat{t} = \Delta\hat{t} + \min\{\alpha_i \Delta t_i, \Delta\hat{t}\}$; $i = i + 1$;

end while

8. $P = \{t_m | t_m = \sum_{i=1}^{m} \Delta t_i \ \ 1 \le m \le M\}$; $P^0 \leftarrow P$.

end while

We introduce the map $F : \mathbb{R}^n \to \mathbb{R}^n$ as follows:

$$F(x) = (\mathbf{M} + \tau_m \mathbf{A})x - \tau_m \mathbf{f}(t_m, x) - g(t_m). \tag{4.35}$$

At time instance $t_m,\ 1 \leq m \leq M$, where $0 = t_0 < t_1 < \cdots < t_M = T$, $\mathbf{y}^m$ and η^m are solutions to $F(x) = 0$ with different $g(t_m)$, where $g(t_m)$ is given by:

$$g(t_m) = \begin{cases} \mathbf{M}\mathbf{y}^{m-1}, & \text{for computing } \mathbf{y}^m \\ \frac{\tau_m}{2}\mathbf{A}(\mathbf{y}^m - \mathbf{y}^{m-1}) - \tau_m \mathbf{f}(t_m, \mathbf{y}^m), & \text{for computing } \eta^m \end{cases} \tag{4.36}$$

$\mathbf{y}^m$ can be solved by Newton's method with initial value $x^0 = \mathbf{y}^{m-1}$, so does η^m with initial value $x^0 = 0$.

We observe that the non-linear function $F(x)$ inherits the Lipschitz condition from f in the sense that:

$$||F(x_1) - F(x_2))|| \leq L_F||(x_1 - x_2||, \quad x_1, x_2 \in \mathbb{R}^n, \tag{4.37}$$

where the Lipschitz constant L_F is given by:

$$L_F = ||\mathbf{M} + \tau_m \mathbf{A}|| + \tau_m L_f. \tag{4.38}$$

The following two lemmas will be used in the proof of theorems thereafter.

Lemma 4.3 (Banach perturbation lemma)**.** *Let $A,\ C \in \mathbb{R}^{n\times n}$ and assume that A is invertible with $\|A\|^{-1} \leq \alpha$. If $\|A - C\| \leq \beta$ and $\alpha\beta < 1$, then C is also invertible and*

$$\|C\|^{-1} \leq \frac{\alpha}{1-\alpha\beta}. \tag{4.39}$$

Proof. We refer to ([1]), Page 45. □

Lemma 4.4 (Neumann lemma)**.** *Let $A \in \mathbb{R}^{n\times n}$ and $\|A\| < 1$. Then $(I - A)$ is invertible and*

$$\|(I-A)^{-1}\| \leq \frac{1}{1-\|A\|}. \tag{4.40}$$

Proof. We refer to ([1]), Page 45. □

It is clear that if f is continuous differentiable in its second argument, then F is also continuous differentiable with regular Jacobian:

$$F'(x) = \mathbf{M} + \tau_m\mathbf{A} + \tau_m\mathbf{f}'_x(t_m, x) \quad x \in \mathbb{R}^{n\times n} \tag{4.41}$$

And it follows from the Banach perturbation lemma(4.3) that $F'(x)$ is invertible for sufficient small τ_m. Hence, we may solve $F(x) = 0$ by Newton's method:

$$F'(x^k)\Delta x^k = -F(x^k), \tag{4.42}$$

$$x^{k+1} = x^k + \Delta x^k, \quad k \geq 0. \tag{4.43}$$

The existence and uniqueness of the Newton iterates at each time step is provided by the following theorem:

Theorem 4.5. *Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable on the domain D. Assume that its Jacobian $F'(x)$ is invertible for some initial guess $x_0 \in D$ and that the following conditions hold true:*

$$||F'(x^0)^{-1}F(x^0)|| \leq \alpha_m, \quad \alpha_m > 0, \tag{4.44}$$

$$||F'(x^0)^{-1}(F'(\mathbf{y}_1) - F'(\mathbf{y}_2))|| \leq \gamma_m||\mathbf{y}_1 - \mathbf{y}_2||, \quad \mathbf{y}_1, \mathbf{y}_2 \in D, \tag{4.45}$$

$$h = \alpha_m\gamma_m < \frac{1}{2}, \tag{4.46}$$

$$\overline{B(x^0, \rho_m)} \subset D, \ \rho_m = \frac{1 - \sqrt{1 - 2h}}{\gamma_m}. \tag{4.47}$$

Then, for the sequence $\{x^k\}_{k\in\mathbb{N}}$ of Newton iterates the following hold true:

(i) $F'(x)$ *is invertible for all Newton iterates* x^k, $k \in N$.

(ii) The sequence $\{x^k\}_{k\in\mathbb{N}}$ *is well defined with* $x^k \in \bar{B}(x_0, \rho)$ *and* $x^k \xrightarrow{k\to\infty} x^*$ *quadratically, where* $F(x^*) = 0$.

(iii) x^* *is unique in* $\bar{B}(x_0, \rho_m) \bigcup (D \bigcap B(x_0, \hat{\rho}_m))$, *where* $\hat{\rho}_m = \frac{1+\sqrt{1-2h}}{\gamma_m}$.

Proof. We refer to the Theorem 2.1 in ([3]). □

This theorem states that Newton's method is quadratically convergent provided the initial iterate x_0 is situated in a sufficiently small neighbourhood of the solution x^*, the so-called Kantorovich neighbourhood. Otherwise, the convergence of Newton's method may be achieved by appropriate globalization techniques such as a damped Newton method combined with a so-called monotonicity test for which we refer to ([3]).

4.2.2 Newton's method for the POD based Reduced Order Model (ROM)

We denote the POD basis matrix by $V \in \mathbb{R}^{n\times l}$, then the implicit Euler method for POD based **ROM** reads:

$$(\mathbf{M}_{rom} + \tau_m \mathbf{A}_{rom})\mathbf{y}^m_{rom} - \tau_m V^T \mathbf{f}(t_m, V\mathbf{y}^m_{rom}) - \mathbf{y}^{m-1}_{rom} = 0, \quad \mathbf{y}_{rom} \in \mathbb{R}^l, \tag{4.48}$$

whereas the POD based **ROM** for $\eta^m_{rom} \in \mathbb{R}^l$ is given by:

$$\begin{aligned}(\mathbf{M}_{rom} + \tau_m \mathbf{A}_{rom})\eta^m_{rom} - &\tau_m V^T \mathbf{f}(t_m, V\mathbf{y}^m_{rom} + V\eta^m_{rom}) \\ &- \frac{\tau_m}{2}\mathbf{A}_{rom}(\mathbf{y}^m_{rom} - \mathbf{y}^{m-1}_{rom}) + \tau_m V^T \mathbf{f}(t_m, V\mathbf{y}^m_{rom}) = 0.\end{aligned} \tag{4.49}$$

Let $\hat{\mathbf{f}}(t,x) : [0,T] \times \mathbb{R}^l \to [0,T] \times \mathbb{R}^l$ such that:

$$\hat{\mathbf{f}}(t,x) = V^T \mathbf{f}(t, Vx). \tag{4.50}$$

Then, the reduced non-linear function $\hat{F} : \mathbb{R}^l \to \mathbb{R}^l$ reads:

$$\hat{F}(x) = (\mathbf{M}_{rom} + \tau_m \mathbf{A}_{rom})x - \tau_m \hat{\mathbf{f}}(t_m, x) - g_{rom}(t_m), \quad x \in \mathbb{R}^l, \tag{4.51}$$

where $\mathbf{M}_{rom} = V^T \mathbf{M} V$, $\mathbf{A}_{rom} = V^T \mathbf{A} V$, $\mathbf{y}^m_{rom} = V^T \mathbf{y}^m$ and function $g_{rom}(t_m)$ is given by:

$$g_{rom}(t_m) = \begin{cases} \mathbf{M}_{rom}\mathbf{y}^{m-1}_{rom}, & \text{for computing } \mathbf{y}^m_{rom} \\ \frac{\tau_m}{2}{}_{rom}(\mathbf{y}^m_{rom} - \mathbf{y}^{m-1}_{rom}) - \tau_m \hat{\mathbf{f}}(t_m, \mathbf{y}^m_{rom}), & \text{for computing } \eta^m_{rom} \end{cases} \tag{4.52}$$

We observe that $\hat{F}$ also inherits the properties of F.

Theorem 4.6. *If F is continuously differentiable in $D \subset \mathbb{R}^n$, then $\hat{F}$ is also continuously differentiable in $\hat{D} = \{x \in \mathbb{R}^l : Vx \in D\}$ and its Jacobian is given by:*

$$\hat{F}'(x) = V^T F'(Vx)V, \quad x \in \hat{D}. \tag{4.53}$$

Proof. For $x \in \hat{D}$, we can deduce from (4.50)-(4.52) that:

$$\hat{F}(x) = V^T F(Vx). \tag{4.54}$$

So, if F is continuously differentiable, $\hat{F}$ is also continuously differentiable and its Jacobian follows directly from (4.54). □

Now, we consider Newton's method for **ROM**.

Theorem 4.7. *Let F be continuously differentiable in $x \in D \subset \mathbb{R}^n$ with Jacobian F' and $F'_V(x) = V^T F'(x) V,\ x \in D$. Assume that $F'(x)$ satisfies the affine covariant Lipschitz condition:*

$$\|F'(x)^{-1}(F'(\mathbf{y}_1) - F'(\mathbf{y}_2))\| \leq \gamma_m \|\mathbf{y}_1 - \mathbf{y}_2\|, \quad \mathbf{y}_1, \mathbf{y}_2 \in D,\ \gamma_m > 0, \tag{4.55}$$

Then, $V^T F'(x)^{-1} V$ is an approximate inverse of $F'_V(x)$ in the sense that:

$$V^T F'(x)^{-1} V = F'_V(x)^{-1} D_l, \tag{4.56}$$

where

$$D_l = I_l - V^T F'(x)(I_N - VV^T) F'(x)^{-1} V. \tag{4.57}$$

In particular, if it holds:

$$||(I_n - VV^T) F'(x)^{-1} V|| \leq q(||V^T F'(x)||)^{-1} \textit{for some } 0 \leq q < 1, \tag{4.58}$$

where I_n denotes the $n \times n$ identity matrix, then we have:

$$||F'_V(x)^{-1}|| \leq \frac{1}{1-q} ||V^T F'(x)^{-1} V||. \tag{4.59}$$

Moreover, assume that $y \in \hat{D}$ satisfies:

$$||x - Vy|| < C_{F'}^{-1}, \textit{ where } C_{F'} = \frac{\gamma}{1-q} ||V^T F'(x)^{-1}|| ||V^T F'(x)||. \tag{4.60}$$

Then it holds:

$$||\hat{F}'(y)^{-1}|| \leq \frac{1}{1-q}||V^T F'(x)^{-1}V||(1 - C_{F'}||x - Vy||)^{-1}. \tag{4.61}$$

Proof. As $F'(x)$ is regular, $F'_V(x)$ is also regular. Note that $V^T V = I_l$, so:

$$\begin{aligned} F'_V(x)V^T F'(x)^{-1}V &= V^T F'(x)VV^T F'(x)^{-1}V \\ &= I_l - V^T F'(x)(I_N - VV^T)F'(x)^{-1}V \\ &= D_l. \end{aligned}$$

Under the assumption (4.58), D_l satisfies:

$$\begin{aligned} \|I_l - D_l\| &= \|V^T F'(x)(I_N - VV^T)F'(x)^{-1}V\| \\ &\leq \|(I_N - VV^T)F'(x)^{-1}V\|\|V^T F'(x)\| \leq q < 1. \end{aligned}$$

As a consequence of Neumann's Lemma (4.4), D_l is invertible and:

$$F'_V(x)^{-1} = V^T F'(x)^{-1}VD_l^{-1},$$

which implies that $V^T F'(x)^{-1}V$ is an approximate inverse of $F'_V(x)$.

Moreover, by the Banach perturbation Lemma (4.3):

$$||D_l^{-1}|| = ||(I_l - (I_l - D_l))^{-1}|| \leq \frac{1}{1 - ||I_l - D_l||} \leq \frac{1}{1-q}, \tag{4.62}$$

it follows that (4.59) is true.

We use the Banach perturbation lemma (4.3) to prove (4.61). Let $S = F'_V(x)$, $E = \hat{F}'(y) - F'_V(x)$, where $x \in D, y \in \hat{D}$. Note that $\|V\| \leq 1$ as V is the POD basis matrix.

Then:

$$\begin{aligned} ||E|| &= ||V^T(F'(x) - F'(Vy))V|| \\ &\leq ||V^T F'(x)||||F'(x)^{-1}(F'(x) - F'(Vy))|| \leq \gamma||V^T F'(x)||||x - Vy||. \end{aligned}$$

Hence,

$$\begin{aligned}||S^{-1}E|| &\leq ||V^T F'(x)^{-1} V D_l^{-1}||||E|| \\ &\leq \frac{\gamma}{1-q}||V^T F'(x)^{-1}||||V^T F'(x)||||x - Vy|| \\ &= C_{F'}||||x - Vy|| < 1.\end{aligned}$$

As a consequence of the Banach perturbation lemma (4.3), $S + E = \hat{F}'(x)$ is invertible and:

$$\begin{aligned}||\hat{F}'(x)^{-1}|| = ||(S+E)^{-1}|| &\leq \frac{||S^{-1}||}{1 - ||S^{-1}E||} \\ &= \frac{\frac{1}{1-q}||V^T F'(x)^{-1} V||}{1 - C_{F'}||x - Vy||} \\ &= \frac{1}{1-q}||V^T F'(x)^{-1}||(1 - C_{F'}||x - Vy||)^{-1}.\end{aligned}$$

□

Remark 4.8. For the POD based **ROM**, the assumption (4.60) in Theorem (4.7) can be expected to hold true provided $\|y - V^T x\|$ is sufficiently small, which can be easily seen from:

$$||x - Vy|| \leq ||x - VV^T x|| + \|V\|||y - V^T x|| = \tau_{pod} + ||y - V^T x|| \tag{4.63}$$

The existence and uniqueness of the Newton iterates at each time step t_m of Newton's method for the POD based **ROM** is provided in the following theorem:

Theorem 4.9. *Assume that for $F : D \subset \mathbb{R}^n \to \mathbb{R}^n$ the assumptions in Theorem 4.5 hold true and (4.58) in Theorem4.7 is satisfied for $x = x_0$. Let $\hat{x}^0 = V^T x^0$. Then, for*

$\hat{F} : \hat{D} \subset \mathbb{R}^l \to \mathbb{R}^l$ the following hold true:

$$||\hat{F}'(\hat{x}_0)^{-1}\hat{F}(\hat{x}_0)|| \leq \hat{\alpha}_m, \tag{4.64}$$

$$||\hat{F}'(\hat{x}_0)^{-1}(\hat{F}'(\mathbf{y}_1) - \hat{F}'(\mathbf{y}_2))|| \leq \hat{\gamma}_m||\mathbf{y}_1 - \mathbf{y}_2||, \quad \mathbf{y}_1, \mathbf{y}_2 \in \hat{D}, \tag{4.65}$$

where the constants are given by:

$$\hat{\alpha}_m = \frac{1}{1-q}||V^T F'(x_0)||||V^T F'(x_0)^{-1}V||(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}(\alpha_m + \epsilon_m), \tag{4.66}$$

$$\hat{\gamma}_m = \frac{\gamma}{1-q}||V^T F'(x_0)||||V^T F'(x_0)^{-1}V||(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}, \tag{4.67}$$

and ϵ_m is given by:

$$\epsilon_m = ||F'(x_0)||^{-1}(||V^T(g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m)||)+$$
$$(||V^T(\mathbf{M} + \tau_m\mathbf{A})|| + \tau_m L_f||V^T||)||x_0 - V\hat{x}_0||), \tag{4.68}$$

Similar to Theorem 4.5, if we further assume:

$$\hat{h}_m = \hat{\alpha}_m\hat{\gamma}_m < \frac{1}{2},\ \overline{\hat{B}(\hat{x}_0, \hat{\rho}_m)} \subset \hat{D},\ \text{where } \hat{\rho}_m = \frac{1 - \sqrt{1 - 2\hat{h}_m}}{\hat{\gamma}_m}. \tag{4.69}$$

Then, for the sequence $\{\hat{x}^k\}_{k\in\mathbb{N}}$ of Newton iterates, the following hold true:

(i) $\hat{F}'(x)$ is invertible for all Newton iterates $\hat{x}^k$, $k \in \mathbb{N}$.

