

LOCAL IMPROVEMENTS TO REDUCED-ORDER APPROXIMATIONS OF
PDE-CONSTRAINED OPTIMIZATION PROBLEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

TUĞBA AKMAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
SCIENTIFIC COMPUTING

AUGUST 2015

Approval of the thesis:

**LOCAL IMPROVEMENTS TO REDUCED-ORDER APPROXIMATIONS
OF PDE-CONSTRAINED OPTIMIZATION PROBLEMS**

submitted by **TUĞBA AKMAN** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Department of Scientific Computing, Middle East Technical University** by,

Prof. Dr. Bülent Karasözen
Director, Graduate School of **Applied Mathematics**

Assoc. Prof. Dr. Ömür Uğur
Head of Department, **Scientific Computing**

Prof. Dr. Bülent Karasözen
Supervisor, **Scientific Computing, METU**

Examining Committee Members:

Prof. Dr. Gerhard Wilhelm Weber
Financial Mathematics, METU

Prof. Dr. Bülent Karasözen
Scientific Computing, METU

Prof. Dr. Tanıl Ergenç
Department of Mathematics, Atılım University

Assoc. Prof. Dr. Murat Manguoğlu
Computer Engineering, METU

Assist. Prof. Dr. Fikriye Nuray Yılmaz
Department of Mathematics, Gazi University

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: TUĞBA AKMAN

Signature :

ABSTRACT

LOCAL IMPROVEMENTS TO REDUCED-ORDER APPROXIMATIONS OF PDE-CONSTRAINED OPTIMIZATION PROBLEMS

Akman, Tuğba

Ph.D., Department of Scientific Computing

Supervisor : Prof. Dr. Bülent Karasözen

August 2015, 152 pages

Optimal control problems (OCPs) governed partial differential equations (PDEs) arise in environmental control problems, optimal control of fluid flow, petroleum reservoir simulation, laser surface hardening of steel, parameter estimation and in many other applications. Although the OCPs governed by elliptic and parabolic problems are investigated theoretically and numerically in several papers, the studies concerning the optimal control of evolutionary diffusion-convection-reaction (DCR) equation and Burgers equation are quite rare.

In this study, we consider the optimal control problem governed by the unsteady diffusion-convection-reaction equation and Burgers equation without control constraints. These problems gain importance, especially when the diffusive term is small. In such cases, the numerical solution exhibit interior/boundary layers and classical finite element method (FEM) is not efficient for derivation of an accurate numerical solution and methods requiring higher regularity of the solution might not be practical. Therefore, we solve these problems using variational time discretization method, which is a stable, superconvergent technique requiring less regularity when compared to the methods of the same order; and symmetric interior penalty Galerkin (SIPG) with upwinding in space, which flexes inter element continuity of the solution. We provide a priori error estimates for space-time discontinuous Galerkin method and present numerical findings.

An accurate and stable numerical solution requires a fine grid/mesh, which increases the dimension of the discrete problem, so the computational time. In case of perturbations in the data, full-order model (FOM) is required to be solved for each new parameter in the data set. In case of optimization problems, FOM associated to the differential equations must be resolved after updating the control. Therefore, we use a model-order reduction (MOR) technique that eliminates the necessity of the solution of the FOM for each parameter and that enables us to solve the problem in a fast way. We use one of the most popular and successful MOR techniques, namely the proper orthogonal decomposition (POD) method. The idea behind the POD method is to derive a new basis spanning the space whose dimension is lower than the finite element space. Then, the FOM is projected onto the low-dimensional space using the new optimal POD basis as we proceed in Galerkin projection. In addition, a priori error estimates associated to reduced-order model (ROM) based on space-time dG method are proven and numerical results are shown.

The POD basis is computed using the snapshots of a particular problem which is interpreted by a mathematical model and data. Because there is a link between the data and the snapshots, some perturbation in the data may lead to larger changes in the snapshots depending on the problem at hand. This leads the nominal/baseline POD basis, which depends on the nominal/baseline parameters, not to approximate the perturbed problem accurately. In such cases, one has to solve the full problem for each parameter in the data set again and regenerate the POD basis. This approach is expensive especially for nonlinear problems or optimal control problems which requires the solution of a set of differential equations. Thus, POD sensitivities are used to enrich the low-dimensional subspace for a wider range of parameters and the quantity of interest is the diffusion term ϵ , the convection term β and the reaction term r . We generate two new bases, i.e. extrapolated POD (ExtPOD) and expanded POD (ExpPOD) and compare these bases in terms of advantages and discuss the main drawbacks of them.

Keywords : Optimal control problems, proper orthogonal decomposition method, space-time discontinuous Galerkin method, sensitivity analysis, a priori error estimates

ÖZ

KİSMİ DİFERANSİYEL DENKLEMLERİN ENİYİLEMELİ KONTROL PROBLEMLERİNİN İNDİRGENMİŞ MERTEBEDEN YAKLAŞIMLARI İÇİN YEREL GELİŞTİRMELER

Akman, Tuğba

Doktora, Bilimsel Hesaplama Bölümü

Tez Yöneticisi : Prof. Dr. Bülent Karasözen

Ağustos 2015, 152 sayfa

Ekolojik kontrol problemleri, akışkanların eniyilemeli kontrol problemleri, petrol birikim simülasyonu, çelik yüzeyinin sertleştirilmesi, değişken tahminleri gibi birçok matematiksel problem, kısmi diferansiyel denklemlerin eniyilemeli kontrol problemleri şeklinde ifade edilmektedir. Eliptik ve parabolik diferansiyel denklemlerin eniyilemeli kontrol problemleri ile ilgili çok sayıda çalışma olmasına karşın, zamana bağlı difüzyon konveksiyon ve reaksiyon terimleri içeren diferansiyel denklemlerin ve Burgers denkleminin eniyilemeli kontrol problemleri ile ilgili çalışma az sayıdadır.

Bu tezde, zamana bağlı difüzyon konveksiyon ve reaksiyon terimleri içeren diferansiyel denklemlerin ve Burgers denkleminin kontrol kısıtı olmayan eniyilemeli kontrol problemleri çalışılmıştır. Bu problem, difüzyon terimi küçük olduğunda önem kazanmaktadır. Bu durumda, çözüm, tanım kümesi içinde veya tanım kümesinin sınırına yakın bölgelerde katmanlar içermektedir. Sürekli sınırlı elemanlar yöntemi bu tip problemlerde iyi sonuçlar vermemektedir. Bu nedenle, kararlı, çok yakınsak ve aynı dereceden yöntemlere göre çözümün daha az türevlenebilir olmasının yeterli olduğu bir yöntem olan varyasyonel zaman ayrıklaştırılması yöntemi ile çözümün elemanlar arasında sürekli olması kısıtını kaldıran simetrik kesintili Galerkin yöntemi bir arada kullanılmıştır. Uzay-zaman kesintili Galerkin yöntemi için, önceden hata tahminleri elde edilmiş ve nümerik sonuçlar gösterilmiştir.

Kesin ve kararlı sayısal çözüm elde edebilmek için, problemin sık bir ağ üzerinde çözülmesi gerekmektedir. Bu da ayrık problemin boyutunu artırır ve daha geç çözülmesine yol açar. Verideki her bozunumda, tam dereceli modelin yeniden çözülmesi gerekmektedir. Eniyileme problemlerinde ise, yeni kontrol bulunduktan sonra, tam dereceli modelle ilişkili birden fazla diferansiyel denklemin nümerik çözümüne gerek duyulmaktadır. Bu nedenle, tam dereceli modelin çözülmesi ihtiyacını ortadan kaldıran ve problemin daha hızlı çözülmesini sağlayan, modelin derecesini azaltma yöntemleri tercih edilmektedir. Bu çalışmada, en yaygın model indirgeme yöntemlerinden biri olan, öz dik ayrışım yöntemi kullanılmaktadır. Bu yöntem ile boyutu indirgenmiş ve kısıtlı elemanlar uzayının içinde kalan yeni bir uzayı üreten tabanlar bulunmaktadır. Bu taban kullanılarak, tam dereceli model, boyutu indirgenmiş uzaya Galerkin yöntemi ile izdüşürülür. Buna ek olarak, indirgenmiş model için uzay-zaman kesintili Galerkin yöntemine dayanarak, önceden hata tahminleri hesaplanmış ve nümerik sonuçlarla güçlendirilmiştir.