(ii) The sequence $\{\hat{x}^k\}_{k\in\mathbb{N}}$ is well defined with $\hat{x}^k \in \overline{B(x_0, \hat{\rho}_m)}$ and $\hat{x}^k \xrightarrow{k\to\infty} \hat{x}^*$ quadratically, where $\hat{F}(\hat{x}^*) = 0$.

(iii) $\hat{x}^*$ is unique in $\overline{B(\hat{x}_0, \hat{\rho}_m)} \bigcup (D \bigcap B(\hat{x}_0, \rho_m))$, where $\bar{\hat{\rho}}_m = \frac{1+\sqrt{1-2h}}{\bar{\hat{\gamma}}_m}$.

Proof. By the definition of $\hat{F}$ given in (4.51):

$$\hat{F}(\hat{x}_0)^{-1}\hat{F}(\hat{x}_0) = \hat{F}(\hat{x}_0)^{-1}(V^T((\mathbf{M}+\tau_m\mathbf{A})V\hat{x}_0)+\tau_m\mathbf{f}(t_m,V\hat{x}_0)-g(t_m,V\hat{\mathbf{y}}_m))$$

$$= \hat{F}(\hat{x}_0)^{-1}V^T\Big(\underbrace{((\mathbf{M}+\tau_m\mathbf{A})x_0+\tau_m\mathbf{f}(t_m,x_0)-g(t_m,\mathbf{y}_m))+g(t_m,\mathbf{y}_m)-g(t_m,V\hat{\mathbf{y}}_m)}_{1}-$$

$$\underbrace{((\mathbf{M}+\tau_m\mathbf{A})(x_0-V\hat{x}_0)+\tau_m(\mathbf{f}(t_m,x_0)-\mathbf{f}(t_m,V\hat{x}_0))}_{2}\Big) \tag{4.70}$$

Under assumption (4.44) in Theorem (4.5) and (4.58) in Theorem 4.7, the term 1 on the right hand side of (4.70) can be estimated as follows:

$$||\hat{F}(\hat{x}_0)^{-1}V^T((\mathbf{M}+\tau_m\mathbf{A})x_0+\tau_m\mathbf{f}(t_m,x_0)-g(t_m,\mathbf{y}_m))+g(t_m,\mathbf{y}_m)-g(t_m,V\hat{\mathbf{y}}_m)||$$
$$= ||\hat{F}(\hat{x}_0)^{-1}V^T(F'(x_0)F'(x_0)^{-1}F'(x_0))+g(t_m,\mathbf{y}_m)-g(t_m,V\hat{\mathbf{y}}_m)||$$
$$\leq ||\hat{F}(\hat{x}_0)^{-1}||\Big(\alpha_m||V^TF'(x_0)||+||V^T(g(t_m,\mathbf{y}_m)-g(t_m,V\hat{\mathbf{y}}_m))||\Big)$$
$$\leq \frac{1}{1-q}||V^TF'(x_0)^{-1}||||V^TF'(x_0)||^{-1}(1-C_{F'}||x_0-V\hat{x}_0||)^{-1}\cdot$$
$$\Big(\alpha_m+||V^TF'(x_0)||^{-1}||V^T(g(t_m,\mathbf{y}_m)-g(t_m,V\hat{\mathbf{y}}_m))||\Big) \tag{4.71}$$

The term 2 on the right hand side of (4.70) can be estimated according to (4.58) in Theorem 4.7:

$$||\hat{F}'(x_0)^{-1}V^T\Big((\mathbf{M}+\tau_m\mathbf{A})(x_0-V\hat{x_0})+\tau_m(\mathbf{f}(t_m,x_0)-\mathbf{f}(t_m,V\hat{x}_0))\Big)||$$
$$\leq \frac{1}{1-q}||V^TF'(x_0)^{-1}||||V^TF'(x_0)||(1-C_{F'}||x_0-V\hat{x}_0\|)\cdot$$
$$(||V^T(\mathbf{M}+\tau_m\mathbf{A})||+\tau_mL_f)||x_0-V\hat{x}_0||. \tag{4.72}$$

Combine the results in (4.70), (4.71) and (4.72), we can obtain 4.64.

In order to prove (4.65), we assume that $\mathbf{y}_1, \mathbf{y}_2 \in \hat{D}$. Then:

$$||\hat{F}'(\hat{x}_0)^{-1}(\hat{F}'(\mathbf{y}_1) - \hat{F}'(\mathbf{y}_2))|| = ||\hat{F}'(\hat{x}_0)^{-1}V^T(F'(V\mathbf{y}_1) - F'(V\mathbf{y}_2))||$$
$$\leq \gamma_m||\hat{F}'(\hat{x}_0)^{-1}V^T F(x_0)||||\mathbf{y}_1 - \mathbf{y}_2||$$
$$\leq \frac{\gamma_m}{1-q}||V^T F'(x_0)^{-1}||||V^T F'(x_0)||(1 - C_{F'}||x_0 - V\hat{x}_0||)||\mathbf{y}_1 - \mathbf{y}_2||$$
$$= \hat{\gamma}_m||\mathbf{y}_1 - \mathbf{y}_2||$$

Results $(i) - (iii)$ are the consequence of Theorem (4.5). □

4.2.3 Newton's method for the POD-DEIM based ROM

Throughout this section, we assume that $f(t, y)$ is continuously differentiable on its second argument, then $\mathbf{f}(t, \mathbf{y}) : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ is also continuously differentiable on $\mathbf{y}$. Moreover, we assume that f_y satisfies the affine invariant Lipschitz condition for $y = x_0$. Then $\mathbf{f}(t, \mathbf{y})$ also satisfies the affine invariant Lipschitz condition as follows:

$$||\mathbf{f}_y(t, x_0)^{-1}(\mathbf{f}_y(t, \mathbf{y}_1) - f_y(t, \mathbf{y}_2))|| \leq \gamma_f||\mathbf{y}_1 - \mathbf{y}_2||, \quad \mathbf{y}_1, \mathbf{y}_2 \in D. \tag{4.73}$$

For DEIM , we perform SVD to the non-linear snapshot matrix, then reduce the complexity of the non-linear term by replacing $\mathbf{f}$ with $\bar{\mathbf{f}}$ such that:

$$\bar{\mathbf{f}} = N(P^T N)^{-1}P^T\mathbf{f}, \tag{4.74}$$

where $P \in \mathbb{R}^{n \times l}$, $N \in \mathbb{R}^{n \times l}$ are matrices introduced in Section (3.2).

To obtain the implicit Euler solution $\mathbf{y}_d^m$ and the solution difference η_d^m of the DEIM based **FOM**, we compute a zero of the following function by Newton's method:

$$F_d(x) = (\mathbf{M} + \tau_m\mathbf{A})x + \tau_m\bar{\mathbf{f}}(t_m, x) - g(t_m), \quad x \in \mathbb{R}^n, \tag{4.75}$$

where $g_m(t_m)$ is the same as in (4.36).

Correspondingly, we solve a zero of the following function for the POD-DEIM based **ROM**:

$$\hat{F}_d(x) = (\mathbf{M}_{rom} + \tau_m \mathbf{A}_{rom})x + \tau_m V^T \bar{\mathbf{f}}(t_m, Vx) - g_{rom}(t_m), \quad x \in \mathbb{R}^l, \tag{4.76}$$

where $g_{rom}(t_m)$ is given in (4.52).

Similar to the previous sections, we are interested in the properties of F_d and $\hat{F}_d$. First of all, (4.75) clearly shows that $F_d(x)$ is continuously differentiable provided $\mathbf{f}(t, \mathbf{y})$ is continuously differentiable with respect to $\mathbf{y}$. The Jacobian of $F_d(x)$ is given by:

$$F_d'(x) = \mathbf{M} + \tau_m \mathbf{A} + \tau_m N(P^T N)^{-1} P^T f_y(t, x), \quad x \in \mathbb{R}^n. \tag{4.77}$$

Therefore, for $x \in \mathbb{R}^l$, as $\hat{F}_d(x) = V^T F_d(Vx)$, Theorem (4.6) implies that $\hat{F}_d(x)$ is also continuously differentiable and its Jacobian is given by:

$$F_d'(x) = \mathbf{M}_{rom} + \tau_m \mathbf{A}_{rom} + \tau_m V^T N(P^T N)^{-1} P^T f_y(t, Vx)V. \tag{4.78}$$

The following theorem provides the invertibility of $\hat{F}'(x)$.

Theorem 4.10. *If the non-linear map $F_d(x)$ for the DEIM based **ROM** satisfies the assumptions of Theorem (4.7), then the non-linear map $\hat{F}_d(x)$ for the POD-DEIM **ROM** is continuously differentiable in $\hat{D} \subset \mathbb{R}^n$ with its Jacobian given by:*

$$\hat{F}_d'(\hat{x}) = V^T(\mathbf{M} + \tau_m \mathbf{A} + \tau_M \bar{\mathbf{f}}(t_m, V\hat{x}))V, \quad \hat{x} \in \hat{D}, \tag{4.79}$$

where $\hat{D} = \{\hat{x} : V\hat{x} \in D\}$.

Further, if $\hat{x} \in \hat{D}$ satisfies:

$$||x - V\hat{x}|| \leq \frac{1}{C_{F'}}, \text{ for some } x \in D, \tag{4.80}$$

where $C_{F'} = \frac{1}{1-q}\tau_m\gamma_f C_D||V^T F_d'(x)^{-1}V||||f_y(t_m,x_0)||$, *then* $\hat{F}_d$ *is regular and invertible and the following holds:*

$$||\hat{F}_d'(x)^{-1}|| \leq \frac{1}{1-q}||V^T F_d'(x)^{-1}V||(1 - C_{F'}||x - V\hat{x}||)^{-1}. \tag{4.81}$$

Proof. We follow the proof of Theorem (4.7) and introduce $F_{d,V}'(x) : D \to \mathbb{R}^{l\times l}$ such that:

$$F_{d,V}'(x) = V^T F_d'(x)V. \tag{4.82}$$

Then $V^T F_d'(x)^{-1}V$ is an approximate inverse of $F_{d,V}'(x)$ in the sense that:

$$\begin{aligned} F_{d,V}'(x)V^T F_d'(x)^{-1}V &= V^T F_d'(x)VV^T F_d'(x)^{-1}V \\ &= I_l - V^T F_d'(x)(I_N - VV^T)F_d'(x)^{-1}V = D_{d,l}. \end{aligned}$$

From (4.62) in the proof of Theorem (4.7), $||D_{d,l}^{-1}|| \leq \frac{1}{1-q}$ and:

$$||F_{d,V}'(x)^{-1}|| \leq \frac{1}{1-q}||V^T F_d'(x)^{-1}V|| \tag{4.83}$$

On the other hand, the difference between $\hat{F}_d'(\hat{x})$ and $F_{d,V}'(x)$ can be estimated as follows:

$$\begin{aligned} ||\hat{F}_d'(\hat{x}) - F_{d,V}'(x)|| &= ||V^T(F_d'(V\hat{x}) - F_d'(x))V|| \\ &= ||\tau_m V^T(\bar{\mathbf{f}}_y(t_m,V\hat{x}) - \bar{\mathbf{f}}_y(t_m,x))V|| \\ &\leq \tau_m||V^T||||N(P^TN)^{-1}P^T f_y(t_m,x_0)f_y(t_m,x_0)^{-1}\cdot \\ &(f_y(t_m,V\hat{x}) - f_y(t_m,x))||||V|| \\ &\leq \tau_m\gamma_f C_D||f_y(t_m,x_0)||||x - V\hat{x}||. \end{aligned} \tag{4.84}$$

Similar to the proof of Theorem (4.7), we combine (4.83) and (4.84), then apply the Banach perturbation Lemma to prove (4.81). □

The existence and uniqueness of Newton iterates at each time step for the POD-DEIM based **ROM** is provided in the following theorem.

Theorem 4.11. *Assume that for $F_d : D \subset \mathbb{R}^n \to \mathbb{R}^n$ the assumptions of Theorem (4.5) are satisfied and that condition (4.60) in Theorem4.7 holds true for $x = x_0$. Then, for non-linear map $\hat{F}_d$ the following hold true:*

$$||\hat{F}_d'(\hat{x}_0)^{-1}\hat{F}_d(\hat{x}_0)|| \le \widetilde{\alpha}_m, \tag{4.85}$$

$$||\hat{F}_d'(\hat{x}_0)^{-1}(\hat{F}_d'(\mathbf{y}_1) - \hat{F}_d'(\mathbf{y}_2))|| \le \widetilde{\gamma}_m||\mathbf{y}_1 - \mathbf{y}_2||, \quad \mathbf{y}_1, \mathbf{y}_2 \in \hat{D}, \tag{4.86}$$

where the constants are given by:

$$\widetilde{\alpha}_m = \frac{1}{1-q}||V^T F_d'(x_0)||||V^T F_d'(x_0)^{-1}V||(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}(\alpha_m + \epsilon_m), \tag{4.87}$$

$$\widetilde{\gamma}_m = \frac{1}{1-q}||V^T F_d'(x_0)||||V^T F_d'(x_0)^{-1}V||(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}(\gamma_m + \hat{\epsilon}_m), \tag{4.88}$$

where ϵ_m, $\hat{\epsilon}_m$ are given by:

$$\epsilon_m = ||V^T F_d'(x_0)||^{-1}\Big(||V^T(g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m))|| + \tau_m C_D||(I_N - NN^T)\mathbf{f}(t_m, x_0)||(||V^T(\mathbf{M} + \tau_m\mathbf{A})|| + \tau_m C_D L_f)||x_0 - V\hat{x}_0||\Big), \tag{4.89}$$

$$\hat{\epsilon}_m = \tau_m\gamma_f C_D||V^T F_d'(x_0)||^{-1}||(I_N - NN^T)f_y(t_m, x_0)||. \tag{4.90}$$

Similar to Theorem (4.5), if we further assume:

$$\hat{h}_m = \widetilde{\alpha}_m\widetilde{\gamma}_m < \frac{1}{2}; \quad \bar{\hat{B}}(\hat{x}_0, \hat{\rho}_m) \subset \hat{D}, \tag{4.91}$$

where $\hat{\rho}_m = \frac{1-\sqrt{1-2\hat{h}_m}}{\widetilde{\gamma}_m}$.

*Then, for the sequence $\{\hat{x}^k\}_{k\in\mathbb{N}}$ of the POD-DEIM based **ROM** Newton iterates, the following hold true:*

(i) $\hat{F}_d'(x)$ is invertible for all Newton iterates $\hat{x}^k$, $k \in \mathbb{N}$.

(ii) The sequence $\{\hat{x}^k\}_{k\in\mathbb{N}}$ is well defined with $\hat{x}^k \in \overline{B(\hat{x}_0, \hat{\rho}_m)}$ and $\hat{x}^k \xrightarrow{k\to\infty} \hat{x}^*$ quadratically, where $\hat{F}_d(\hat{x}^*) = 0$.

(iii) $\hat{x}^*$ *is unique in* $\overline{B(\hat{x}_0, \hat{\rho}_m)} \bigcup (D \bigcap B(\hat{x}_0, \tilde{\rho}_m))$, *where* $\tilde{\rho}_m = \frac{1+\sqrt{1-2\hat{h}_m}}{\hat{\gamma}_m}$.