Öz dik ayrışım yöntemi, üretildiği sayısal çözüm ile ilgili bilgi taşımaktadır. Bu sayısal çözümün elde edildiği verideki bozunumlar, öz dik ayrışım yöntemi ile elde edilen tabanın iyi sonuç vermemesine yol açar. Bu durumda, verideki her değişim için tam boyutlu problemin yeniden çözülmesi ve tabanın yeniden üretilmesi gerekmektedir. Özellikle lineer olmayan problemler için pahalı olan bu seçenek yerine, çözümün difüzyon, adveksiyon ve reaksiyon değişkenine göre hassasiyeti kullanılarak, boyutu indirgenmiş uzay zenginleştirilebilir. Bu bilgiyi kullanarak iki yeni taban, dışdeğerbiçim tabanı ve genişletilmiş taban, üretilmiştir. Sayısal sonuçlar kullanılarak, bu bazların faydaları ve eksikleri incelenmiştir. Her bir problem için, en uygun baza karar verilmiştir.

Anahtar Kelimeler: Eniyilemeli kontrol problemleri, öz dik ayrışım yöntemi, uzay-zaman kesintili Galerkin yöntemi, hassasiyet analizi, önceden hata tahminleri

To my grandparents

ACKNOWLEDGMENTS

I would like to express my gratitude to all those people who have helped and supported me during my studies.

I'm grateful to my supervisor Prof. Dr. Bülent Karasözen for his guidance and for introducing me the field of PDE-constrained optimization and discontinuous Galerkin methods.

I'd like to thank Prof. Dr. Gerhard Wilhelm Weber for his kindness and proofreading.

I'm grateful to my thesis defense committee members for their comments and discussions.

I especially thank to my family for their warm love, endless support and encouragement throughout my life. Deepest thanks to my best friend, Burak Yıldız, who has supported me in many respects and who has been with me during my tough times. Without them, I could not have been such strong and hopeful.

TABLE OF CONTENTS

ABSTRACT	vii
ÖZ	ix
ACKNOWLEDGMENTS	xiii
TABLE OF CONTENTS	xv
LIST OF FIGURES	xix
LIST OF TABLES	xxiii
LIST OF ABBREVIATIONS	xxv
CHAPTERS	
1 INTRODUCTION	1
2 OPTIMAL CONTROL PROBLEMS	7
2.1 Preliminaries	8
2.2 Optimality Conditions	11
2.2.1 Optimality Conditions for Diffusion Convection Reaction Equation	12
2.2.2 Optimality Conditions for Burgers Equation	12
2.3 Optimization Algorithm	13
2.3.1 First-order Derivatives	14
2.3.2 Second-order Derivatives	16
3 DISCONTINUOUS GALERKIN METHODS	19

3.1	Discontinuous Galerkin Methods in One-Dimension	20
3.1.1	Semi-discrete Optimal Control Problem	23
3.2	Discontinuous Galerkin Methods in Two-Dimension	24
3.2.1	Semi-discrete Optimal Control Problem	28
4	VARIATIONAL TIME DISCRETIZATION METHODS	29
4.1	Continuous Galerkin-Petrov (cGP(q+1)) Method	31
4.2	Discontinuous Galerkin (dG(q)) Methods	34
4.3	A Priori Error Estimates for Optimal Control of Diffusion- Convection-Reaction Equation	37
4.3.1	Some Auxiliary Results	38
4.3.2	Discrete Characteristic Function	39
4.3.3	Main Result	40
4.4	Computational Aspects	46
4.4.1	Numerical Results	50
4.5	A Priori Error Estimates for Optimal Control of Burgers Equa- tion	54
4.5.1	Some Auxiliary Results	54
4.5.2	Main Result	55
4.6	Computational Aspects	58
4.6.1	Numerical Results	60
5	MODEL REDUCTION USING PROPER ORTHOGONAL DECOM- POSITION METHOD	63
5.1	Continuous POD Method	65
5.2	Discrete POD Method	67

5.2.1	Spatial Discretization of POD Method	67
5.2.2	Semi-discrete Optimality System Based on POD	68
5.2.3	Temporal Discretization POD Method	69
5.2.4	Fully-discrete Optimality System Based on POD	71
5.3	A Priori Error Estimates for Optimal Control of Diffusion- Convection-Reaction Equation	73
5.3.1	Auxiliary Results	73
5.3.2	Main Result	79
5.3.3	Numerical Results	84
5.4	Discrete Empirical Interpolation Method	90
5.5	A Priori Error Estimates for Optimal Control of Burgers Equa- tion	92
5.5.1	Auxiliary Results	93
5.5.2	Main Result	93
5.5.3	Numerical Results	95
6	LOCAL IMPROVEMENTS TO REDUCED-ORDER SOLUTION	99
6.1	Proper Orthogonal Decomposition Sensitivities	100
6.2	Sensitivity Equations for Optimal Control of Diffusion-Convection- Reaction Equation	102
6.2.1	Sensitivity with respect to the diffusion term ϵ	102
6.2.2	Sensitivity with respect to the convective term β	103
6.2.3	Sensitivity with respect to the reaction term r	104
6.2.4	Numerical Results	104
6.3	Sensitivity Equations for Optimal Control of Burgers Equation	128

6.3.1	Numerical Results	128
7	CONCLUSION	135
	REFERENCES	137
	CURRICULUM VITAE	149

LIST OF FIGURES

Figure 4.1 DCR Eqn: Numerical solution of the state (<i>left</i>), error (<i>right</i>) for Example 1 at $t=0.5$ with $h = k = 1/80$ by dG(1) method	51
Figure 4.2 DCR Eqn: Numerical solution of the control (<i>left</i>), error (<i>right</i>) for Example 1 at $t=0.5$ with $h = k = 1/80$ by dG(1) method	51
Figure 4.3 DCR Eqn: Numerical solution of the state (<i>left</i>), error (<i>right</i>) for Example 2 at $t=0.5$ with $h = k = 1/80$ by dG(1) method	52
Figure 4.4 DCR Eqn: Numerical solution of the control (<i>left</i>), error (<i>right</i>) for Example 2 at $t=0.5$ with $h = k = 1/80$ by dG(1) method	53
Figure 4.5 Burgers Eqn: Numerical solution of the state (<i>top-left</i>), the adjoint (<i>top-right</i>) and the control (<i>bottom</i>) for Example 1 with $h = k = 1/128$ by dG(1) method	61
Figure 5.1 Convection field of the state and the adjoint, respectively	85
Figure 5.2 DCR Eqn: Numerical solution of the state at $t = 0.2, 0.6, 1$, respectively	86
Figure 5.3 DCR Eqn: Numerical solution of the adjoint at $t = 0.8, 0.4, 0$, respectively	86
Figure 5.4 DCR Eqn: Eigenvalues for snapshot ensemble Y, P and $Y \cup P$	86
Figure 5.5 DCR Eqn: Error versus number of POD basis functions for state (<i>top-left</i>), adjoint (<i>top-right</i>) and control (<i>bottom</i>)	87
Figure 5.6 DCR Eqn: Order convergence for state (<i>top-left</i>), adjoint (<i>top-right</i>) and control (<i>bottom</i>) $h = k = 1/80$	89
Figure 5.7 Burgers Eqn: Eigenvalues for snapshot ensemble Y, P and $Y \cup P$	96
Figure 5.8 Burgers Eqn: Error versus number of POD basis functions for state (<i>top-left</i>), adjoint (<i>top-right</i>) and control (<i>bottom</i>)	97
Figure 5.9 Burgers Eqn: Order convergence for state (<i>top-left</i>), adjoint (<i>top-right</i>) and control (<i>bottom</i>)	98
Figure 6.1 Convection field of the state $\beta = [y - 0.5, -x + p]$ with $p = 0.1, 0.5, 0.9$, respectively	103

Figure 6.2	Convection field of the adjoint $\beta = [-y + 0.5, x - p]$ with $p = 0.1, 0.5, 0.9$, respectively	103
Figure 6.3	DCR Eqn - Case I: State sensitivities computed with CSE method (<i>first row</i>) and FD approximation (<i>second row</i>) at $t = 0.2, 0.6, 1$, respectively	106
Figure 6.4	DCR Eqn - Case I: Adjoint sensitivities computed with CSE method (<i>first row</i>) and FD approximation (<i>second row</i>) at $t = 0.2, 0.6, 1$, respectively	107
Figure 6.5	DCR Eqn - Case I: Eigenvalue sensitivities for snapshot ensemble Y, P and $Y \cup P$	108
Figure 6.6	DCR Eqn - Case I: Local improvements in the error of the control with 7 (<i>first column</i>) and 14 (<i>second column</i>) POD basis functions	110
Figure 6.7	DCR Eqn - Case I: Local improvements in the error of the state with 7 (<i>first column</i>) and 14 (<i>second column</i>) POD basis functions	111
Figure 6.8	DCR Eqn - Case I: Error versus number of POD basis functions for control with $\epsilon = 1/80$ (<i>first column</i>) and $\epsilon = 1/120$ (<i>second column</i>)	112
Figure 6.9	DCR Eqn - Case I: Error versus number of POD basis functions for state with $\epsilon = 1/80$ (<i>first column</i>) and $\epsilon = 1/120$ (<i>second column</i>)	113
Figure 6.10	DCR Eqn - Case II: State sensitivities computed with CSE method (<i>first row</i>) and FD approximation (<i>second row</i>) at $t = 0.2, 0.6, 1$, respectively	114
Figure 6.11	DCR Eqn - Case II: Adjoint sensitivities computed with CSE method (<i>first row</i>) and FD approximation (<i>second row</i>) at $t = 0.8, 0.4, 0$, respectively	115
Figure 6.12	DCR Eqn - Case II: Eigenvalue sensitivities for snapshot ensemble Y, P and $Y \cup P$	116
Figure 6.13	DCR Eqn - Case II: Local improvements in the error of the control with 7 (<i>first column</i>) and 14 (<i>second column</i>) POD basis functions	117
Figure 6.14	DCR Eqn - Case II: Local improvements in the error of the state with 7 (<i>first column</i>) and 14 (<i>second column</i>) POD basis functions	118
Figure 6.15	DCR Eqn - Case II: Error versus number of POD basis functions for control with $p = 0.1$ (<i>first column</i>) and $p = 0.9$ (<i>second column</i>)	119
Figure 6.16	DCR Eqn - Case II: Error versus number of POD basis functions for state with $p = 0.1$ (<i>first column</i>) and $p = 0.9$ (<i>second column</i>)	120
Figure 6.17	DCR Eqn - Case III: State sensitivities computed with CSE method (<i>first row</i>) and FD approximation (<i>second row</i>) at $t = 0.2, 0.6, 1$, respectively	121
Figure 6.18	DCR Eqn - Case III: Adjoint sensitivities computed with CSE method (<i>first row</i>) and FD approximation (<i>second row</i>) at $t = 0.8, 0.4, 0$, respectively	122

Figure 6.19 DCR Eqn - Case III: Eigenvalue sensitivities for snapshot ensemble Y, P and $Y \cup P$	122
Figure 6.20 DCR Eqn - Case III: Local improvements in the error of the control with 7 (<i>first column</i>) and 14 (<i>second column</i>) POD basis functions	124
Figure 6.21 DCR Eqn - Case III: Local improvements in the error of the state with 7 (<i>first column</i>) and 14 (<i>second column</i>) POD basis functions	125
Figure 6.22 DCR Eqn - Case III: Error versus number of POD basis functions for control with $p = 0.6$ (<i>first column</i>) and $p = 1.4$ (<i>second column</i>)	126
Figure 6.23 DCR Eqn - Case III: Error versus number of POD basis functions for state with $p = 0.6$ (<i>first column</i>) and $p = 1.4$ (<i>second column</i>)	127
Figure 6.24 Burgers Eqn: State (<i>left</i>) and adjoint (<i>right</i>) sensitivities	129
Figure 6.25 Burgers Eqn: Eigenvalue sensitivities for snapshot ensemble Y, P and $Y \cup P$	130
Figure 6.26 DEIM interpolation points	130
Figure 6.27 Burgers Eqn: Local improvements in the error of the state with 5 (<i>first column</i>) and 10 (<i>second column</i>) POD basis functions	132
Figure 6.28 Burgers Eqn: Local improvements in the error of the control with 5 (<i>first column</i>) and 10 (<i>second column</i>) POD basis functions	134

LIST OF TABLES

Table 4.1 DCR Eqn: Example 1 by cGP(2) and dG(1) (<i>in parenthesis</i>) method.	51
Table 4.2 DCR Eqn: Example 2 by cGP(2) and dG(1) (<i>in parenthesis</i>) method.	53
Table 4.3 Burgers Eqn: Example 1 by dG(1) method.	61
Table 6.1 DCR Eqn - Case I: Computation times in seconds for the full-order problem	106
Table 6.2 DCR Eqn - Case I: Computation times in seconds for the computation of POD basis functions	106
Table 6.3 DCR Eqn - Case I: Computational speedup for the reduced-order model	107
Table 6.4 Burgers Eqn: Computation times for the full-order problem	130
Table 6.5 Burgers Eqn: Computation times in seconds for the computation of POD basis functions	131
Table 6.6 Burgers Eqn: Computational speedup for the reduced-order model with/out DEIM	131

LIST OF ABBREVIATIONS

BPOD	Baseline proper orthogonal decomposition
CG	Conjugate gradient
cGP	Continuous Galerkin/Petrov
CSE	Continuous sensitivity equation
DCR	Diffusion-convection-reaction
DEIM	Discrete empirical interpolation method
dG	Discontinuous Galerkin
DO	Discretize-then-optimize
dof	Degrees of freedom
ExpPOD	Expanded proper orthogonal decomposition
ExtPOD	Extrapolated proper orthogonal decomposition
FD	Finite difference
FEM	Finite element method
FOM	Full-order model
MOR	Model-order reduction
OCP	Optimal control problem
OD	Optimize-then-discretize
PDE	Partial differential equation
POD	Proper orthogonal decomposition
ROM	Reduced-order model
SIPG	Symmetric interior penalty Galerkin
SVD	Singular value decomposition

CHAPTER 1

INTRODUCTION

Optimal control problems (OCPs) governed partial differential equations (PDEs) arise in environmental control problems [38], optimal control of fluid flow [69, 126], petroleum reservoir simulation [78], laser surface hardening of steel [106], parameter estimation [125], biology [90] and in many other applications. In addition to real-life applications, OCPs governed by elliptic and parabolic problems are investigated theoretically and numerically in several studies [1, 17, 29, 34, 68, 97, 101, 142]. However, studies concerning the optimal control of evolutionary diffusion-convection-reaction (DCR) equation and Burgers equation are quite rare [5, 57, 141, 131, 132]. Therefore, in this study, firstly, we consider the following distributed optimal control problem governed by unsteady diffusion-convection-reaction equation

$$\underset{u \in L^2(0,T;L^2(\Omega))}{\text{minimize}} \quad J(y, u) := \frac{1}{2} \int_0^T (\|y - y_d\|_{L^2(\Omega)}^2 + \alpha \|u\|_{L^2(\Omega)}^2) dt, \quad (1.1a)$$

$$\text{subject to } \partial_t y - \epsilon \Delta y + \beta \cdot \nabla y + ry = f + u \quad (x, t) \in \Omega \times (0, T], \quad (1.1b)$$

$$y(x, t) = 0 \quad (x, t) \in \partial\Omega \times [0, T], \quad (1.1c)$$

$$y(x, 0) = y_0(x) \quad x \in \Omega, \quad (1.1d)$$

where $\Omega = (0, 1) \times (0, 1)$ and $I = (0, T)$. The source function and the desired state are denoted by $f \in L^2(0, T; L^2(\Omega))$ and $y_d \in L^2(0, T; L^2(\Omega))$, respectively. The initial condition is also defined as $y_0(x) \in H_0^1(\Omega)$. The diffusion and reaction coefficients are $\epsilon > 0$ and $r \in L^\infty(\Omega)$, respectively. The velocity field $\beta \in (W^{1,\infty}(\Omega))^2$ satisfies the incompressibility condition, i.e. $\nabla \cdot \beta = 0$. Furthermore, we assume the existence of the constant C_0 such that $r \geq C_0$ a.e. in Ω so that well-posedness of the optimal control problem (1.1) is guaranteed.

Secondly, we solve the following distributed optimal control problem governed by viscous Burgers equation without control constraints

$$\underset{u \in L^2(0,T;L^2(\Omega))}{\text{minimize}} \quad J(y, u) := \int_0^T \left(\frac{1}{2} \|y - y^d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 \right) dt, \quad (1.2a)$$

subject to

$$\partial_t y - \epsilon y_{xx} + yy_x = f + u, \quad x \in \Omega, \quad t \in I, \quad (1.2b)$$

$$y(0, t) = y(1, t) = 0, \quad t \in I, \quad (1.2c)$$

$$y(x, 0) = y_0(x), \quad x \in \Omega, \quad (1.2d)$$

where $\Omega = (0, 1)$ and $I = (0, T)$. Here, $f, y^d \in L^2(0, T; L^2(\Omega))$, $y_0(x) \in L^2(\Omega)$ are the source function, the desired state and the initial condition with $\epsilon > 0$.