Proof. Similar to the proof of Theorem (4.9), we can estimate $\hat{F}_d'(\hat{x}_0)^{-1}\hat{F}_d(\hat{x}_0)$ as follows:

$$\hat{F}_d'(\hat{x}_0)^{-1}\hat{F}_d(\hat{x}_0) = \hat{F}_d(\hat{x}_0)^{-1}V^T\Big((\mathbf{M}+\tau_m\mathbf{A})V\hat{x}_0 + \tau_m N(P^TN)^{-1}P^T\mathbf{f}(t_m, V\hat{x}_0)$$
$$- g(t_m, V\hat{\mathbf{y}}_m)\Big) \tag{4.92}$$
$$= \hat{F}_d'(\hat{x}_0)^{-1}V^T\Big(\Big(\underbrace{(\mathbf{M}+\tau_m\mathbf{A})x_0 + \tau_m\mathbf{f}(t_m, x_0) - g(t_m, \mathbf{y}_m) + g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m)}_{\text{term (1)}}\Big)$$
$$- \Big(\underbrace{(\mathbf{M}+\tau_m\mathbf{A})(x_0 - V\hat{x}_0) + \tau_m(\mathbf{f}(t_m, x_0) - (P^TN)^{-1}P^T\mathbf{f}(t_m, x_0))}_{\text{term (2)}}\Big)$$
$$- \underbrace{\tau_m N(P^TN)^{-1}P^T(\mathbf{f}(t_m, x_0) - \mathbf{f}(t_m, V\hat{x_0}))}_{\text{term (3)}}\Big).$$

Now we estimate the term (1)-term (3) separately.

By the conditions in Theorem (4.7) and (4.81) in Theorem (4.10), the term (1) can be estimated as follows:

$$||\hat{F}_d'(\hat{x}_0)^{-1}V^T\Big((\mathbf{M}+\tau_m\mathbf{A})x_0 + \tau_m\mathbf{f}(t_m, x_0) - g(t_m, \mathbf{y}_m)\Big) + g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m)||$$
$$= ||\hat{F}_d'(\hat{x}_0)^{-1}V^T(F_d'(x_0) + g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m))||$$
$$= ||\hat{F}_d'(\hat{x}_0)^{-1}V^T(F_d'(x_0)F_d'(x_0)^{-1}F_d'(x_0) + g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m))||$$
$$\leq ||\hat{F}_d'(\hat{x}_0)^{-1}||\Big(\alpha_m||V^TF_d'(x_0)|| + ||V^T(g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m))||\Big)$$
$$\leq \frac{1}{1-q}||V^TF_d'(x_0)^{-1}V||||V^TF_d'(x_0)||^{-1}(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}\Big(\alpha_m + \tag{4.93}$$
$$||V^TF'(x_0)||^{-1}||V^T(g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m))||\Big).$$

Similar to the term (1), the term (2) can be estimated as follows:

$$
\begin{aligned}
&||\hat{F}_d'(x_0)^{-1}V^T\Big((\mathbf{M}+\tau_m\mathbf{A})(x_0-V\hat{x_0})+\tau_m(\mathbf{f}(t_m,x_0)-N(P^TN)^{-1}P^T\mathbf{f}(t_m,\hat{x}_0))\Big)||\\
&\leq \frac{1}{1-q}||V^TF_d'(x_0)^{-1}V||(1-C_{F'}||x_0-V\hat{x}_0||)^{-1}\cdot\\
&\Big(||V^T\Big((\mathbf{M}+\tau_m\mathbf{A})(x_0-V\hat{x_0})+\tau_m(\mathbf{f}(t_m,x_0)-N(P^TN)^{-1}P^T\mathbf{f}(t_m,\hat{x}_0))||\Big)\Big)\\
&\leq \frac{1}{1-q}||V^TF_d'(x_0)^{-1}V||(1-C_{F'}||x_0-V\hat{x}_0||)^{-1}\cdot\\
&(||V^T(\mathbf{M}+\tau_m\mathbf{A})||||x_0-V\hat{x}_0||+\tau_mC_D||V^T||||(I_N-NN^T)\mathbf{f}(t_m,x_0)||.
\end{aligned}
\tag{4.94}
$$

Under the assumption that condition (4.60) in Theorem 4.7 holds true, we can use Lemma (3.4) in Section (3.2) to estimate the term (3):

$$
\begin{aligned}
&||\hat{F}_d'(\hat{x_0})^{-1}V^T(\tau_mN(P^TN)^{-1}P^T(\mathbf{f}(t_m,x_0)-\mathbf{f}(t_m,V\hat{x}_0)))||\leq\tau_m||\hat{F}_d'(\hat{x_0})^{-1}||\\
&||N(P^TN)^{-1}P^T||L_f||x_0-V\hat{x}_0||\\
&\leq\frac{1}{1-q}\tau_mC_DL_f||V^T\hat{F}'(x_0)^{-1}V||||V^T||(1-C_{F'}||x_0-V\hat{x}_0||)||x_0-V\hat{x}_0||.
\end{aligned}
\tag{4.95}
$$

Combining the results (4.93), (4.94), (4.95), we obtain:

$$
\begin{aligned}
&||\hat{F}_d'(\hat{x}_0)^{-1}\hat{F}_d'\hat{x}_0)||\leq\frac{1}{1-q}||V^TF'(x_0)^{-1}V||(1-C_{F'}||x_0-V\hat{x}_0||)\cdot\\
&\Big(\alpha_m||V^TF'(x_0)||+(||V^T(\mathbf{M}+\tau_m\mathbf{A})||+\tau_mC_DL_f||V^T||)||x_0-V\hat{x}_0||\\
&+||V^T(g(t_m,\mathbf{y}_m)-g(t_m,V\hat{\mathbf{y}}_m)||+\tau_mC_D||(I_N-NN^T)\mathbf{f}(t_m,x_0)||\Big),
\end{aligned}
$$

which proves (4.85).

(4.86) can be proved in a similar way:

$$
\begin{aligned}
&\hat{F}_d'(\hat{x}_0)^{-1}(\hat{F}_d'(\mathbf{y}_1) - \hat{F}_d'(\mathbf{y}_2)) = \hat{F}_d'(\hat{x}_0)^{-1}\Big(V^T(\mathbf{M} + \tau_m \mathbf{A})V(\mathbf{y}_1 - \mathbf{y}_2) + \\
&\quad \tau_m V^T N(P^T N)^{-1} P^T (f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2))V\Big) \\
&\quad = \hat{F}_d'(\hat{x}_0)^{-1}\Big(V^T(\mathbf{M} + \tau_m \mathbf{A})V(\mathbf{y}_1 - \mathbf{y}_2)V + \tau_m V^T (f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2))V - \\
&\quad \tau_m V^T \Big((f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2)) - N(P^T N)^{-1} P^T (f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2))\Big)V\Big) \\
&\quad = \hat{F}_d'(\hat{x}_0)^{-1}\Big(\underbrace{V^T(F'(V\mathbf{y}_1) - F'(V\mathbf{y}_2))V}_{\text{term } (a)} \qquad (4.96) \\
&\quad - \underbrace{\tau_m V^T (I_N - N(P^T N)^{-1} P^T)(f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2))V}_{\text{term } (b)}\Big).
\end{aligned}
$$

We estimate the terms (a) and (b) separately.

Assume that $\hat{x}^0 \in \hat{D}$ satisfies the condition in Theorem (4.7) such that $||x^0 - V\hat{x}^0|| \leq \frac{1}{C_{F'}}$. Then from (4.81) in Theorem (4.10), term (a) can be estimated as follows:

$$
\begin{aligned}
&||\hat{F}_d'(\hat{x}_0)^{-1}\Big(V^T(F'(V\mathbf{y}_1) - F'(V\mathbf{y}_2))V\Big)|| \\
&= ||\hat{F}_d'(\hat{x}_0)^{-1}(V^T F'(x_0) F'(x_0)^{-1}(F'(V\mathbf{y}_1) - F'(V\mathbf{y}_2))|| \\
&\leq \gamma_m ||\hat{F}_d'(\hat{x}_0)^{-1}|| ||V^T F'(x_0)|| ||V\mathbf{y}_1 - V\mathbf{y}_2|| \\
&\leq \frac{\gamma_m}{1-q} ||V^T F_d'(x_0)^{-1} V|| ||V^T F'(x_0)|| (1 - C_{F'} ||x_0 - V\hat{x}_0||)^{-1} ||\mathbf{y}_1 - \mathbf{y}_2||. \qquad (4.97)
\end{aligned}
$$

Similarly, the term (b) can be estimated by using condition (4.60) in Theorem4.7 and

Lemma (3.4) in Section (3.2), as shown below:

$$\begin{aligned}
&||\tau_m \hat{F}_d'(\hat{x}_0)^{-1}V^T(I_N - N(P^TN)^{-1}P^T)(f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2))V|| \\
&\leq \tau_m||\hat{F}_d'(\hat{x}_0)^{-1}||||(I_N - N(P^TN)^{-1}P^T)(f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2))|| \\
&\leq \frac{\tau_m}{1-q}||V^TF_d'(x_0)^{-1}V||(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}\cdot \\
&||(I_N - N(P^TN)^{-1}P^T)f_y(t_m, x_0)f_y(t_m, x_0)^{-1}(f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2))|| \\
&\leq \frac{\tau_m}{1-q}C_D||V^TF_d'(x_0)^{-1}V||(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}\cdot \\
&||(I_N - NN^T)f_y(t_m, x_0)||||f_y(t_m, x_0)^{-1}(f_y(t_m, V\mathbf{y}_1) - f_y(t_m, V\mathbf{y}_2))|| \\
&\leq \frac{\tau_m}{1-q}C_D\gamma_f||V^TF_d'(x_0)^{-1}V||(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}\cdot \\
&||(I_N - NN^T)f_y(t_m, x_0)||||V\mathbf{y}_1 - V\mathbf{y}_2|| \\
&\leq \frac{\tau_m}{1-q}C_D\gamma_f||V^TF_d'(x_0)^{-1}V||(1 - C_{F'}||x_0 - V\hat{x}_0||)^{-1}\cdot \qquad (4.98) \\
&||(I_N - NN^T)f_y(t_m, x_0)||||\mathbf{y}_1 - \mathbf{y}_2||.
\end{aligned}$$

Combining (4.97) and (4.98), we obtain (4.86). □

4.3 Error equilibration for the POD and the POD-DEIM based ROM

So far, we have examined the existence and uniqueness of the Newton iterates for both the **FOM** and the **ROM** . Now, we are in a position to show that the error equilibration of the **FOM** is inherited by its **ROM**. That is to say, if $0 = t_0 < t_1 < \cdots < t_M = T$ is the time partitioning such that the error estimators for **FOM** satisfy:

$$e^m \approx e_{m'},\ 1 \leq m \neq m' \leq M,$$

then the error estimators for **ROM** e^m_{rom}'s are also of the same order of magnitude, provided certain conditions are satisfied.

$$e^m_{rom} \approx e^{m'}_{rom},\ 1 \le m \ne m' \le M.$$

We consider the POD based **ROM** first and assume that $F(x)$ satisfies the assumptions in Theorem (4.5) and (4.7) so that Newton's method generates a unique solution at each time step.

We recall that at each time instance t_m,

$$e^m = \|\eta^m\| = \|\hat{\mathbf{y}}^m - \mathbf{y}^m\|,$$

and η^m is solved by Newton's method from the following equation:

$$F(x) = (\mathbf{M} + \tau_m \mathbf{A})x - \tau_m \mathbf{f}(t_m, x) - g(t_m) = 0,\ x \in D,$$

with initial value $x^0 = 0$.

On the other hand, the error estimator for the POD based **ROM** reads:

$$e^m_{rom} = \|\eta^m_{rom}\| = \|\hat{\mathbf{y}}^m_{rom} - \mathbf{y}^m_{rom}\|,$$

and η^m_{rom} is solved by Newton's method from:

$$\hat{F}(\hat{x}) = (\mathbf{M}_{rom} + \tau_m \mathbf{A}_{rom})\hat{x} - \tau_m \hat{\mathbf{f}}(t_m, \hat{x}) - g_{rom}(t_m) = 0,\ \hat{x} \in \hat{D},$$

with initial value $\hat{x}^0 = 0$.

From (3.5) in Section (3.1), we know that there exists a constant $C > 0$ such that:

$$\begin{aligned} ||\eta^m - V\eta^m_{rom}|| &= ||(\mathbf{y}^m - \hat{\mathbf{y}}^m) - V(\mathbf{y}^m_{rom} - \hat{\mathbf{y}}^m_{rom})|| \\ &\le ||(\mathbf{y}^m - V\mathbf{y}^m_{rom})|| + ||\hat{\mathbf{y}}^m - V\hat{\mathbf{y}}^m_{rom}|| \le C\tau_{pod}, \end{aligned} \tag{4.99}$$

where $\tau_{pod} = (\sum_{i=l+1}^{n} \lambda_i)^{\frac{1}{2}}$ is the POD state error.

Therefore, the equilibration of error is preserved by the **ROM** if the POD state error τ_{pod} is sufficient small.

Theorem 4.12. *Assume that $F(x)$ satisfies the assumptions in Theorem (4.5) and (4.7) for initial value $x = 0$. Let us further assume that: $\delta^m = C\tau_{pod}(e^m)^{-1} \ll 1$ for $1 \leq m \leq M$. Then, let $e^m = e(1-\epsilon^m)$, $0 < \epsilon^m \ll 1$ and $r_k = max(\hat{\rho}_m, \rho_m | \ 1 \leq m \leq M)$, where ρ_m, $\hat{\rho}_m$ are the Kantorovich radii for the **FOM** and the POD based **ROM** respectively, the following is true:*

$$(r_k)^{-1}(1-\epsilon^m)(1-\delta^m)e \leq \frac{e^m_{rom}}{e^{m-1}_{rom}} \leq r_k(1-\epsilon^{m-1})^{-1}(1-\delta^{m-1})^{-1}e^{-1}. \tag{4.100}$$

Proof.

$$||\eta^m - V\eta^m_{rom}|| \geq ||\eta^m|| - ||V\eta^m_{rom}|| \geq ||\eta^m|| - ||\eta^m_{rom}|| = e^m - e^m_{rom}$$

From (4.99), we can get:

$$e^m_{rom} \geq e^m - ||\eta^m - \eta^m_{rom}|| \geq e^m(1 - C\frac{\tau_{pod}}{e^m}).$$

By Theorem (4.5) and (4.9), we have:

$$e^m_{rom} \leq \hat{\rho}_m, \quad 1 \leq m \leq M.$$

Hence,

$$\frac{e^m_{rom}}{e^{m-1}_{rom}} \geq \frac{e^m(1 - C\tau_{pod}(e^m)^{-1})}{\hat{\rho}^{m-1}} \geq \frac{e(1-\epsilon^m)(1 - C\tau_{pod}(e^m)^{-1})}{r_k} \tag{4.101}$$

and

$$\frac{e^{m-1}_{rom}}{e^m_{rom}} \geq \frac{e^{m-1}(1 - C\tau_{pod}(e^{m-1})^{-1})}{\hat{\rho}_m} \geq \frac{e(1-\epsilon^{m-1})(1 - C\tau_{pod}(e^{m-1})^{-1})}{r_k}. \tag{4.102}$$

The inequality (4.100) follows directly from (4.101) and (4.102). □

Remark 4.13. For the constant e in Theorem (4.12), we can take the average value of the error estimators, i.e., $e = \frac{1}{M}\sum_{m=1}^{M} e^m$. If the error equilibration is satisfied for $\{e^m\}$, $1 \leq m \leq M$, then $\epsilon^m \ll 1$ holds true.

A similar result can be derived for the POD-DEIM based **ROM**. We first apply Theorem (4.11) to $\hat{F}_d(x)$ with initial value $x = 0$ in the following corollary:

Corollary 4.14. *Assume that $F_d(x)$ satisfies the assumptions of Theorem (4.5) and (4.7) for $x = 0$. Then, for POD-DEIM non-linear map $\hat{F}_d(x)$, the following hold true:*

$$
\begin{aligned}
&||\hat{F}'_d(0)^{-1}\hat{F}_d(0)|| \leq \hat{\alpha}_m, \\
&||\hat{F}'_d(0)^{-1}(\hat{F}'_d(\mathbf{y}_1) - \hat{F}'_d(\mathbf{y}_2))|| \leq \gamma||\mathbf{y}_1 - \mathbf{y}_2||, \quad \mathbf{y}_1, \mathbf{y}_2 \in \hat{D},
\end{aligned}
$$

where the constants are given by:

$$\hat{\alpha}_m = \alpha(1 + \epsilon_m^{(1)}); \quad \hat{\gamma}_m == \gamma(1 + \epsilon_m^{(2)});$$

where, α, γ, $\epsilon_m^{(1)}$, $\epsilon_m^{(2)}$ are given by:

$$
\begin{aligned}
\alpha &= \frac{1}{1-q}||V^T F'_d(0)||||V^T F'_d(0)^{-1}V||\alpha_m, \\
\gamma &= \frac{1}{1-q}||V^T F'_d(0)||||V^T F'_d(0)^{-1}V||\gamma_m, \\
\epsilon_m^{(1)} &= \frac{1}{\alpha(1-q)}||V^T F'_d(0)^{-1}V||||V^T(g(t_m, \mathbf{y}_m) - g(t_m, V\hat{\mathbf{y}}_m))||, \\
\epsilon_m^{(2)} &= \frac{\tau_m \gamma_f C_D}{\gamma(1-q)}||V^T F'_d(0)^{-1}V||||(I_N - NN^T)f_y(t_m, 0)||.
\end{aligned}
$$

Let us further assume:

$$\hat{h}_m = \hat{\alpha}_m \hat{\gamma}_m < \frac{1}{2}; \quad \bar{\hat{B}}(\hat{x}_0, \hat{\rho}_m) \subset \hat{D},$$

where $\hat{\rho}_m = \frac{1-\sqrt{1-2\hat{h}_m}}{\hat{\gamma}_m}$.

*Then, the sequence $\{\hat{x}^k\}$ of the POD-DEIM based **ROM** Newton iterates is well defined with $\{\hat{x}^k\} \in \overline{B(0,\hat{\rho}_m)}$ and $\hat{x}^k \xrightarrow{k\to\infty} \hat{x}^*$ quadratically, where $\hat{F}_d(\hat{x}^*) = 0$. $\hat{x}^*$ is unique in $\overline{B(0,\hat{\rho}_m)} \bigcup (D \bigcap B(0,\tilde{\rho}_m))$, where $\tilde{\rho}_m = \frac{1+\sqrt{1-2\hat{h}_m}}{\hat{\gamma}_m}$.*

Moreover, if we let $\hat{h} = \alpha\gamma$, $\hat{\rho} = \frac{1-\sqrt{1-2\hat{h}}}{\gamma}$, then the following hold true:

$$(1+\epsilon_m^{(1)})\hat{\rho} \le \hat{\rho}_m \le 2(1+\epsilon_m^{(1)})\hat{\rho}. \tag{4.103}$$

Proof. On one hand, as $0 < \alpha\gamma \le \hat{\alpha}_m\hat{\gamma}_m$:

$$\begin{aligned}\frac{\hat{\rho}_m}{\hat{\rho}} &= \frac{\gamma}{\hat{\gamma}_m}\frac{1-\sqrt{1-2\hat{\alpha}_m\hat{\gamma}_m}}{1-\sqrt{1-2\alpha\gamma}} = \frac{\gamma}{\hat{\gamma}_m}\frac{(1-\sqrt{1-2\hat{\alpha}_m\hat{\gamma}_m})(1+\sqrt{1-2\alpha\gamma})}{2\alpha\gamma}\\ &\ge \frac{\gamma}{\hat{\gamma}_m}\frac{2\hat{\alpha}_m\hat{\gamma}_m}{2\alpha\gamma} = 1+\epsilon_m^{(1)}.\end{aligned}$$

On the other hand, as $\hat{\alpha}_m\hat{\gamma}_m < \frac{1}{2} < 1$, we have:

$$\frac{\hat{\rho}_m}{\hat{\rho}} \le \frac{\gamma}{\hat{\gamma}_m}\frac{1-(1-2\hat{\alpha}_m\hat{\gamma}_m)}{1-\sqrt{1-2\alpha\gamma}} = \frac{2\hat{\alpha}_m\gamma(1+\sqrt{1-2\alpha\gamma})}{2\alpha\gamma} \le 2(1+\epsilon_m^{(1)}).$$

The proof for the rest of the results follows directly from the proof for Theorem (4.11). □

From (3.4) in Section (3.2) we can deduce that there exists a constant $C_D > 0$ such that:

$$\begin{aligned}||\eta^m - V\eta^m_{rom}|| &= ||(\mathbf{y}^m - \hat{\mathbf{y}}^m) - V(\mathbf{y}^m_{rom} - \hat{\mathbf{y}}^m_{rom})||\\ &\le ||(\mathbf{y}^m - \mathbf{y}^m_{rom})|| + ||\hat{\mathbf{y}}^m - \hat{\mathbf{y}}^m_{rom}|| \le C_D\tau_D,\end{aligned} \tag{4.104}$$

where $\tau_D = (\sum_{k=l+1}^{n}\lambda_k + \sum_{k=l+1}^{n} s_k)^{\frac{1}{2}}$ is the POD-DEIM state error.

Theorem 4.15. *Assume that $F(x)$ satisfies the assumptions in Theorem (4.5) and (4.7) for initial value $x = 0$. If we further assume that: $\delta^m = C\tau_D(e^m)^{-1} \ll 1$ for $1 \le m \le M$ and let $e^m = e(1-\epsilon^m)$, $0 < \epsilon^m \ll 1$, $r_k = max(\hat{\rho}_m, \hat{\rho}|\ 1 \le m \le M)$. Then, the following is true:*

$$\frac{(1-\epsilon^m)(1-\delta^m)e}{2r_k(1+\epsilon_m^{(1)})} \le \frac{e^m_{rom}}{e^{m-1}_{rom}} \le \frac{2r_k(1+\epsilon_m^{(1)})}{(1-\epsilon^{m-1})(1-\delta^{m-1})e}. \tag{4.105}$$

Proof. By (4.103) in Corollary (4.14), (4.105) can be proved in much the same way as in the proof of Theorem (4.12). □

4.4 Snapshot location based on optimization

In [13], the authors consider the linear case of the evolution equation. They show that instead of taking the FE solution at fixed time instances $Y = [y(t_0, x), \cdots, y(t_M, x)]$ as the POD snapshot matrix, one can use a different POD snapshot matrix Y_{new} by adding some new snapshots $\{y(\bar{t}_1, x), \cdots, y(\bar{t}_{\bar{k}}, x)\}$ to the existing snapshot matrix Y. Then, POD can be performed to the new snapshot matrix Y_{new}. The elements of $\{y(\bar{t}_1, x), \cdots, y(\bar{t}_{\bar{k}}, x)\}$ can be the same as the elements of Y, that is to say, if $y(\bar{t}_k, x) = y(t_i, x)$, where $1 \leq k \leq \hat{k}$, $1 \leq i \leq M$, then we double $y(t_i, x)$ in the snapshot matrix Y, otherwise the new snapshot is added to Y according to its time instance $\bar{t}_k, 1 \leq k \leq \hat{k}$ in order to construct the new snapshot Y_{new}.

With the new snapshot matrix Y_{new}, the authors proceed by solving the optimal location of the additional snapshots, i.e., solving for the time instances $\bar{t}_k,\ 1 \leq k \leq \bar{k}$ such that the following cost function is minimized:

$$J(\mathbf{y}^l, \bar{t}, \phi) = \int_0^T \|\mathbf{y}(t, x) - \sum_{i=1}^{l} \mathbf{y}^l(t_i, x)\phi_i\| \; dt, \tag{4.106}$$

where $\mathbf{y}(t, x)$ and $\mathbf{y}^l(t, x)$ are the FE solution of the **FOM** and the POD based **ROM** based on the new snapshot matrix Y_{new}, respectively. $\mathbf{y}(t, x) = [y(t_1, x), \cdots, y(t_l, x)]^T$, $\mathbf{y}^l = [y^l(t_1, x), \cdots, y^l(t_l, x)]^T$, $\bar{t} = [\bar{t}_1, \cdots, \bar{t}_{\bar{k}}]^T$ and $\phi = [\phi_1, \cdots, \phi_l]$ is the matrix whose columns are POD basis functions.

As we have shown in section 3.1 that the POD basis functions are the solutions of an eigenvalue problem, the authors suggest that we can define the self-adjoint operator **R** for computing the POD basis functions based on the new snapshot matrix as follows:

$$\mathbf{R}(\bar{t})\phi = \lambda\phi, \tag{4.107a}$$

$$\mathbf{R}(\bar{t})\phi = \sum_{j=0}^{M} < y(t_j, x), \phi > y(t_j, x) + \sum_{k=1}^{\bar{k}} < y(\bar{t}_k, x), \phi > y(\bar{t}_k, x). \tag{4.107b}$$

On the other hand, as $\mathbf{y}^l(t_i) = \mathbf{y}^l(t_i, x)$'s are the FE solutions to the POD based **ROM**, so we also have the following constrains:

$$\frac{d\mathbf{y}^l(t)}{dt} + \mathbf{A}^l\mathbf{y}^l(t) = \mathbf{f}^l(t), \quad 0 \leq t \leq T, \tag{4.108a}$$

$$\mathbf{y}^l(0) = \mathbf{y}_0^l. \tag{4.108b}$$

Where the vectors $\mathbf{y}_0^l$ and $f^l(t)$ are given by:

$$\vec{\psi}_0 = \begin{pmatrix} < y_0, \phi_1 > \\ \vdots \\ < y_0, \phi_l > \end{pmatrix}, \quad f^l(t) = \begin{pmatrix} < f(t), \phi_1 > \\ \vdots \\ < f(t), \phi_l > \end{pmatrix}.$$

And the coefficient matrix:

$$\mathbf{A}^l = ((\mathbf{A}_{ij}^l)) \in \mathbb{R}^{l \times l}, \text{with } \mathbf{A}_{ij}^l = < A\phi_j, \phi_i > .$$

Moreover, recall that the POD basis functions are orthonormal and $\bar{t}_k \in (0, T]$, so we also have:

$$\|\phi_i\| = 1, \; 1 \leq i \leq l, \quad 0 < \bar{t}_k \leq T \, 1 \leq i \leq \bar{k}. \tag{4.109}$$

With constrains (4.107a), (4.107b), (4.108a), (4.108b) and (4.109), the optimal location of the additional snapshots can be formulated as a non-linear minimization problem of the cost function $J(\mathbf{y}^l, \bar{t}, \phi)$ and can be solved by the SQP (Sequential Quadratic Programming)-type algorithm with BFGS updates based on the second order sufficient optimality conditions given in their paper. In our numerical example, a linear evolution equation is also examined and both the error equilibration method and the optimal location method invented by Kunisch/Volkwein are performed for comparison.

Chapter 5

Numerical Results

5.1 Linear parabolic equations

The first example is taken from [13] and features a solution that exhibits a rapid change only in the initial time interval.

Example 1. Let $Q := \Omega \times (0,T)$, where $\Omega = (0,1)^2$, and $\Sigma := \Gamma \times (0,T)$, where $\Gamma := \partial\Omega$. Consider the following parabolic initial-boundary value problem

$$\begin{cases} \dfrac{\partial y}{\partial t} - c\Delta y + \boldsymbol{\beta} \cdot \nabla y + y = f \quad \text{in } Q, \\ c\dfrac{\partial y}{\partial \mathbf{n}} + qy = g \quad \text{on } \Sigma, \\ y(\cdot, 0) = y^0 \quad \text{in } \Omega, \end{cases}$$

where the source terms f, g, the coefficient function q, and the initial data y^0 are chosen according to:

$$f(x,t) = \begin{cases} 4,\ x=(x_1,x_2),\ (x_1-0.25)^2+(x_2-0.65)^2 \le 0.05,\ t\in[0,T], \\ 0,\ \text{otherwise} \end{cases},$$

$$g(x,t) = \begin{cases} 1,\ x=(x_1,1),\ 0<x_1<1,\ t\in[0,T], \\ 0,\ x=(1,x_2),\ 0<x_2<1,\ t\in[0,T], \\ 0,\ x=(0,x_2),\ 0<x_2<1,\ t\in[0,T], \\ -1,\ x=(x_1,0),\ 0<x_1<1,\ t\in[0,T] \end{cases},$$

$$q(x,t) = \begin{cases} 1,\ x=(x_1,1),\ 0<x_1<1,\ t\in[0,T], \\ x_2,\ x=(1,x_2),\ 0<x_2<1,\ t\in[0,T], \\ -2,\ x=(x_1,0),\ 0<x_1<1,\ t\in[0,T], \\ 0,\ x=(0,x_2),\ 0<x_2<1,\ t\in[0,T] \end{cases},$$

$$y^0(x) = \sin(\pi x_1)\cos(\pi x_2),\ x=(x_1,x_2)\in\Omega.$$

Further, we choose $T=1$, $c=0.1$, and $\boldsymbol{\beta}=(0.1,-10)^T$.

We discretize the spatial domain Ω by a uniform grid with mesh size $h=\frac{1}{40}$ and construct a simplicial triangulation by right isosceles. As a result, the degree of freedom for the spatial discretization is 41^2. Further, we utilize piecewise linear FE ansatz functions and solve the discretized system by the implicit Euler method with a fixed time step length $\Delta t=\frac{1}{20}$. The FE solutions $\mathbf{y}(t,x)$ at time instances $t=[0,\ 0.15,\ 0.5,\ 0.75,\ 1]$ are shown in Figure 5.1.

5.1.1 Error equilibration in time

In this numerical example, we consider the error equilibration in time. We take $TOL = 1e-2$ as the tolerance for the error estimators e^m and apply Algorithm 2 introduced in