Here, the state and the control are denoted by y and u , respectively. The aim is to minimize the difference between the state y and the desired state y_d by finding a minimum control u . Optimal control of evolutionary DCR equation has been investigated numerically and theoretically, especially when the convective term is more dominant than the diffusive term. In such cases, numerical solution exhibit interior/boundary layers and classical finite element method (FEM) is not efficient for derivation of an accurate numerical solution. To tackle such problems, several methods have been developed. For example, symmetric stabilization strategies [23], characteristic finite element method [57], streamline Petrov/Galerkin method [34], discontinuous Galerkin (DG) method [91] and adaptive mesh refinement strategies [142] are employed for this problem to substitute with the FEM. As a different strategy, we investigate these problems using variational time discretization combined with symmetric interior penalty Galerkin (SIPG) with upwinding in space [3].

Numerical treatment of OCPs consists of two ingredients. One of them is the optimization algorithm, which is used to find the optimum control; the other one is the discretization method which converts the continuous/infinite dimensional problem into a discrete/finite dimensional one. In [16], continuous and discontinuous Galerkin method are applied to continuous OCPs and optimization techniques for an efficient solution of parabolic OCPs are compared. A posteriori error estimates are given in [100] for parabolic optimization problems using space-time FEM. In [101, 102, 117], a priori error estimates are derived for this problem for unconstrained and control-constrained problems. Discontinuous in time and conforming elements in space are used for discretization of the parabolic OCPs in [29, 30] and theoretical estimates are proven. As a different and accurate strategy [3], we utilize variational time discretization method and symmetric interior penalty Galerkin (SIPG) with upwinding in space [13, 14, 112].

Discontinuous Galerkin method flexes the continuity condition of the approximate solution, that is, the trial space and the test space consist of polynomials which are discontinuous at the end points of each interval for one-dimensional case or discontinuous along the edges of the triangles/rectangles for two-dimensional case [112]. In addition, the boundary conditions can be imposed weakly which is not the case for continuous finite element method. On the other hand, this method can be applied to complicated meshes with unstructured grids or hanging nodes. As opposed to the finite difference or finite element methods, DG method are locally mass conservative which gains importance solving coupled flow and transport problems in porous media.

Variational time discretization technique is based on Galerkin projection on the time domain where the test functions are piecewise polynomials which are discontinuous at the endpoints of each time interval. Depending on the choice of the trial space, method is called differently. For continuous Galerkin-Petrov (cGP($q+1$)) method, solution space consists of continuous piecewise polynomials of degree $q+1$ and test functions are piecewise discontinuous polynomials of degree q . For discontinuous Galerkin dG(q) methods, both of the test and the trial spaces are piecewise discontinuous polynomials of degree q [120, Chap. 12]. Advantages of variational time discretization are stability, nodal superconvergence, and applicability of space-time adaptivity. Both continuous and discontinuous Galerkin methods are A -stable; discontinuous Galerkin methods are even L -stable (strongly stable). Convergence order of cGP($q+1$) methods are of one order higher than dG(q) methods. A priori error estimates of optimal order can be obtained with respect to the size of time steps and regularity requirements of the solutions [49], whereas dG methods require less regular solutions than the cG methods. dG(q) methods are superconvergent at the nodal points of order $2q + 1$ when the order of the method is q and the solution of the problem is sufficiently regular. For dG method, long-time integration is achievable without accumulating the error on the current time interval [49].

There exist two different approaches for solving OCPs: *optimize-then-discretize* (OD) and *discretize-then-optimize* (DO). In the OD approach, through continuous Lagrangian, infinite dimensional optimality system is derived containing the state and the adjoint equations and the variational inequality. Then, optimality system is discretized by using a suitable discretization method in space and time. In DO approach, infinite dimensional OCP is discretized and then finite-dimensional optimality system is derived by constructing the discrete Lagrangian. DO and OD approaches do not commute in general for OCPs governed by diffusion-convection-reaction equation [34]. However, commutativity is achieved in the case of SIPG discretization for steady state problems [91]. For discontinuous Galerkin time discretization, where both trial and test spaces are discontinuous, we show that OD and DO approaches commute, i.e. the adjoint state is discretized as we do for the state variable. For continuous Galerkin time discretization, where trial spaces are continuous and test spaces are discontinuous, OD and DO approaches do not commute.

An accurate and stable numerical solution requires a fine grid/mesh, which increases the dimension of the discrete problem, so the computational time. In case of perturbations in the data, full-order model (FOM) is required to be solved for each new parameter in the data set. In case of optimization problems, FOM associated to the differential equations must be resolved after updating the control. Therefore, a method, that eliminates the necessity of the solution of the FOM for each parameter or that enables us to solve the problem in a fast way, is required. Here, model-order reduction (MOR) techniques are used to replace the FOM with the reduced-order model (ROM). In the literature, there are several MOR techniques and each of them has different advantages and disadvantages depending on the problem at hand. One of the most popular and successful MOR techniques is the proper orthogonal decomposition (POD) method, also known as Karhunen-Loève decomposition or principal component analysis [71, 83, 133]. The idea behind the POD method is to derive a new basis spanning the subspace whose dimension is lower than the finite element space. Then, FOM

is projected onto the low-dimensional subspace using the new optimal POD basis as we proceed in Galerkin projection. POD basis is computed using the full-order solution at discrete time steps, which are called as the snapshots. Because the snapshots might be linearly dependent or almost linearly dependent, they cannot be directly used as a basis. Therefore, we construct a snapshot matrix and employ singular value decomposition or eigenvalue decomposition to derive the POD basis. In the literature, there are several studies concerning the reduced solution of the PDE-constrained optimization based on POD. However, the studies on the optimal control of diffusion-convection-reaction equation, compared to parabolic PDEs, or Burgers equation are not so popular cases. Although there are some studies analysing the error in the reduced solution for continuous finite element method and finite difference approximation in time, a priori error estimates based on space-time dG method is a new topic, where several differences arise during the derivation of the estimates and in the resulting error bound.

POD basis is computed using the snapshots of a particular problem which is interpreted by a mathematical model and data. Because there is a link between the data and the snapshots, some perturbation in the data may lead to larger changes in the snapshots depending on the problem at hand. This leads the nominal/baseline POD basis, which depends on the nominal/baseline parameters, not to approximate the perturbed problem accurately. In such cases, one has to solve the full problem for each parameter in the data set again and regenerate the POD basis. This approach is expensive especially for nonlinear problems or optimal control problems which requires the solution of a set of differential equations. Another choice is to use the sensitivities of the trajectories which require the solution of the sensitivity equations. They can be obtained by applying continuous sensitivity equation (CSE) method or finite difference (FD) approximation. Because sensitivity equations are always linear, the former method is especially preferable for nonlinear problems. The latter one requires the computation of the full problem at least one more time, so it is expensive for nonlinear case. In this study, motivated by the papers [62, 63, 64], POD sensitivities are used to enrich the low-dimensional subspace for a wider range of parameters and the quantity of interest is the diffusion term ϵ , the convective term β and the reactive term r . We generate two new bases, i.e. extrapolated POD (ExtPOD) and expanded POD (ExpPOD) and compare these bases in terms of advantages and discuss the main drawbacks of them.

The goal of this thesis is to develop a new discretization scheme for optimal control of evolutionary diffusion convection reaction equation and Burgers equation, derive error estimates for the error between the exact and the fully-discrete solution requiring less regularity of the exact solution. In addition, these problems are solved using a different basis, called proper orthogonal decomposition basis, which enables us to solve the problem in a fast way. Then, error estimates for the error between the exact and the reduced-order solution are proven based on space-time discontinuous Galerkin discretization. In addition, sensitivity information is used to increase the accuracy of nominal POD basis in case of perturbation of the data. To the best of our knowledge, this is the first study combining the sensitivity analysis techniques and model-order reduction strategies for PDE-constrained optimization problems. This thesis is organized as follows:

In Chapter 2, we explain how to derive the optimality conditions of the associated

problems, after stating the spaces and norms used in the study. Then, we proceed with the optimization algorithm by writing the OCPs in an abstract setting and explain the derivation of the first-order and second-order derivatives. Then, the computation of the gradient and the Hessian-times-vector product are formulated. Newton conjugate gradient method, which is employed to find the optimum control, is explained step by step.

In Chapter 3, we explain the symmetric interior penalty Galerkin method, which is used for spatial discretization, for one and two-dimensional problems separately. For linear problem, the upwind scheme is used to discretize the convection term; for non-linear case, Lax-Friedrichs flux is employed. Then, we derive semi-discrete optimality system in weak form and in matrix-vector form.