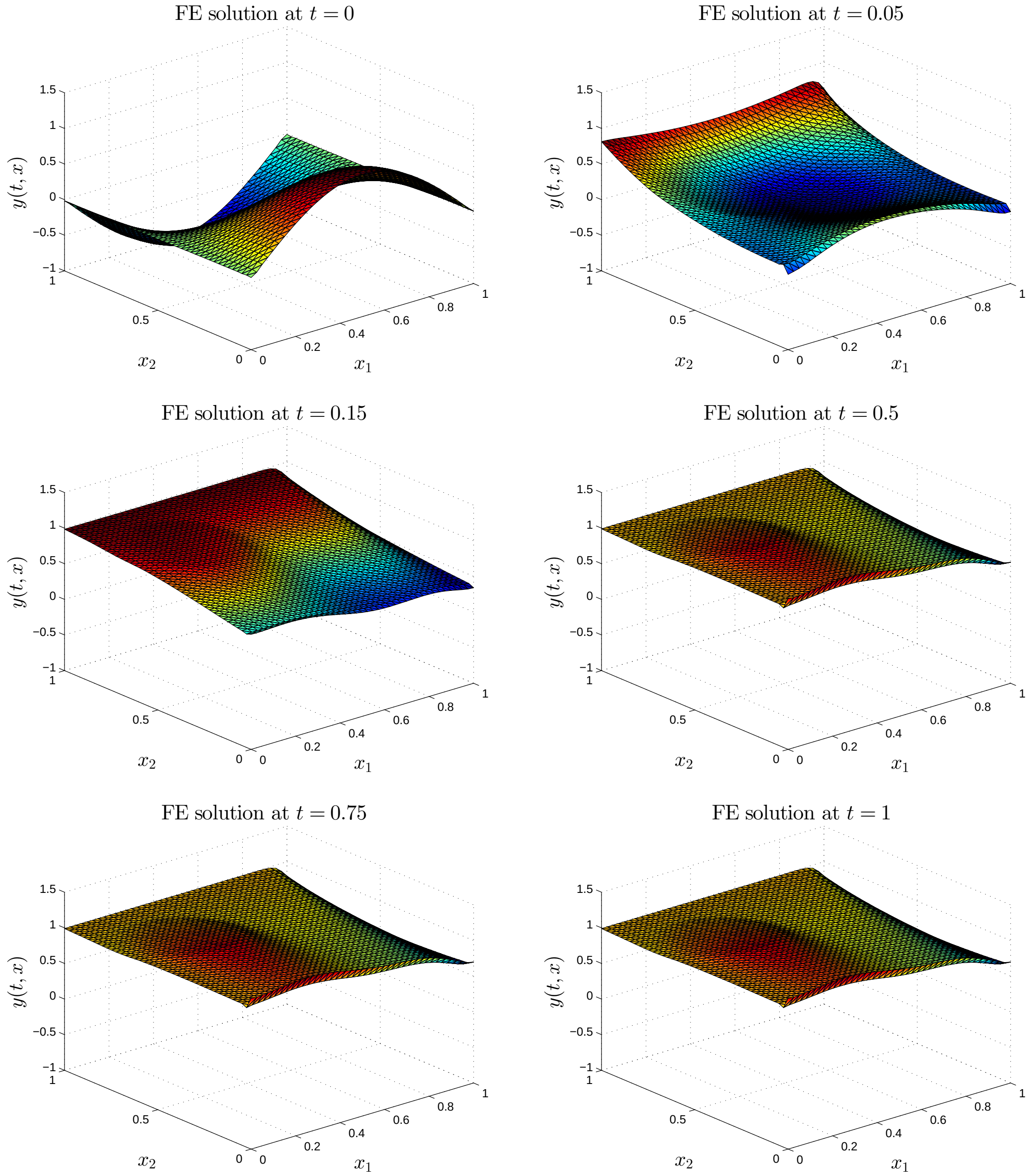

Figure 5.1: Example 1: Implicit Euler FE solutions at different time instances

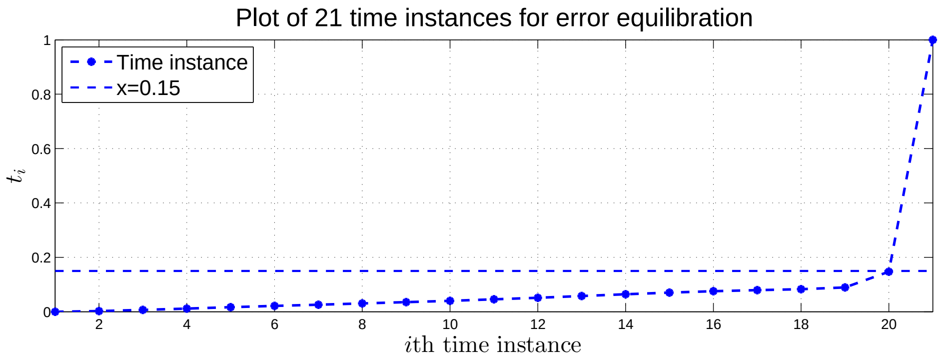

Figure 5.2: Example 1: 21 time instances for the error equilibration for **FOM**

section 4.2 to the **FOM** with initial time partitioning $P_0 = \{t_j = j\Delta t, 0 \leq j \leq M\}$, where $M = \frac{T}{\Delta t}$. We observe that the finite element solution does not change significantly from $t = 0.15$ to $t = 1$. Therefore, if we denote the set of the time instances for the error equilibration by $P_{eq} = \{0 = t_0,\ t_1, \cdots,\ t_{20} = 1\}$, then one can expect to see that most of the elements in P_{eq} are located in the initial sub-interval $[0, 0.15]$, as shown in Figure 5.2.

Now we use the implicit Euler FE solutions to construct the POD snapshot matrix $Y = [\mathbf{y}(0,x), \mathbf{y}(t_1,x), \cdots, \mathbf{y}(t_{20},x)]$, then apply the POD method to the **FOM** with $l = 6,\ 10,\ 14, 21$ POD basis, respectively. The singular values for POD snapshot matrix Y is plotted in Figure 5.3. The first 6 POD basis functions are shown in Figure 5.4.

By Theorem 4.12 in Section 4.3, the error estimators e^m_{rom} for POD-based **ROM** is also expected to be of the same order of magnitude, provided they are computed from the time partitioning that satisfies the error equilibration for the **FOM**. That is to say, if we let the $e^m_{rel}, \hat{e}^m_{rel}$ be the relative errors for **FOM** and **ROM**, respectively, as introduced in (4.31). Then, for sufficiently small POD state error τ_{pod}, we have:

$$\hat{e}^m_{rel} < TOL,\ if\ \ e^m_{rel} < TOL \tag{5.2}$$

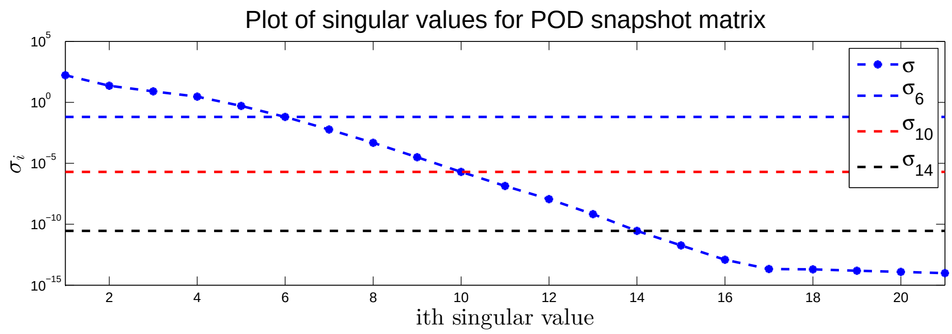

Figure 5.3: Example 1: Singular values $\{\sigma_i\}_{i=1}^{21}$ for the POD snapshot matrix.

We take $l = 6,\ 10,\ 14,\ 21$ as dimensions for the POD-based **ROM** and compute the error estimators for **ROM** accordingly. The behaviour of the error estimators for **ROM** and **FOM** are plotted in Figure 5.5-5.8. A table containing the average and the standard deviation of the error estimator under different POD dimension l is provided in Table 5.1.

POD dim	**FOM**	$l = 6$	$l = 10$
$\mu(\{e^m\}) \pm \sigma(e^m)$	$2.855e-1 \pm 7.333e-4$	$5.53e-1 \pm 3.171e-1$	$2.852e-1 \pm 3.53e-2$
POD dim	**FOM**	$l = 14$	$l = 21$
$\mu(\{e^m\}) \pm \sigma(e^m)$	$2.855e-1 \pm 7.333e-4$	$2.849e-1 \pm 4.6e-3$	$2.855e-1 \pm 1.5e-3$

Table 5.1: Example 1: Mean $(\mu(\{e^m\}))$ $\pm$ standard deviation $(\sigma(e^m))$ of the error estimators under POD dimension l

We can see from Figure 5.5 that for the **ROM** with dimension $l = 6$, e^m_{rom}'s are completely different from e^m's because of the large POD state error τ_{pod} as shown in Figure 5.3. However, as we increase the POD dimension l, the relative error $\hat{e}^m_{rel}$ will decrease. Figure

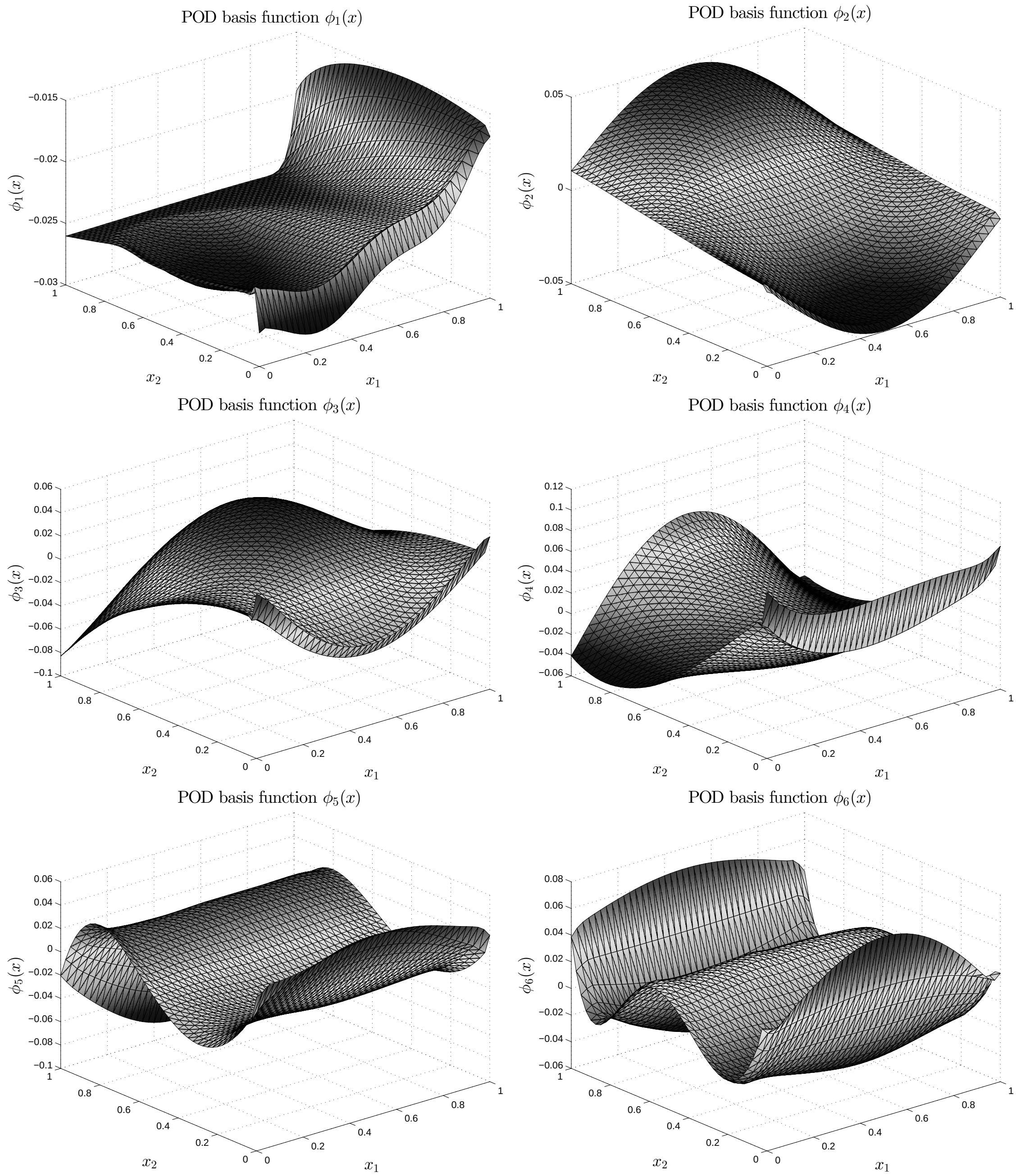

Figure 5.4: Example 1: The first 6 POD basis functions

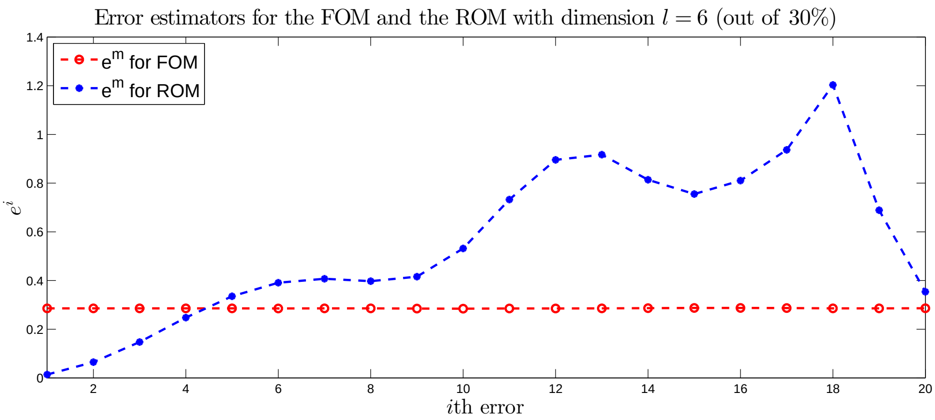

Figure 5.5: Example 1: Error estimators e^m and e^m_{rom} at time instance t_m, $1 \leq m \leq M$ for POD dimension $l = 6$.

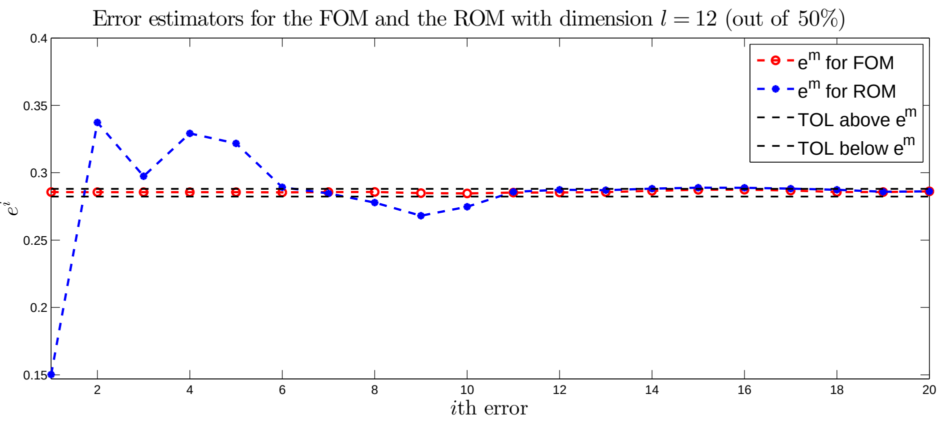

Figure 5.6: Example 1: Error estimators e^m and e^m_{rom} at time instance t_m, $1 \leq m \leq M$ for POD dimension $l = 10$

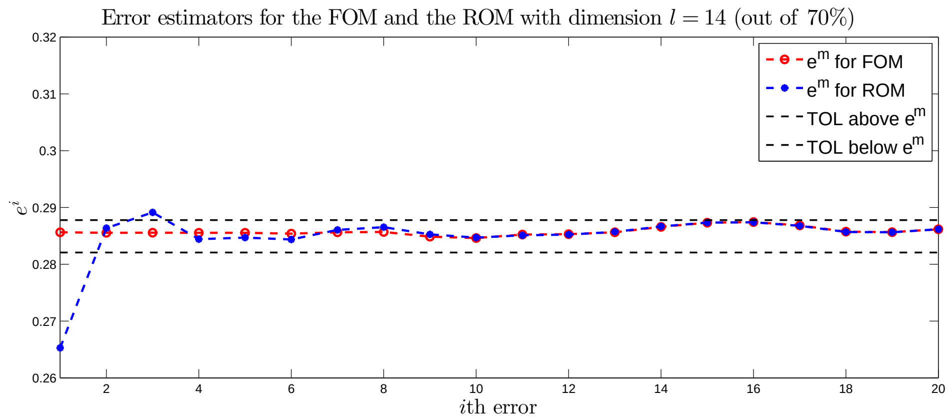

Figure 5.7: Example 1: Error estimators e^m and e^m_{rom} at time instance t_m, $1 \leq m \leq M$ for POD dimension $l = 14$.

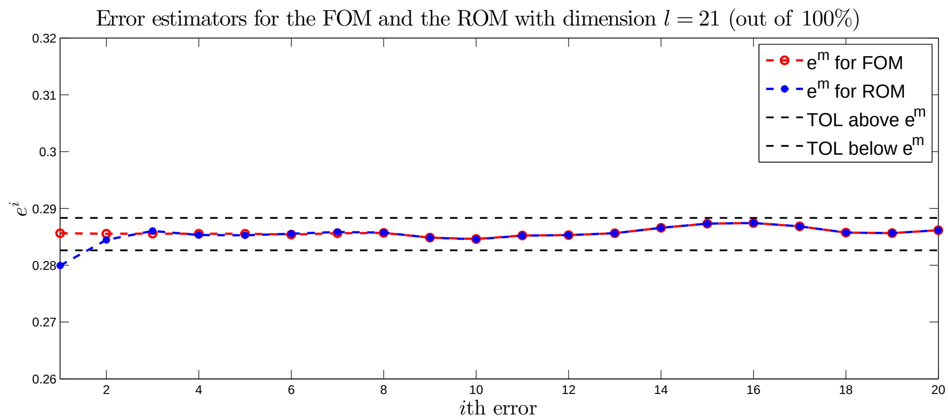

Figure 5.8: Example 1: Error estimators e^m and e^m_{rom} at time instance t_m, $1 \leq m \leq M$ for POD dimension $l = 21$.

5.5 shows that for sufficiently large POD dimension l, $\{e^m\}_{m=1}^M$ and $\{e_{rom}^m\}_{m=1}^M$ are almost the same.

5.1.2 Comparison of two strategies for snapshot location

In this numerical example, we compare the error equilibration method with the approach for computing optimal snapshot locations introduced in Section 4.4. We use the same spatial discretization as in Section **??**, but change the time step to $\Delta t = \frac{1}{14}$, i.e., the initial time partitioning is set to be $P_0 = \{t_j = \frac{j}{14},\ 0 \le j \le 14\}$. We first compute the time partitioning $P_{eq} = \{0 = t_0,\ t_1, \cdots,\ t_{14} = 1\}$ for the error equilibration, the result is plotted in Figure 5.1.2.

Now, we fix the POD dimension to be $l = 3$ and take off $t_1,\ \cdots,\ t_4$ from P_{eq}. We use the rest 11 time instances as the fixed locations for POD snapshots and compute the optimal locations for 4 additional time instances $\bar{\mathbf{t}} = [\bar{t}_1, \bar{t}_2, \bar{t}_3, \bar{t}_4]^T$ by solving the minimization problem that has been introduced in Section (4.4). For initial guess of the additional snapshot locations, we take $\bar{\mathbf{t}}_0 = [0.03, 0.05, 0.08, 0.1]^T$ The optimization is computed by SQP algorithm with BFGS updates using matlab function 'fmincon' according to the following settings:

```
options=optimset('Display','iter','Algorithm','sqp','TolX',...
         1e-9,'TolFun',1e-7,'TolCon',1e-6,'Diagnostics','on',...
         'MaxFunEvals',10^4,'MaxIter',10^4,'GradObj','on',...
         'LargeScale', 'off');
```

The optimal locations are $\bar{\mathbf{t}}^* = [5.21e-003, 8.98e-002, 2.34e-002, 1.74e-002]^T$. $\bar{t}$'s are

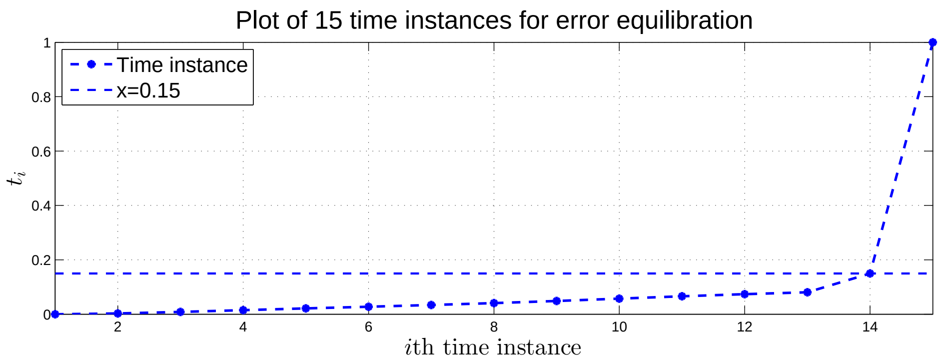

Figure 5.9: Example 1: 15 time instances for the error equilibration for **FOM**.

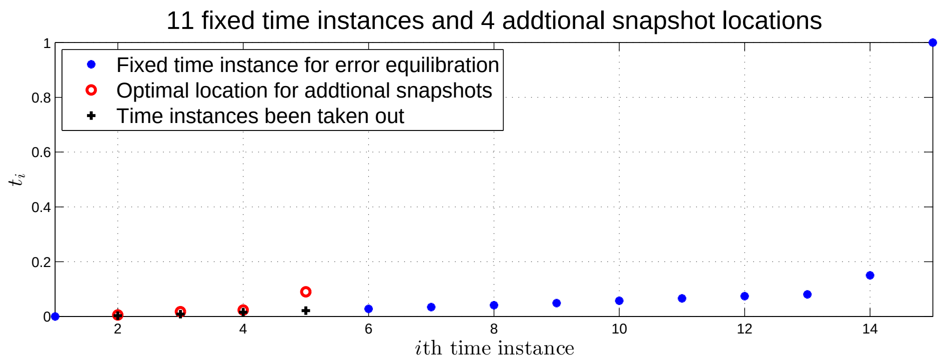

Figure 5.10: Example 1: 11 fixed time instances (blue), the deleted 4 time instances and the optimal locations for 4 additional time instances (red).

plotted together with the fixed snapshot locations in P_{eq} in Figure 5.10. The values of the cost function $J(y, \bar{\mathbf{t}}, \vec{\phi})$ under different time partitioning are also listed in Table 5.2. From Figure 5.10, we can see that the optimal locations of the POD snapshot (Kunisch/Volkwein approach) are close to the locations computed by means of the error equilibration. Table (5.2) shows that the evaluation of the cost function for the optimal snapshot locations (Kunisch/Volkwein approach) is smaller than that of error equilibration in time but there is no significant difference between the two costs.

Time partitioning	P_{eq}	$\bar{t}_0$	$\bar{t}^*$
Cost $J(y, \bar{t}, \vec{\phi})$	1.1629	1.3107	1.1224

Table 5.2: Example 1: Values of J for the time partitioning of error equilibration P_{eq}, initial additional locations $\bar{t}_0$ and optimal locations $\bar{t}^*$

5.2 Non-linear parabolic equations

As examples for non-linear parabolic problems we choose two initial-boundary value problems for semi-linear second order parabolic partial differential equations.

5.2.1 Example 2: Error equilibration in time

Let $Q := \Omega \times (0, T)$, where $\Omega = (0, 1)$ and $\Sigma := \Gamma \times (0, T)$, with $\Gamma := \partial\Omega$. Consider the following semi-linear parabolic initial-boundary value problem:

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + y^2 = f \quad \text{in } Q, & (5.3a) \\ y = 0 \quad \text{on } \Sigma, & (5.3b) \\ y(\cdot, 0) = 0 \quad \text{in } \Omega, & (5.3c) \end{cases}$$

where the right-hand side f is given such that:

$$y_e = x^3(1-x)^3 t^2 (1-t)^2 \arctan(60((x - \frac{5}{4})^2 + (t + \frac{1}{4})^2)^{\frac{1}{2}} - 1)$$

is the solution to 5.3a-5.3c.

It is obvious that $y(t,x)^2$ is Lipschitz continuous if $y(t,x)$ is bounded, i.e., the solution to system ((5.3a))-((5.3c)) exists globally. On the other hand, it has been shown in Theorem 5.2 in [27] that the solution exists globally if the initial value is bounded above by the

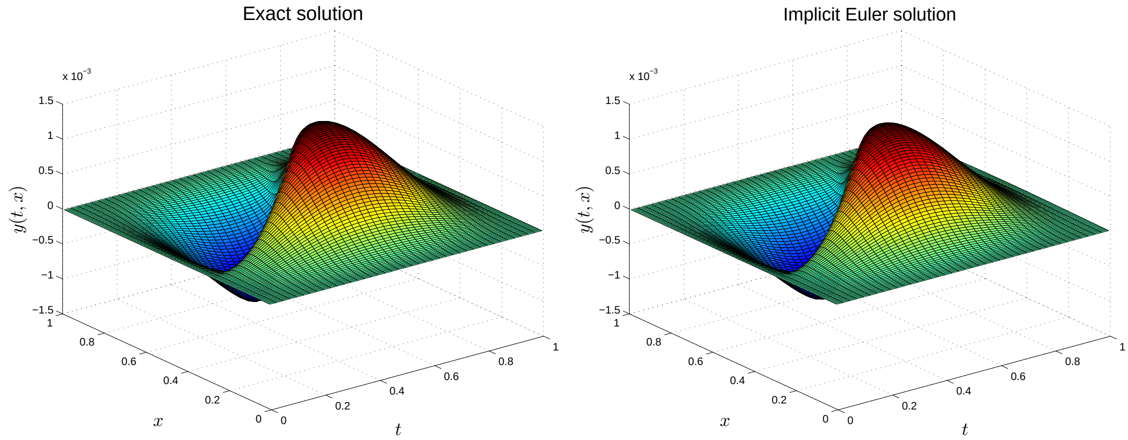

Figure 5.11: Example 2: Exact solution and the implicit Euler FE solution for **FOM** .

first eigenfunction of $-\Delta$. As our initial value is 0, so the solution exists globally and is bounded. Therefore, $y(t,x)^2$ is Lipschitz continuous and conditions in Theorem (4.5) are satisfied. So, the newton's method applies to system ((5.3a))-((5.3c)).

With $T = 1$, we uniformly discretize the interval Ω with mesh size $h = \frac{1}{128}$ and denote the grid points by $\{0 = x_0 \leq x_1 < \cdots < x_{128} = 1\}$. Then we utilize piece-wise linear FE functions and discretize the PDE in time by the implicit Euler method with fixed time step length $\Delta t = \frac{1}{40}$. The exact and the implicit Euler solutions are displayed in Figure 5.11. As in previous example, the tolerance for error equilibration is taken to be $TOL = 1e-2$, then we apply Algorithm 2 from Section (4.2) to the **FOM** with initial time partitioning $P_0 = \{t_j = \frac{j}{40}, 0 \leq j \leq M\}$. Figure 5.11 shows that the finite element solution does not perform any significant change along the time interval $[0,1]$. Therefore, different from Example 1, the time instances for error equilibration for Example 2 are expected to be distributed all over the entire time interval $[0,1]$, as shown in Figure 5.12. We use the implicit Euler FE solutions as the snapshot matrix to compute both the POD and the

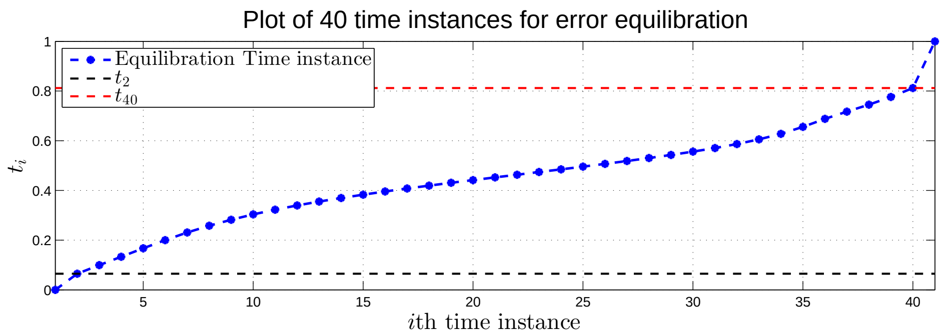

Figure 5.12: Example 2: 40 time instances for error equilibration for non-linear **FOM** .

POD-DEIM basis functions and singular values. To visualize the solution of the **ROM**, we take $l = 3,\ 12$ as the dimensions for both the POD based **ROM** and the POD-DEIM based **ROM**, respectively. The singular values of POD snapshot matrix and DEIM non-linear snapshot matrix are plotted in Figure 5.13 and Figure 5.14. The solutions of the POD and POD-DEIM **ROM** are shown in Figure 5.15 and Figure 5.16. We also provide the first 6 POD basis functions and DEIM basis functions in Figure 5.1 and Figure 5.1.

Figure 5.15 shows that when the POD dimension l is small (3 out of 40), the solution of POD based **ROM** might be inaccurate. This may result in the large difference between the error estimators for the **FOM** and the **ROM**. To examine the behaviour of the error estimators of **ROM**, we take POD dimension $l = 12,\ 20,\ 28,\ 40$ and compute the error estimators for the **ROM** according to the locations given by the error equilibration of the **FOM**. The error estimators are shown in Figure 5.19-5.21.

Figure 5.19-5.21 show that as the dimension l increases, the error estimators for the **ROM** tend to 'converge' to within the interval bounded by the given TOL. The average of the error estimators and their standard deviation under different dimensions l are listed in Table 5.3.

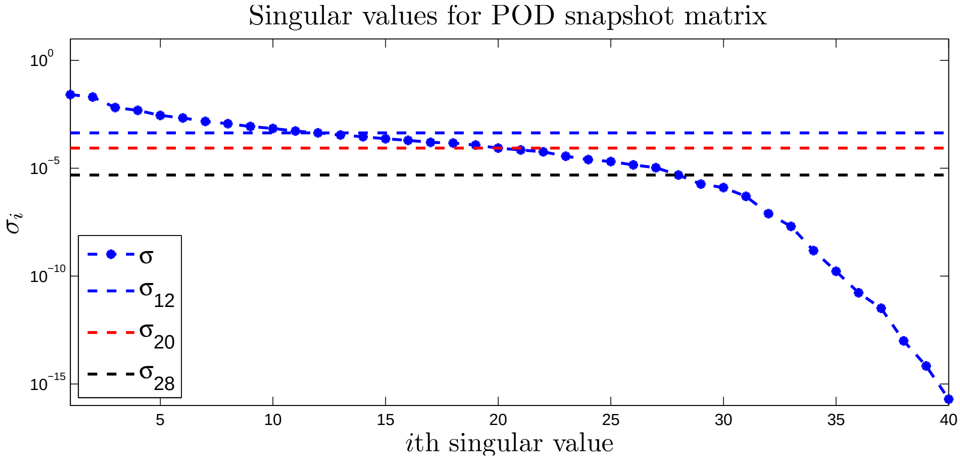

Figure 5.13: Example 2: Singular values $\{\sigma_i\}_{i=1}^{40}$ for POD snapshot matrix.

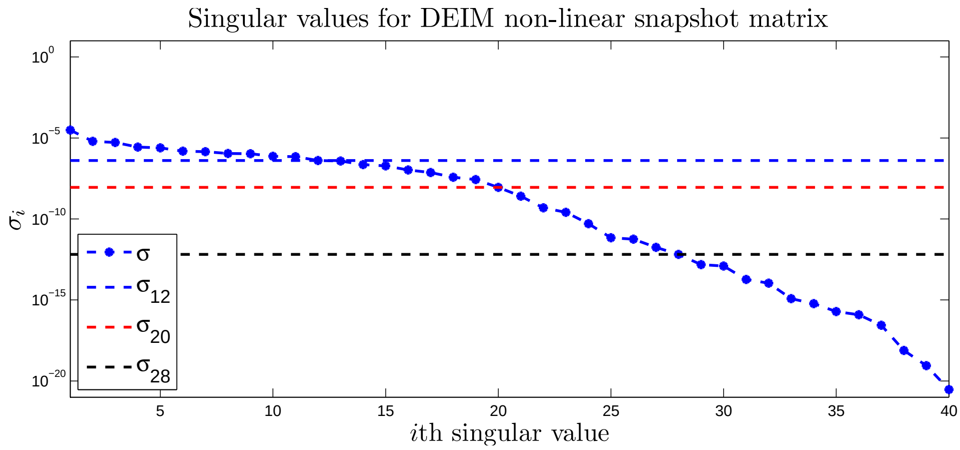

Figure 5.14: Example 2: Singular values $\{\sigma_i\}_{i=1}^{40}$ for DEIM non-linear snapshot matrix.

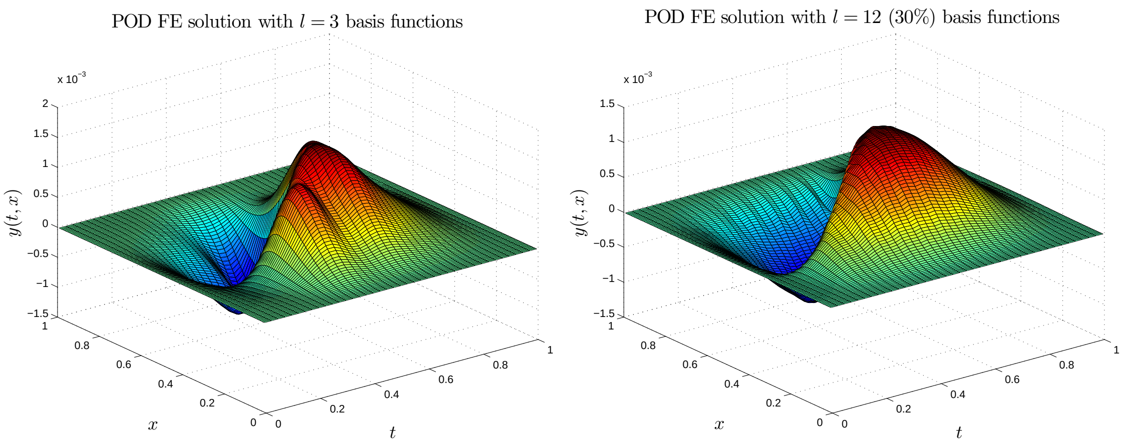

Figure 5.15: Example 2: Implicit Euler FE solution for POD based **ROM** with dimension $l = 3$ (Left), 12 (Right).