In Chapter 4, variational time discretization methods, namely continuous Galerkin-Petrov and discontinuous Galerkin methods, are introduced. Fully-discrete optimality systems are derived and the differential equations which must be solved to compute the gradient and the Hessian-times-vector product are specified. Then, we discuss similarities and differences of these two time discretization methods in terms of two approaches, namely OD and DO. We proceed with a priori error estimates for DCR equation, particularly for space-time DG method. The resulting error bound is suboptimal in space and optimal in time. The linear systems associated to the fully-discrete optimality system are computed and numerical results are presented. Then, we derive a priori error estimates for OCP of Burgers equation, which are suboptimal in space and optimal in time. The matrix systems for linearised state equation and the adjoint equation are formulated, and numerical results are given.

In Chapter 5, we use model-order reduction technique to reduce the dimension of the OCPs using proper orthogonal decomposition method. Here, solution of the full-order model is used and a new basis depending on the baseline parameters is computed. Then, Galerkin projection leads to weak form of each equation in the optimality system in a low-dimensional subspace. We explain continuous and discrete POD method and mention POD truncation error. Semi-discrete and fully-discrete optimality systems are written in this space and associated matrix forms are given, firstly, for linear PDE constraint. Then, a priori error estimates for reduced-order solution is derived based on space-time DG method and numerical results are presented to test the resulting error bound. Then, we proceed with the discrete empirical interpolation method, which is used to increase the efficiency of POD in case of nonlinear state equation. A priori error estimates for reduced solution of the Burgers equation follow and numerical results are discussed.

In Chapter 6, the aim is to increase the robustness of POD method in case of parameter perturbation. We utilize the sensitivity information of the snapshot matrix and compute POD basis sensitivities. Then, two new bases, namely extrapolated and expanded POD basis, are computed. Numerical results indicate the efficiency of this approach for PDE-constrained optimization.

Lastly, we draw a conclusion to underline vital parts of each study and comment on possible future work.

CHAPTER 2

OPTIMAL CONTROL PROBLEMS

Optimal control problems have several applications, such as petroleum reservoir simulation [78], environmental modeling [38], control of fluid flow [69, 126] and parameter estimation [125]. They are formulated through an optimization problem governed by a PDE. Numerical treatment of PDE-constrained optimization problems requires an optimization algorithm and a discretization method. Since we have a PDE-constrained optimization problem, one firstly attempts to construct the Lagrangian of the problem. Depending on the structure of the Lagrangian, i.e. continuous or discrete; two approaches, namely *optimize-then-discretize* (OD) and *discretize-then-optimize* (DO), are employed to find a numerical solution of the OCPs. The type of PDE constraint and the symmetric properties of the numerical discretization method might result in the same variational formulations [34, 91]. Depending on the problem, differences of these approaches might affect the accuracy and the rate of convergence.

OCPs governed by parabolic PDEs have gained interest, recently [12]. For example, this problem is solved by space-time multigrid techniques while discretization is performed with finite difference scheme in [21]. A nested multigrid method is utilized for time-periodic parabolic OCPs in [1]. OCP of one-dimensional heat equation with pointwise control constraints is investigated theoretically in [113], while optimal control of two-dimensional parabolic problems is studied in [93]. In addition to these, unconstrained and control-constrained OCPs governed by parabolic PDEs are solved with different types of control discretizations and error estimates are derived in [101, 102]. Finite volume method is combined with backward Euler time discretization and error estimates in different norms are derived in [97]. A variational discretization scheme is suggested in [68] where the relation between the adjoint and the control is used and the space of controls is not discretized which leads to an increase in the rate of convergence. A priori error estimates for discontinuous Galerkin time discretization for unconstrained parabolic OCPs are proposed in [29].

In case of optimal control of steady diffusion-convection-reaction (DCR) equation, symmetric stabilization has been integrated with finite element method and approaches OD and DO lead to the same optimality condition in [17]. Variational discretization concept is applied to this problem in [73]. Numerical result of this problem is derived using symmetric interior penalty Galerkin discretization in [92, 139], while streamline upwind Petrov/Galerkin method is used in [137]. When the convective term dominates the diffusion term, the problem becomes highly convection dominated and it might

contain boundary or interior layers. For such cases, adaptive refinement strategies can be used to find an accurate numerical solution. For example, local DG method is employed for this problem in [142] and a posteriori error estimates are proven, while a residual type error estimator with symmetric interior penalty Galerkin method is derived in [140].

Studies on OCPs governed by evolutionary DCR equations are quite rare when compared to steady-state case. For example, local DG approximation of the OCP which is discretized by backward Euler in time is studied in [141] and a priori error estimates for semi-discrete OCP is provided in [119]. In [57, 58], characteristic finite element solution of the OCP is discussed and numerical results are provided using uniform and adaptive refinement strategies, respectively. Crank-Nicolson time discretization with symmetric stabilization is applied in [23]. Control-constrained case is investigated in one of our previous study [5] for SIPG method combined with backward Euler discretization, while we investigated variational time discretization methods in [3].

In addition, OCP governed by Burgers equation is studied numerically and theoretically. For example, distributed and boundary control problem are solved using discontinuous Galerkin method in space and the fourth order accurate Runge-Kutta time integration in [28], while finite element method combined with Crank-Nicolson time discretization is used in [65]. In addition, augmented Lagrangian sequential quadratic programming (SQP) method is investigated by combining the theoretical analysis with some numerical results in [131, 132]. Lagrange-Newton SQP method is discussed for control-constrained OCP and the convergence of the method is proven and the study is enriched with numerical examples in [123]. Three different algorithms for control of Burgers equation with point-wise constraints are investigated in [35], while different adjoint techniques for Neumann boundary control are discussed in [104] and two time discretization methods, i.e. backward Euler and Crank-Nicolson, are compared.

In this chapter, we firstly mention the spaces and the norms used in this study in Sec. 2.1. Then, we derive the optimality conditions for OCP governed by DCR equation and Burgers equation, respectively in Sec. 2.2. Then, we explain the optimization method and the equations solved in each step in an abstract form in Sec. 2.3.

2.1 Preliminaries

Let Ω be a bounded domain in \mathbb{R}^1 for one-dimensional problems and let it be a convex polygonal domain in \mathbb{R}^2 with Lipschitz boundary $\partial\Omega$ for two-dimensional problems.

Lebesgue space $L^p(\Omega)$ for $1 \leq p < \infty$ is defined as

$$L^p(\Omega) = \{v \text{ Lebesgue measurable} : \|v\|_{L^p(\Omega)} < \infty\},$$

where

$$\|v\|_{L^p(\Omega)} = \left(\int_{\Omega} |v(x)|^p \right)^{1/p}.$$

In this study, particularly for $p = 2$, inner product in $L^2(\Omega)$ is denoted by (\cdot, \cdot) and L^2 -norm is written as $\|\cdot\|_{L^2(\Omega)} = \|\cdot\|$.

The space $L^\infty(\Omega)$ is defined as

$$L^\infty(\Omega) = \{v \text{ Lebesgue measurable} : \|v\|_{L^\infty(\Omega)} < \infty\},$$

where

$$\|v\|_{L^\infty(\Omega)} = \text{ess sup} \{|v(x)| : \text{a.e. } x \in \Omega\}.$$

We introduce L^2 inner product on the inflow or outflow boundaries as follows

$$(w, v)_{\Gamma^-} = \int_{\Gamma^-} |\boldsymbol{\beta} \cdot \mathbf{n}| w v \, ds$$

with analogous definition of $(\cdot, \cdot)_{\Gamma^+}$ and associated norms $\|\cdot\|_{\Gamma^-}$ and $\|\cdot\|_{\Gamma^+}$.

We proceed with Sobolev space denoted by $W^{k,p}(\Omega)$. Let $k \geq 0$ be an integer and $1 \leq p \leq \infty$ be a real number. Sobolev space is defined as

$$W^{k,p}(\Omega) = \{v \in L^p(\Omega) : \forall \alpha \in D_d^k, \partial^\alpha v \in L^p(\Omega)\},$$

where $D_d^k = \{\alpha \in \mathbb{N}^d : |\alpha| \leq k\}$. Equivalently, functions with derivatives up to order k in $L^p(\Omega)$ span the space $W^{k,p}(\Omega)$.

Sobolev space $W^{k,p}(\Omega)$ is a Banach space equipped with the norm

$$\|v\|_{W^{k,p}(\Omega)} = \left(\sum_{\alpha \in D_d^k} \|\partial^\alpha v\|_{L^p(\Omega)}^p \right)^{1/p}, \quad 1 \leq p < \infty,$$

and $\|v\|_{W^{k,\infty}(\Omega)} = \max_{\alpha \in D_d^k} \|\partial^\alpha v\|_{L^\infty(\Omega)}$.

We define the semi-norm $|\cdot|_{W^{k,p}}$ by fixing the condition $\bar{D}_d^k = \{\alpha \in \mathbb{N}^d : |\alpha| = k\}$, which means that only the derivative of order k is needed.

In particular, for $p = 2$, $H^k(\Omega) = W^{k,2}(\Omega)$, which is a Hilbert space, is written as

$$H^k(\Omega) = \{v \in L^2(\Omega) : \forall \alpha \in \bar{D}_d^k, \partial^\alpha v \in L^2(\Omega)\}.$$

On this space, we define the norm and semi-norm as follows:

$$\|v\|_{H^k(\Omega)} = \left(\sum_{\alpha \in D_d^k} \|\partial^\alpha v\|^2 \right)^{1/2}, \quad |v|_{H^k(\Omega)} = \left(\sum_{\alpha \in \bar{D}_d^k} \|\partial^\alpha v\|^2 \right)^{1/2}.$$

The broken Sobolev space is defined as

$$H^k(\Omega, \mathcal{T}_h) = \{v : v|_K \in H^k(K) \quad \forall K \in \mathcal{T}_h\},$$

with the semi-norm defined by

$$|v|_{H^k(\Omega, \mathcal{T}_h)} = \left(\sum_{K \in \mathcal{T}_h} |v|_{H^k(K)}^2 \right)^{1/2}, \quad v \in H^k(\Omega, \mathcal{T}_h).$$

We use DG energy norm given in [130, Section 4]

$$|||v|||_{DG}^2 = |v|_{H^1(\Omega, \mathcal{T}_h)}^2 + J_\sigma(v, v). \quad (2.1)$$

Further, we consider the space of functions mapping the time interval $(0, T)$ to a normed space X in which the norm $\|\cdot\|_X$ is defined. For $r \geq 1$, we define

$$L^r(0, T; X) = \{z : [0, T] \rightarrow X \text{ measurable} : \int_0^T \|z(t)\|_X^r dt < \infty\}$$

with

$$\|z(t)\|_{L^r(0, T; X)} = \begin{cases} \left(\int_0^T \|z(t)\|_X^r dt \right)^{1/r}, & \text{if } 1 \leq r < \infty, \\ \text{ess sup}_{t \in (0, T)} \|z(t)\|_X, & \text{if } r = \infty. \end{cases}$$

Bochner space of functions whose k th time derivative is bounded almost everywhere on $(0, T)$ with values in X is denoted by $W^{k, \infty}(0, T; X)$.

We need the discrete time-dependent norms for $1 \leq q < \infty$,

$$\|v\|_{L^q(0, T; L^2(\Omega))} = \left(\frac{1}{N} \sum_{i=1}^N \|v_i\|_{L^2(\Omega)}^q \right)^{1/q}, \quad \|v\|_{L^2(0, T; DG)} = \left(\frac{1}{N} \sum_{i=1}^N |||v_i|||_{DG}^2 \right)^{1/2}. \quad (2.2)$$

We give the multiplicative trace inequality for all $K \in \mathcal{T}_h$, for all $v \in H^1(K)$ as follows:

$$\|v\|_{L^2(\partial K)}^2 \leq C_M \left(\|v\|_{L^2(K)} \|v\|_{H^1(K)} + h_K^{-1} \|v\|_{L^2(K)}^2 \right), \quad (2.3)$$

where C_M is a positive constant independent of v, h and K . We refer the reader to the study [43, Lemma 3.1] for the proof.

In addition, generalization of Poincaré inequality to the broken Sobolev space $H^1(\Omega, \mathcal{T}_h)$ is given as [112, Section 3.1.4]

$$\|v\|_{L^2(\Omega)}^2 \leq C_S \left(|v|_{H^1(\Omega, \mathcal{T}_h)}^2 + \sum_{E \in \mathcal{E}_h} \frac{1}{h_E} \|[[y]]\|_{L^2(E)}^2 \right). \quad (2.4)$$

We mention the discrete trace inequality for all $K \in \mathcal{T}_h$, for all $v \in V_{h,p}$ in [40, Lemma 1.46]:

$$\|v\|_{L^2(\partial K)} \leq C_{tr} h^{-1/2} \|v\|_{L^2(K)}, \quad (2.5)$$

where C_{tr} is a constant independent of v, h and K , but it depends on the polynomial degree p .

We mention Cauchy-Schwarz's and Young's inequalities which are frequently used for the derivation of the error estimates.

- Cauchy-Schwarz's inequality:

$$\forall f, g \in L^2(\Omega), \quad |(f, g)| \leq \|f\| \|g\|.$$

- Young's inequality:

$$\forall \epsilon > 0, \forall \alpha, \beta, \quad \alpha\beta \leq \frac{\epsilon}{2}\alpha^2 + \frac{1}{2\epsilon}\beta^2.$$

We recall continuous and discrete Gronwall inequalities:

- Continuous Gronwall inequality:

Suppose that f, g, h are piecewise continuous nonnegative functions defined on (a, b) and let g be nondecreasing. If there is a positive constant C independent of t such that

$$\forall t \in (a, b), \quad f(t) + h(t) \leq g(t) + C \int_a^t f(s) ds,$$

then

$$\forall t \in (a, b), \quad f(t) + h(t) \leq \exp C(t - a)g(t).$$

- Discrete Gronwall inequality:

Suppose that $\Delta t, B, C > 0$ and $(a_n)_n, (b_n)_n, (c_n)_n, (d_n)_n$ are sequences with nonnegative numbers satisfying the relation

$$\forall n \geq 0, \quad a_n + \Delta t \sum_{i=0}^n b_i \leq B + C\Delta t \sum_{i=0}^n a_i + \Delta t \sum_{i=0}^n c_i.$$

Then, under the condition $C\Delta t \leq 1$,

$$\forall n \geq 0, \quad a_n + \Delta t \sum_{i=0}^n b_i \leq \exp C(n + 1)\Delta t \left(B + \Delta t \sum_{i=0}^n c_i \right).$$

2.2 Optimality Conditions

In this section, we obtain the optimality conditions for OCP of DCR equation and Burgers equation, respectively.

2.2.1 Optimality Conditions for Diffusion Convection Reaction Equation

Firstly, we consider the problem (1.1). We fix trial and test spaces as $Y = V = H_0^1(\Omega)$, $\forall t \in (0, T]$. We construct the continuous Lagrangian as follows:

$$\begin{aligned} \mathcal{L}(y, u, \lambda) &= \int_0^T \left(\frac{1}{2} \|y - y^d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 \right) dt \\ &\quad + (\partial_t y, \lambda) + a(y, \lambda) - (f + u, \lambda) + (y(x, 0) - y_0, \lambda(x, 0)). \end{aligned}$$

By differentiating the Lagrangian with respect to λ, y and u , we obtain the following (the first order necessary) optimality condition or Karush-Kuhn-Tucker conditions:

$$(\partial_t y, v) + a(y, v) = (f + u, v) \quad \forall v \in V, \quad y(x, 0) = y_0, \quad (2.6a)$$

$$-(\partial_t \lambda, \psi) + a(\psi, \lambda) = -(y - y_d, \psi) \quad \forall \psi \in V, \quad p(x, T) = 0, \quad (2.6b)$$

$$\alpha u = \lambda, \quad (2.6c)$$

with

$$a(y, v) = \int_{\Omega} (\epsilon \nabla y \cdot \nabla v + \beta \cdot \nabla y v + r y v) dx, \quad (w, v) = \int_{\Omega} w v dx.$$

It is well known that the functions $(y, u) \in H^1(0, T; L^2(\Omega)) \cap L^2(0, T; Y) \times L^2(0, T; L^2(\Omega))$ solve (1.1) if and only if there is an adjoint $\lambda \in H^1(0, T; L^2(\Omega)) \cap L^2(0, T; Y)$ such that (y, u, λ) is the unique solution of the following optimality system [122].