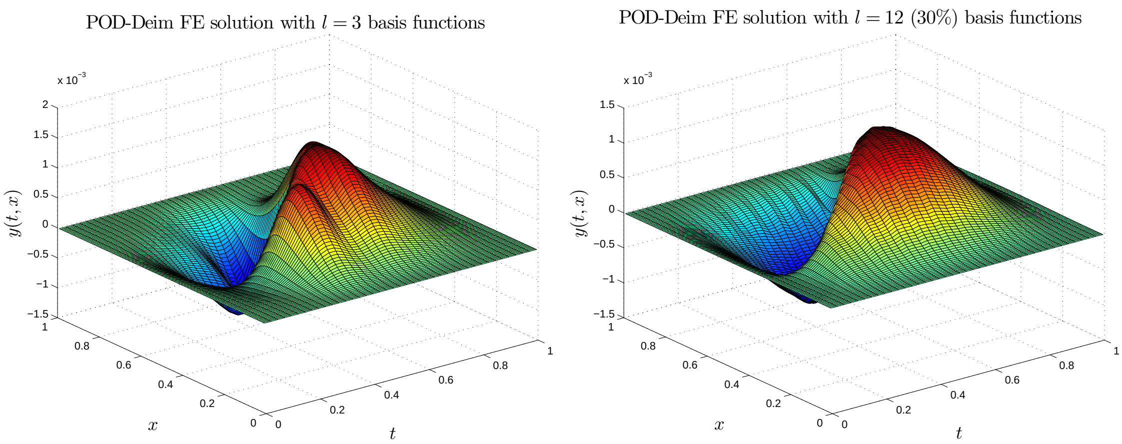

Figure 5.16: Example 2: Implicit Euler FE solution for POD-DEIM based **ROM** with dimension $l = 3$ (Left), 12 (Right).

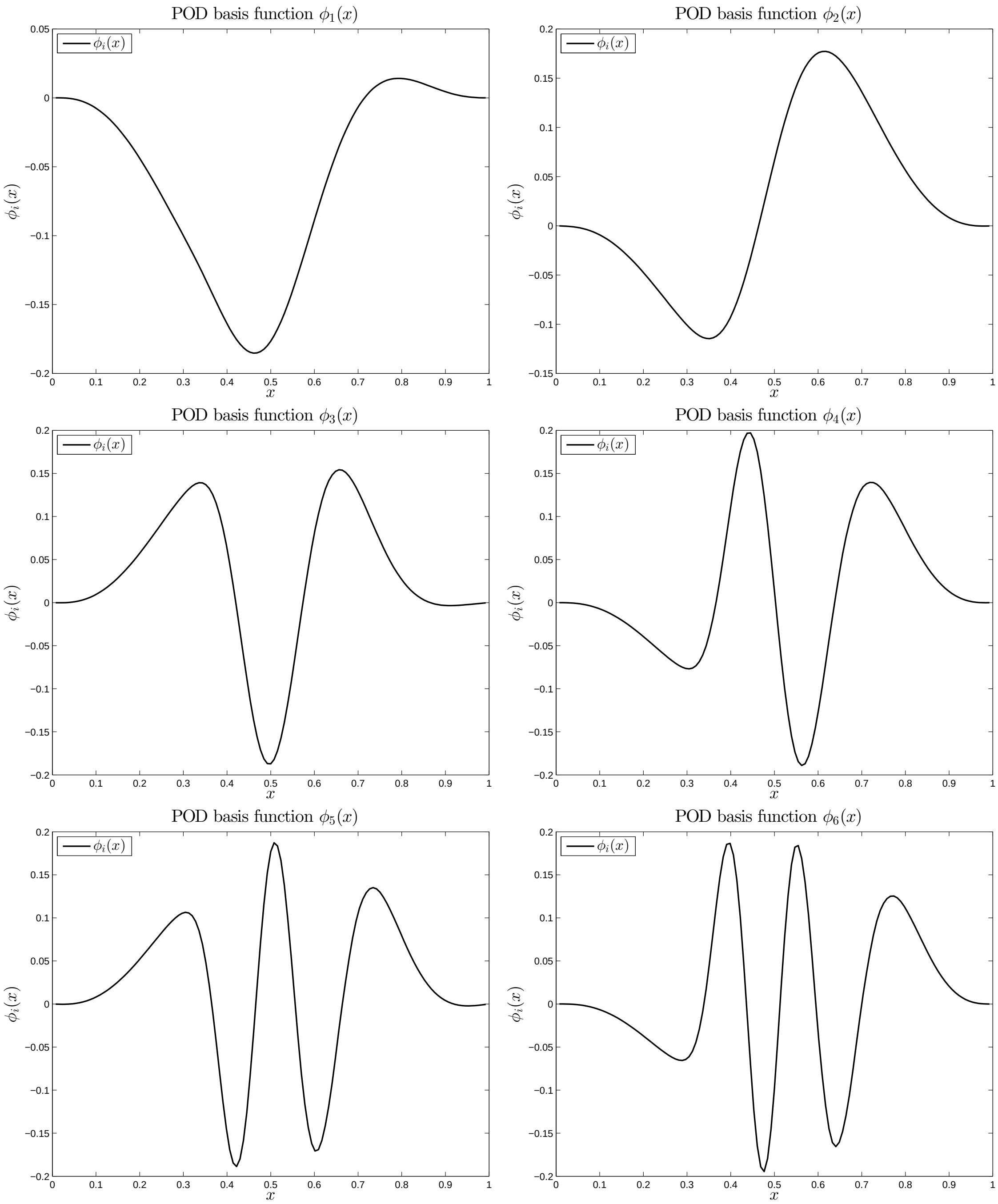

Figure 5.17: Example 2: Plot of the first 6 POD basis functions

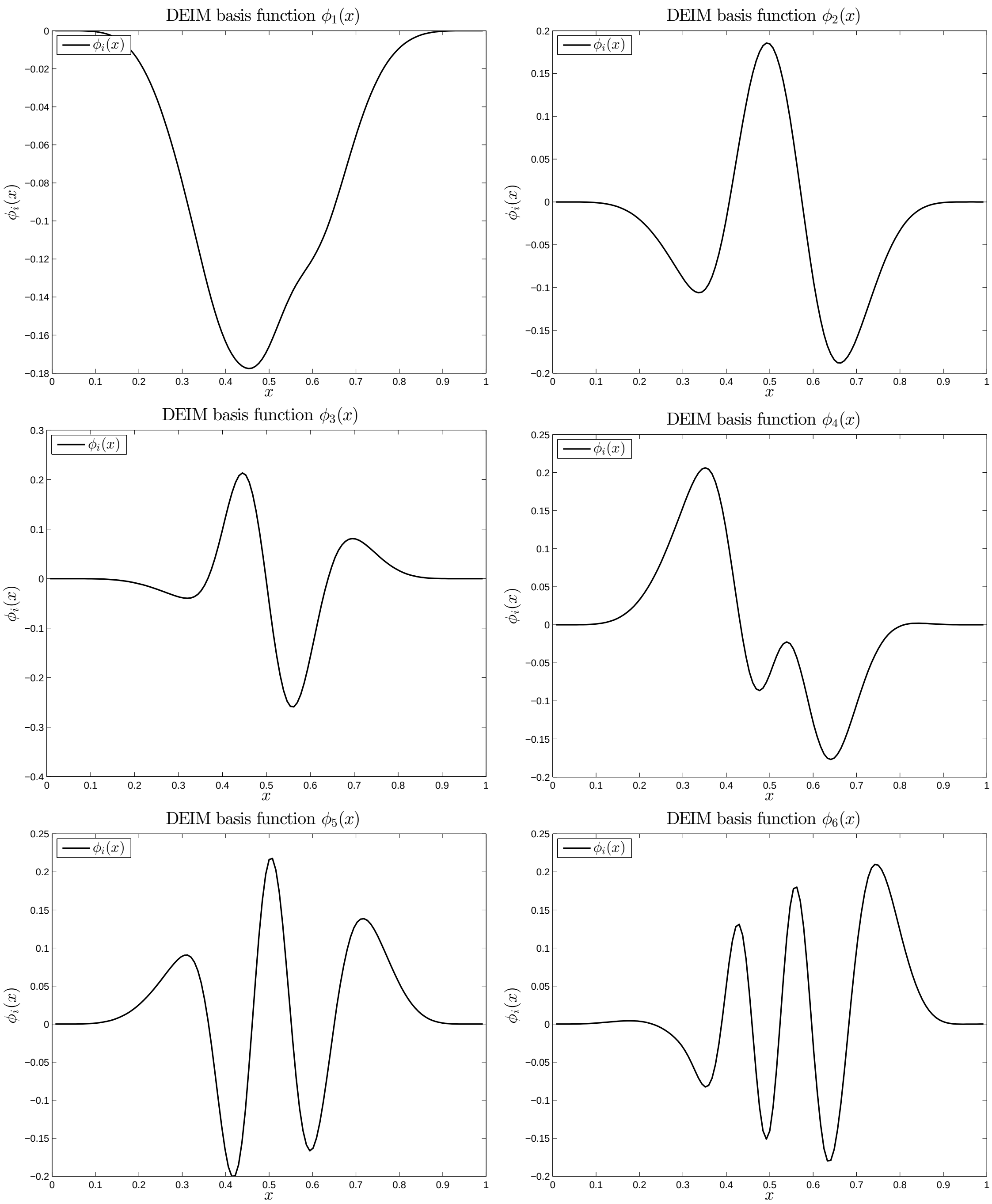

Figure 5.18: Example 2: Plot of the first 6 basis functions for DEIM

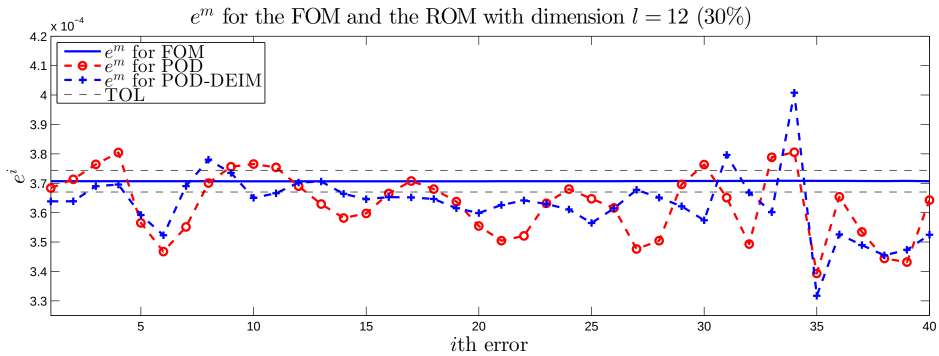

Figure 5.19: Example 2: Error estimators e^m and e^m_{rom} for the **ROM** with dimension $l = 12,\ 20$

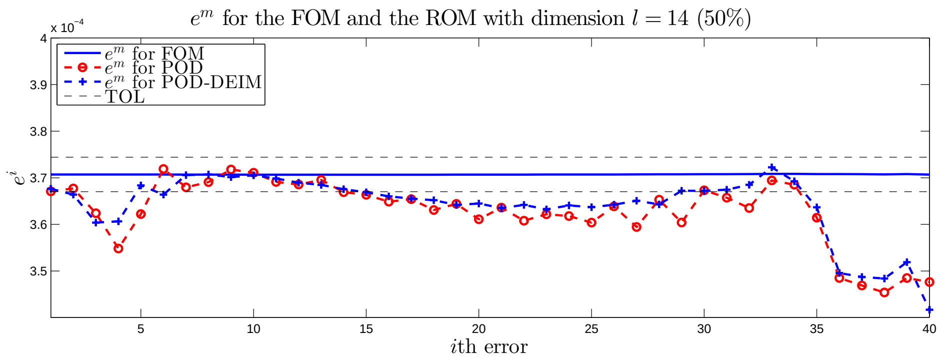

Figure 5.20: Example 2: Error estimators e^m and e^m_{rom} for the **ROM** with dimension $l = 20$

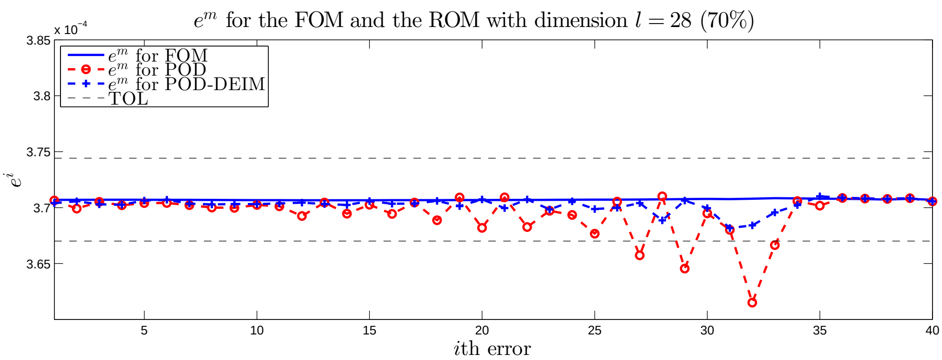

Figure 5.21: Example 2: Error estimators e^m and e^m_{rom} for the **ROM** with dimension $l = 28$

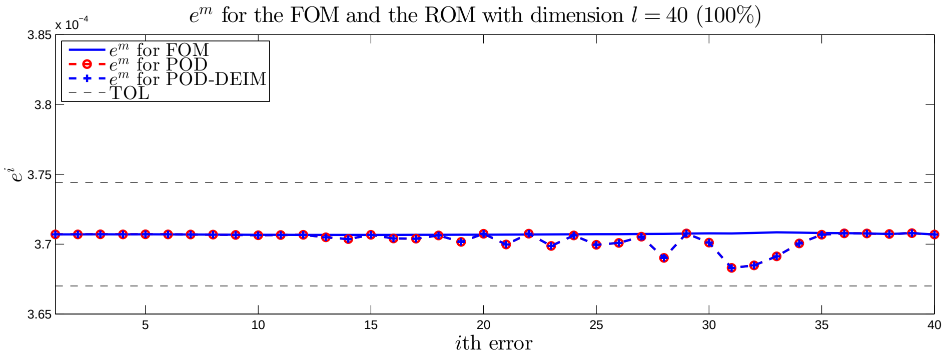

Figure 5.22: Example 2: Error estimators e^m and e^m_{rom} for the **ROM** with dimension $l = 40$

$\mu(\{e^m\}) \pm \sigma(e^m)$	POD	POD-DEIM	**FOM**
$l = 12$	$3.6287e-4 \pm 1.0639e-4$	$3.6314e-4 \pm 1.0940e-4$	$3.707e-4 \pm 5.0249e-8$
$l = 20$	$3.6290e-4 \pm 6.7517e-06$	$3.6417e-4 \pm 6.9064e-5$	NA
$l = 28$	$3.6945e-4 \pm 5.9643e-07$	$3.7026e-4 \pm 1.9100e-06$	NA
$l = 40$	$3.6963e-4 \pm 6.0801e-07$	$3.6963e-4 \pm 6.0801e-07$	NA

Table 5.3: Example 2: Average of the error estimators ($\mu(\{e^m\})$) and its standard deviation ($\sigma(e^m)$) for the **ROM** with dimension l.