In this optimality system (2.6), the first equation is the weak form of the state equation in (1.1). The second one is the adjoint equation which is also a PDE with the negative convection field. In addition, initial condition of this equation is defined at $t = T$ and it is solved backward in time. PDE corresponding to (2.6b) is written as follows:

$$-\partial_t \lambda - \epsilon \Delta \lambda - \beta \cdot \nabla \lambda + r \lambda = -(y - y_d) \quad (x, t) \in \Omega \times (0, T], \quad (2.7a)$$

$$\lambda(x, t) = 0 \quad (x, t) \in \partial \Omega \times [0, T], \quad (2.7b)$$

$$\lambda(x, T) = 0 \quad x \in \Omega. \quad (2.7c)$$

2.2.2 Optimality Conditions for Burgers Equation

Secondly, we proceed with OCP of Burgers equation (1.2). We fix state and test space are $Y = V = H_0^1(\Omega)$ and we define the space

$$W(0, T) = \{v \in L^2(0, T; V) : v_t \in L^2(0, T; V')\}.$$

We construct the continuous Lagrangian as follows:

$$\begin{aligned} \mathcal{L}(y, u, \lambda) &= \int_0^T \left(\frac{1}{2} \|y - y^d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 \right) dt \\ &\quad + (\partial_t y, \lambda) + a(y, \lambda) + n(y, \lambda) - (f + u, \lambda) + (y(x, 0) - y_0, \lambda(x, 0)). \end{aligned}$$

By differentiating the Lagrangian with respect to λ, y and u , we obtain the following (the first order necessary) optimality condition:

$$(\partial_t y, v) + a(y, v) + n(y, v) = (f + u, v), \quad \forall v \in V, \quad y(x, 0) = y_0, \quad (2.8a)$$

$$-(\partial_t \lambda, q) + a(q, \lambda) - n^\lambda(\lambda, y, v) = -(y - y^d, q), \quad \forall q \in V, \quad p(x, T) = 0, \quad (2.8b)$$

$$\alpha u = \lambda, \quad (2.8c)$$

where

$$a(y, v) = \int_{\Omega} \epsilon y_x v_x \, dx, \quad n(y, v) = \int_{\Omega} y y_x v \, dx, \quad n^\lambda(\lambda, y, v) = \int_{\Omega} y \lambda_x v \, dx.$$

In the optimality system (2.8), the first equation is the weak form of the state equation in (1.2). The second one is the adjoint equation which is written as follows:

$$-\partial_t \lambda - \epsilon \lambda_{xx} - y \lambda_x = -(y - y^d) \quad (x, t) \in \Omega \times (0, T], \quad (2.9a)$$

$$\lambda(x, t) = 0 \quad (x, t) \in \partial\Omega \times [0, T], \quad (2.9b)$$

$$\lambda(x, T) = 0 \quad x \in \Omega. \quad (2.9c)$$

The pair $(y, u) \in W(0, T)$ together with an adjoint $\lambda \in W(0, T)$ must satisfy the first-order necessary optimality conditions [123]. To guarantee that the pair (y, u) is a local solution to (1.2), we assume that the second order optimality condition holds.

Therefore, we obtain the second Fréchet derivative of the Lagrangian with respect to $x = (y, u)$ in the direction of $h_i = (y_i, u_i)$ for $i = 1, 2$ as

$$\mathcal{L}''(y, u, \lambda)(h_1, h_2) = \int_0^T (y_1 y_2 + u_1 u_2 - \lambda_x y_1 y_2) \, dx dt.$$

The second-order sufficient optimality condition assures that [123]

$$\mathcal{L}''(y^*, u^*, \lambda^*)((y, u), (y, u)) \geq \kappa \|u\|_{L^2(0, T; L^2(\Omega))}^2,$$

for all $(y, u) \in W(0, T) \times L^2(0, T; L^2(\Omega))$ where the state y solves the following linearised PDE:

$$\begin{aligned} \partial_t y - \epsilon y_{xx} + (y^* y)_x &= u, & x \in \Omega, & \quad t \in I, \\ y(0, t) = y(1, t) &= 0, & t \in I, \\ y(x, 0) &= y_0(x), & x \in \Omega. \end{aligned}$$

2.3 Optimization Algorithm

We write the OCP (1.1) and (1.2) in an abstract form as in (2.10). The cost functional J is minimized over the control u where the relation between the state y and the control u is satisfied through a partial differential equation $c(y, u)$:

$$\min_{u \in U} J(y, u), \quad (2.10a)$$

$$\text{subject to } c(y, u) = 0. \quad (2.10b)$$

Alternatively, the minimization problem (2.10) can be rewritten as an unconstrained optimization problem as follows:

$$\min_{u \in U} \widehat{J}(u) = J(y(u), u). \quad (2.11)$$

The solution of the optimization problem (2.11) can be achieved through different solvers (see, e.g. [39, 80, 105]). Apart from those, Newton-Conjugate Gradient method enables us to find an approximate solution to Newton equation $\nabla^2 \widehat{J}(u_k) s_k = -\nabla \widehat{J}(u_k)$ by applying conjugate gradient (CG) method with initiating a control, i.e. u_0 . The terms s_k , $\nabla \widehat{J}(u_k)$, $\nabla^2 \widehat{J}(u_k)$ correspond to the search direction, the gradient and the Hessian of the reduced cost functional, respectively. CG method is applied until the residual of the Newton system is small enough, namely,

$$\|\nabla^2 \widehat{J}(u_k) s_k + \nabla \widehat{J}(u_k)\|_2 \leq \eta_k \|\nabla \widehat{J}(u_k)\|_2$$

with $\eta_k \in (0, 1)$ or a direction of negative curvature is found. After determining a search direction s_k , the step size α is computed through Armijo line-search method. Then, the control is updated as $u_{k+1} = u_k + \alpha s_k$ [105]. We give the details of Newton-CG method with Armijo line-search in Algorithm 2.1.

Following [65], we explain how to derive the gradient and perform Hessian times vector multiplication.

2.3.1 First-order Derivatives

For an efficient computation of the gradient, we construct the Lagrangian of the problem (2.10) so that we have an unconstrained optimization problem at hand.

$$\mathcal{L}(y, u, \lambda) := J(y, u) + \lambda^T c(y, u), \quad (2.12)$$

with Lagrange multiplier λ . The first order necessary optimality condition in terms of the reduced cost functional $\widehat{J}(u)$ guarantees the existence of a triple (y, u, λ) which solves the optimality system:

$$\mathcal{L}_\lambda(y, u, \lambda)|_{(y:=y(u), \lambda:=\lambda_u)} = 0, \quad (\text{State equation}), \quad (2.13a)$$

$$\mathcal{L}_y(y, u, \lambda)|_{(y:=y(u), \lambda:=\lambda_u)} = 0, \quad (\text{Adjoint equation}), \quad (2.13b)$$

$$\mathcal{L}_u(y, u, \lambda)|_{(y:=y(u), \lambda:=\lambda_u)} = 0, \quad (\text{Gradient equation}). \quad (2.13c)$$

The first equation (2.13a) gives the PDE-constraint, i.e., $c(y, u) = 0$. Differentiating the state equation, we obtain the relation

$$c_y(y(u), u) y_u(u) + c_u(y(u), u) = 0. \quad (2.14)$$

From this relation, the derivative $y_u(u)$ or the sensitivity of y with respect to u is represented by

$$y_u(u) = -c_y(y(u), u)^{-1} c_u(y(u), u). \quad (2.15)$$

Algorithm 2.1 Newton-Conjugate Gradient Method with Armijo Line-Search

Require: Initial control u_0 , tolerance $\epsilon_{gtol} > 0$.

```
1: for  $k = 0$  to  $k_{\max}$  do
2:   Compute  $\nabla \hat{J}(u_k)$ .
3:   if  $\|\nabla \hat{J}(u_k)\| < \epsilon_{gtol}$  then return
4:   end if
5:   Compute  $\nabla^2 \hat{J}(u_k)$ .
6:   Solve the Newton equation  $\nabla^2 \hat{J}(u_k) s_k = -\nabla \hat{J}(u_k)$  using CG method:
Require: Set  $\eta_k \in (0, 1)$ ,  $s_k = 0$  and  $p_{k,0} = r_{k,0} = -\nabla \hat{J}(u_k)$ .
7:   for  $i = 0, 1, 2, \dots$  to  $i_{\max}$  do
8:     if  $\|r_{k,i}\|_2 < \eta_k \|r_{k,0}\|_2$  then
9:       go to Line (21).
10:    end if
11:    Compute  $q_{k,i} = \nabla^2 \hat{J}(u_k) p_i$ .
12:    if  $p_{k,i}^T q_{k,i} < 0$  then
13:      go to Line (21).
14:    end if
15:     $\gamma_{k,i} = \|r_{k,i}\|^2 / p_{k,i}^T q_{k,i}$ .
16:     $s_k = s_k + \gamma_{k,i} p_{k,i}$ .
17:     $r_{k,i+1} = r_{k,i} - \gamma_{k,i} q_{k,i}$ .
18:     $\beta_{k,i} = \|r_{k,i+1}\|^2 / \|r_{k,i}\|^2$ .
19:     $p_{k,i+1} = r_{k,i+1} + \beta_{k,i} p_{k,i}$ .
20:    if  $i = 0$  then
21:      set  $s_k = -\nabla \hat{J}(u_k)$ .
22:    end if
23:  end for
24:  Perform Armijo line-search.
Require: Set  $\alpha_k = 1$  and evaluate  $f(u_k + \alpha_k s_k)$ 
25:  while  $f(u_k + \alpha_k s_k) > f(u_k) + 10^{-4} \alpha_k s_k^T \nabla \hat{J}(u_k)$  do
26:    Set  $\alpha_k = \alpha_k / 2$  and evaluate  $f(u_k + \alpha_k s_k)$ .
27:  end while
28:  Set  $u_{k+1} = u_k + \alpha_k s_k$ .
29: end for
```

The second equation (2.13b) is equivalent to

$$c_y(y(u), u)^T \lambda + \nabla_y J(y(u), u) = 0. \quad (2.16)$$

The third equation (2.13c) can be written as

$$\begin{aligned} \nabla \widehat{J}(u) &= \nabla_u J(y(u), u) + y_u(u)^T \nabla_y J(y(u), u) \\ &\stackrel{(2.15)}{=} \nabla_u J(y(u), u) - c_u(y(u), u)^T c_y(y(u), u)^{-T} \nabla_y J(y(u), u) \\ &\stackrel{(2.16)}{=} \nabla_u J(y(u), u) + c_u(y(u), u)^T \lambda(u). \end{aligned} \quad (2.17)$$

In Algorithm 2.2, the gradient computation for given control u is explained.