5.2.2 Example 3: Error equilibration in time

In this numerical example, we use the same settings as in equation (5.3a) but change the non-linear term to y^3 and consider the error equilibration in time accordingly, i.e., we consider the following equation:

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + y^3 = f \quad \text{in } Q, & (5.4a) \\ y = 0 \quad \text{on } \Sigma, & (5.4b) \\ y(\cdot, 0) = 0 \quad \text{in } \Omega, & (5.4c) \end{cases}$$

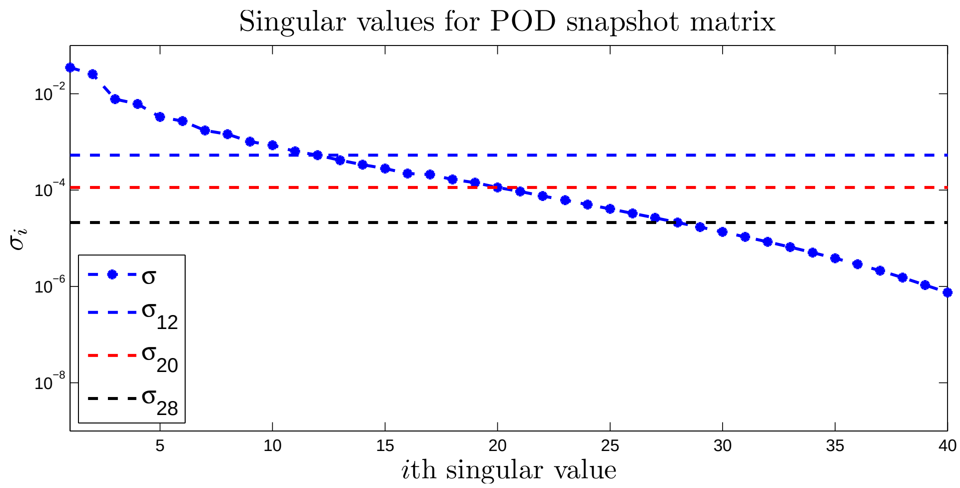

Figure 5.23: Example 3: Singular values $\{\sigma_i\}_{i=1}^{40}$ for POD snapshot matrix.

Similar to Example 2, we compute the time instances for error equilibration in time for the **FOM**. Then, we solve the equation (5.4a)-(5.4c) by implicit Euler method with a fixed time step length $\Delta t = \frac{1}{40}$ and use the solutions as the POD snapshot matrix to compute the error estimators for the **ROM** with POD dimensions $l = 12,\ 20,\ 28,\ 40$.

As the solution of (5.4a) is the same as that of (5.3a), the plot of solutions and the singular values of POD snapshot matrix are omitted. The singular values of POD snapshot matrix and DEIM non-linear snapshot matrix are plotted in Figure (5.23) and (5.24)

Compared to the singular values of POD snapshot matrix, the singular values of DEIM non-linear snapshot matrix are much smaller and are dominated by that of POD snapshot matrix. Therefore, the error estimators of POD based **ROM** are overlapped with that of POD-DEIM based **ROM**, as shown in Figure (5.25)-(5.28). However, DEIM method can significantly reduce the computing time for solving the **ROM**. Table (5.4) shows the average computing time for solving the two different types of **ROM** based on two different

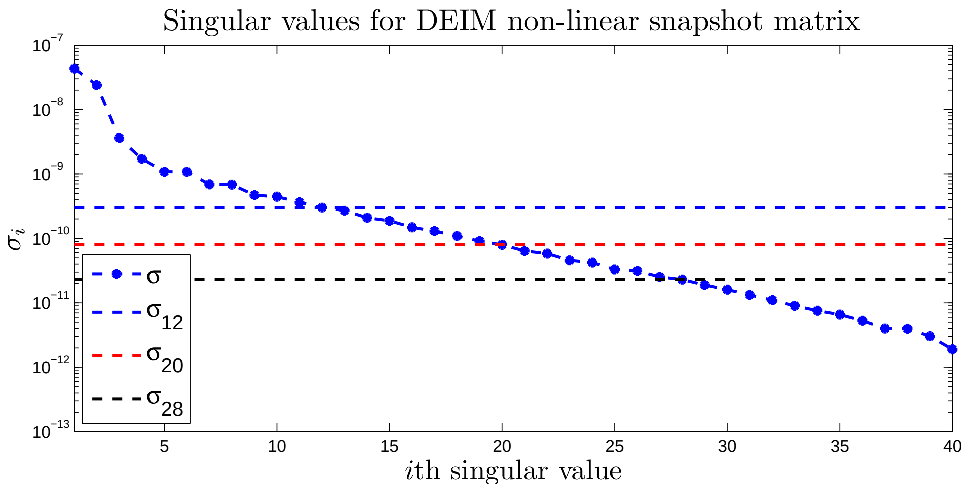

Figure 5.24: Example 3: Singular values $\{\sigma_i\}_{i=1}^{40}$ for DEIM non-linear snapshot matrix.

discretization in time for a given time partitioning.

	Implicit Euler	Modified Trapezoidal
POD	16.926	40.4198
POD-DEIM	5.0544	12.8544

Table 5.4: Example 3: Average computing time for solving types of **ROM** for equilibration time instances

The average of the error estimators and their standard deviation under different dimensions l are listed in Table (5.5), which shows how the error estimators of **ROM** converge to within the tolerance as we increase POD dimension.

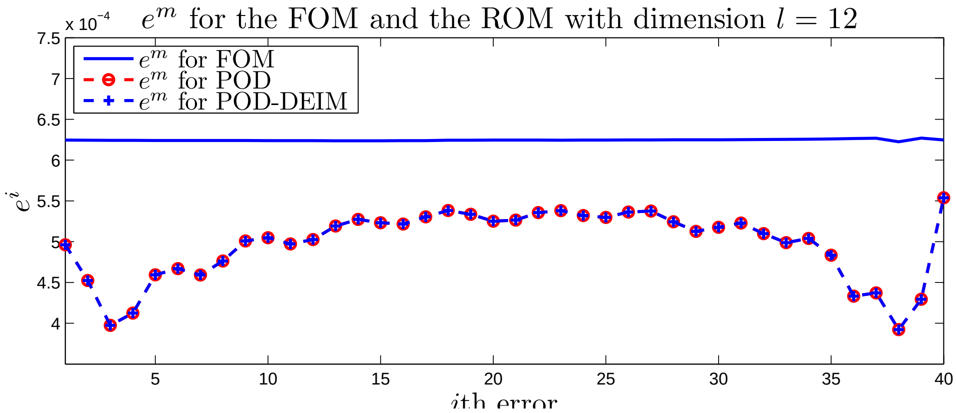

Figure 5.25: Example 3: Error estimators e^m and e^m_{rom} for the **ROM** with dimension $l = 12$

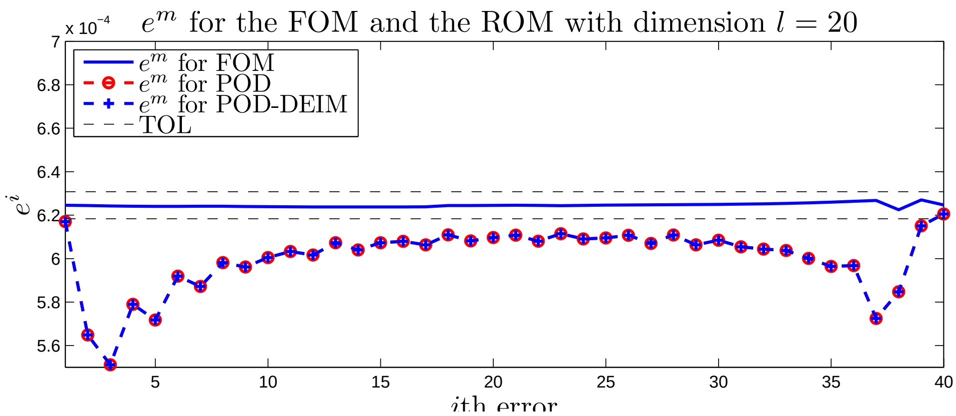

Figure 5.26: Example 3: Error estimators e^m and e^m_{rom} for the **ROM** with dimension $l = 20$

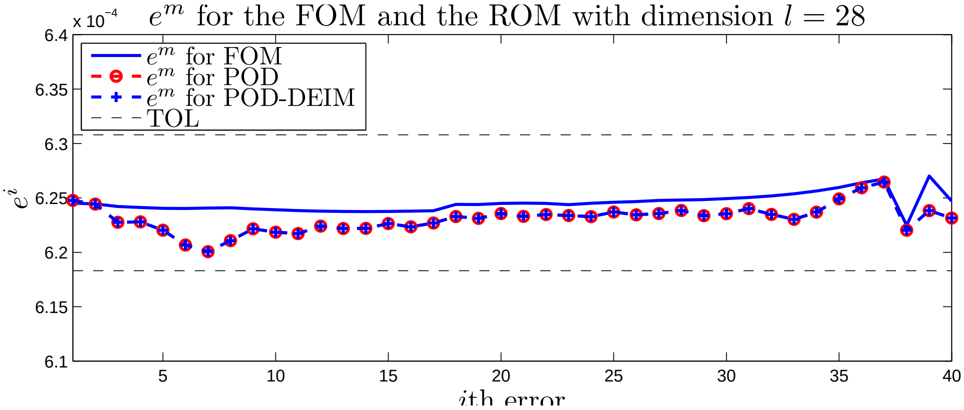

Figure 5.27: Example 3: Error estimators e^m and e^m_{rom} for the **ROM** with dimension $l = 28$

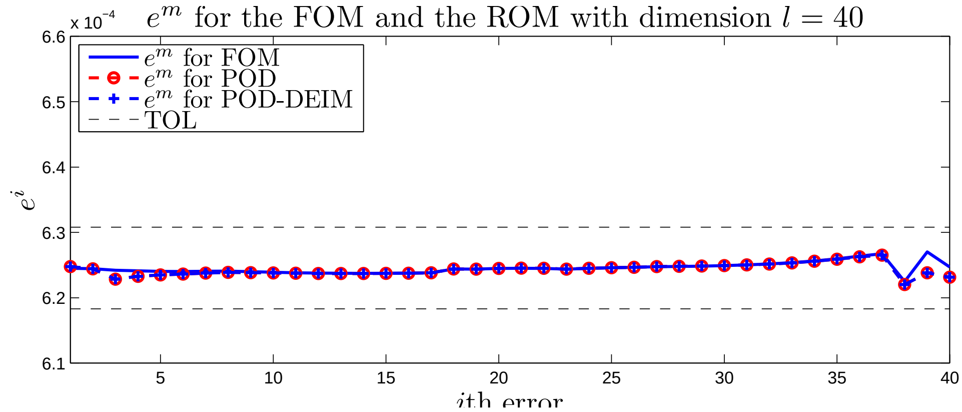

Figure 5.28: Example 3: Error estimators e^m and e^m_{rom} for the **ROM** with dimension $l = 40$

$\mu(\{e^m\}) \pm \sigma(e^m)$	POD	POD-DEIM	**FOM**
$l = 12$	$4.9751e-4 \pm 4.1643e-5$	$4.9751e-4 \pm 4.1643e-5$	$6.2455e-4 \pm 8.7215e-7$
$l = 20$	$6.0040e-4 \pm 1.4586e-5$	$6.0040e-4 \pm 1.4586e-5$	NA
$l = 28$	$6.2310e-4 \pm 1.2263e-6$	$6.2310e-4 \pm 1.2263e-6$	NA
$l = 40$	$6.2432e-4 \pm 8.8220e-7$	$6.2432e-4 \pm 8.8220e-7$	NA

Table 5.5: Example 3: Average of the error estimators ($\mu(\{e^m\})$) and its standard deviation ($\sigma(e^m)$) for the **ROM** with dimension l.

Chapter 6

Conclusions

So far, we have developed an approach for determining the optimal snapshot locations for time dependent parabolic PDEs from the ideal of automated time stepping, the so-called error equilibration in time. The goal of this approach is to determine the time instances such that on each time sub-interval the error estimators are almost of the same order of magnitude. Different from the automated time stepping, the total number of time steps in this method is fixed and the length of each time sub-interval is adaptive according to the corresponding error estimator. This error equilibration approach can be formulated as an optimization problem about the variance of the error estimators and a bi-section type algorithm is provided. This approach applies to both **FOM** and **ROM**. The property of error equilibration in time for **FOM** is preserved by the POD and POD-DEIM based **ROM** provided the dimension of the **ROM** is large enough such that the POD or POD-DEIM error is sufficiently small, as shown in numerical examples 1, 2, 3. Moreover, as stated in Theorem 4.12 and 4.15, if the error estimators for **ROM** are computed from the time instances such that the error estimators for **FOM** are equilibrated, then they are bounded by some constants. This result is shown in Figure (5.5), (5.6), (5.19), (5.20), (5.25) and (5.26). Although the error estimators for **ROM** do not located within the given tolerance for equilibration TOL, they are bounded above and below.

Bibliography

[1] J.M.Ortega, and W.C.Rheinboldt; Iterative Solution of Nonlinear Equations in Several Variables Academic Press, 1970.

[2] H. Brézis and A. Pazy; Convergence and approximation of semigroups of nonlinear operators, J. Funct. Anal. **9**, 63–74, 1972.

[3] P. Deuflhard; Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms, Springer, Berlin-Heidleberg-New York, 2004.

[4] Chaturantabut. Saifon, D. C. Sorensen; Nonliear model reduction via discrete emperical interpolation SIAM Journal on Scientific Computing **32**, 2737-2764, 2010.

[5] T. Henri; Geometric Theory of Semilinear Parabolic Equations. Springer, Berlin-Heigelberg-New York, 1981.

[6] Earl A. Coddington, Norman Levinson (1955); Theory of Ordinary Differential Equations, New York: McGraw-Hill, 1995.

[7] T. Kato; Nonlinear semigroups and evolution equations, J. Math. Soc. Japan **19**, 503–520, 1967.

[8] N. Kenmochi and S. Oharu; Difference approximation of nonlinear evolution equations and semigroups of nonlinear operators. Publ. RIMS, Kyoto Univ. **10**, 147–207, 1974.

[9] K. Kunisch and S. Volkwein; Proper orthogonal decomposition and singular value decomposition. J. Optim. Theory and Appl. **102**, 345–371, 1999.

[10] K. Kunisch and S. Volkwein; Galerkin proper orthogonal decomposition methods for parabolic problems. Numer. Math. **90**, 117–148, 2001.

[11] K. Kunisch and S. Volkwein; Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. SIAM J. Numer. Anal. **40**, 492–515, 2002.

[12] K. Kunisch and S. Volkwein; Proper orthogonal decomposition for optimality systems. M2AN, **42**, 1–23, 2008.

[13] K. Kunisch and S. Volkwein; Optimal snapshot location for computing POD basis functions. Preprint. Department of Mathematics. Karl-Franzens University of Graz, 2009.

[14] O.A. Ladyzenskaja, V.A. Solonnikow, and N.N. Ural'ceva. Linear and quasilinear Equations of Parabolic Type. AMS, Providence, 1968

[15] H.V. Ly and H.T. Tran; Modelling and control of physical processes using proper orthogonal decomposition. Math. and Comput. Modelling **33**, 223–236, 2001.

[16] B. C. Moore; Principal component analysis in linear systems: controllability, observability, and model reduction, IEEE Trans. Automat. Control **26**, 17–32, 1981.

[17] K. Sun, R. Glowinski, M. Heinkenschloss, and D. C. Sorensen; Domain decomposition and model reduction of systems with local nonlinearities, In: Numerical Mathematics and Advanced Applications. ENUMATH 2007 (K. Kunisch, G. Of, and O. Steinbach, eds.), pp. 389–396, Springer, Berlin-Heidelberg-New York, 2008.

[18] F. Tröltzsch and S. Volkwein; POD a posteriori error estimates for linear-quadratic optimal control problems. to appear in Comp. Opt. and Appl., 2009

[19] K. Zhou, J. C. Doyle, and K. Glover; Robust and Optimal Control, Prentice Hall, Englewood Cliffs, 1996.

[20] M.Fahl; Computation of POD basis functions for fluid flows with Lanczos methods. Mathematical and Computer Modelling, 2000

[21] K.Fukunaga; Introduction to Statistical Recognition, Academic Press, 1999

[22] P. Astrid, Reduction of Process Simulation Models; A Proper Orthogonal Decomposition Approach, Ph.D. thesis, Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands, 2004.

[23] M. J. Rewienski ; A Trajectory Piecewise-Linear Approach to Model Order Reduction of Nonlinear Dynamical Systems, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2003.

[24] M. Rewienski and J. White; Model order reduction for non-linear dynamical systems based on trajectory piecewise-linear approximations, Linear Algebra Appl., 415 (2006), pp. 426454

[25] S.C.Brenner and L.Ridgway Scott; The mathematical Theory of Finite Flement Methods. 3rd Edition, Springer, New York, 2008

[26] P.F.Ciarlet; The Finite Element Method for Elliptic Problems. SIAM, Philadelphia, 2002

[27] Bei Hu; Blow-up Theories for Semilinear Parabolic Equations Springer Heidelberg Dordrecht London New York, 2011