Algorithm 2.2 Gradient Computation Using Adjoint

- 1: Given u , solve $c(y, u) = 0$ for $y := y(u)$.
 - 2: Solve the adjoint equation $c_y(y(u), u)^T \lambda = -\nabla_y J(y(u), u)$ for $\lambda := \lambda(u)$.
 - 3: Compute $\nabla \widehat{J}(u) = \nabla_u J(y(u), u) + c_u(y(u), u)^T \lambda(u)$.
-

2.3.2 Second-order Derivatives

After deriving an expression for $\nabla \widehat{J}(u)$, we proceed with

$$\begin{aligned} \nabla^2 \widehat{J}(u) &= \mathcal{L}_{yy}(y(u), u, \lambda(u)) y_u(u) \\ &\quad + \mathcal{L}_{uu}(y(u), u, \lambda(u)) + \mathcal{L}_{u\lambda}(y(u), u, \lambda(u)) \lambda_u(u). \end{aligned} \quad (2.18)$$

Now, the steps required to compute Hessian times vector multiplication through adjoints is explained. Firstly, we differentiate the adjoint equation (2.13b) with respect to u to obtain

$$\mathcal{L}_{yy}(y(u), u, \lambda(u)) y_u(u) + \mathcal{L}_{yu}(y(u), u, \lambda(u)) + \mathcal{L}_{y\lambda}(y(u), u, \lambda(u)) \lambda_u(u) = 0. \quad (2.19)$$

Then, we note that

$$\mathcal{L}_{u\lambda}(y(u), u, \lambda(u)) = c_u(y(u), u)^T, \quad (2.20a)$$

$$\mathcal{L}_{y\lambda}(y(u), u, \lambda(u)) = c_y(y(u), u)^T. \quad (2.20b)$$

In addition,

$$\begin{aligned} \lambda_u(u) &= -\mathcal{L}_{y\lambda}(y(u), u, \lambda(u))^{-1} (\mathcal{L}_{yy}(y(u), u, \lambda(u)) y_u(u) + \mathcal{L}_{yu}(y(u), u, \lambda(u))) \\ &\stackrel{(2.20b)}{=} c_y(y(u), u)^{-T} (-\mathcal{L}_{yy}(y(u), u, \lambda(u)) y_u(u) - \mathcal{L}_{yu}(y(u), u, \lambda(u))) \\ &\stackrel{(2.15)}{=} c_y(y(u), u)^{-T} \mathcal{L}_{yy}(y(u), u, \lambda(u)) (c_y(y(u), u)^{-1} c_u(y(u), u)) \\ &\quad - c_y(y(u), u)^{-T} \mathcal{L}_{yu}(y(u), u, \lambda(u)) \end{aligned} \quad (2.21)$$

We consider the Hessian again

$$\begin{aligned}
\nabla^2 \widehat{J}(u) &= \mathcal{L}_{yy}(y(u), u, \lambda(u))y_u(u) + \mathcal{L}_{uu}(y(u), u, \lambda(u)) + \mathcal{L}_{u\lambda}(y(u), u, \lambda(u))\lambda_u(u) \\
&\stackrel{(2.20a)}{=} \mathcal{L}_{yy}(y(u), u, \lambda(u))y_u(u) + \mathcal{L}_{uu}(y(u), u, \lambda(u)) + c_u(y(u), u)^T \lambda_u(u) \\
&\stackrel{(2.21)}{=} \mathcal{L}_{yy}(y(u), u, \lambda(u))y_u(u) + \mathcal{L}_{uu}(y(u), u, \lambda(u)) \\
&\quad + c_u(y(u), u)^T c_y(y(u), u)^{-T} \mathcal{L}_{yy}(y(u), u, \lambda(u)) (c_y(y(u), u)^{-1} c_u(y(u), u)) \\
&\quad - c_u(y(u), u)^T c_y(y(u), u)^{-T} \mathcal{L}_{yu}(y(u), u, \lambda(u)) \\
&\stackrel{(2.15)}{=} -\mathcal{L}_{yy}(y(u), u, \lambda(u)) \underbrace{(c_y(y(u), u)^{-1} c_u(y(u), u))}_{:=w} + \mathcal{L}_{uu}(y(u), u, \lambda(u)) \\
&\quad + (c_u(y(u), u)^T c_y(y(u), u)^{-T} \mathcal{L}_{yy}(y(u), u, \lambda(u)) \underbrace{(c_y(y(u), u)^{-1} c_u(y(u), u))}_{:=w}) \\
&\quad - (c_u(y(u), u)^T c_y(y(u), u)^{-T}) \mathcal{L}_{yu}(y(u), u, \lambda(u)) \\
&\stackrel{(2.15)}{=} -\mathcal{L}_{yy}(y(u), u, \lambda(u))w + \mathcal{L}_{uu}(y(u), u, \lambda(u)) \\
&\quad + (c_u(y(u), u)^T \underbrace{c_y(y(u), u)^{-T} (\mathcal{L}_{yy}(y(u), u, \lambda(u))w - \mathcal{L}_{yu}(y(u), u, \lambda(u)))}_{:=p}) \\
&= -\mathcal{L}_{yy}(y(u), u, \lambda(u))w + \mathcal{L}_{uu}(y(u), u, \lambda(u)) + c_u(y(u), u)^T p. \tag{2.22}
\end{aligned}$$

The equality (2.22) reveals that the Hessian times vector multiplication, namely $\nabla^2 \widehat{J}(u)\nu$ can be computed without computing the Hessian which requires less computational time. In Algorithm 2.3, the computation of Hessian vector multiplication is given step by step.

Algorithm 2.3 Hessian-Times-Vector Computation

- 1: Given u , solve $c(y, u) = 0$ for $y := y(u)$.
- 2: Solve the adjoint equation $c_y(y(u), u)^T \lambda = -\nabla_y J(y(u), u)$ for $\lambda := \lambda(u)$.
- 3: Solve the equation $c_y(y(u), u)w = c_u(y, u)\nu$ for w .
- 4: Solve the equation

$$c_y(y(u), u)^T p = \nabla_{yy} \mathcal{L}(y(u), u, \lambda(u))w - \nabla_{yu} \mathcal{L}(y(u), u, \lambda(u))\nu$$

for p .

- 5: Compute

$$\nabla^2 \widehat{J}(u)\nu = c_u(y(u), u)^T p - \nabla_{yy} \mathcal{L}(y(u), u, \lambda(u))w + \nabla_{uu} \mathcal{L}(y(u), u, \lambda(u))\nu.$$

We note that the Hessian times vector multiplication requires the solution of the state equation, which might be a linear and nonlinear constraint. Other unknowns λ , w and p are derived through three linear equations. If one has the solution of the state and the adjoint equation at hand, then only the solutions of two linear equation are required.

CHAPTER 3

DISCONTINUOUS GALERKIN METHODS

Discontinuous Galerkin (dG) methods was firstly introduced by Reed and Hill in 1973 [110] for hyperbolic problems. Then, there has been an increasing trend in the studies discussing dG methods for elliptic and parabolic problems and some variations of this method have been invented. For example, local dG (LDG) method converts the original second order equation into a systems of first order equations and discretize it [33]. A variation of LDG method is compact dG method (CDG) where these two methods are equivalent for one-dimensional problem, while for high dimensional problems, computation of the approximate gradient differs. CDG is more stable and less sensitive to element orientation than LDG [108]. Bubble stabilized dG method is derived in [24]. Another scheme is explicit DG with local time stepping [96] where a space-time weak formulation and an explicit approximate solver is used as a predictor. For nonlinear problems, an implicit and explicit Runge-Kutta method and nonsymmetric interior penalty method is combined in [129]. For DCR equation, this method has been successfully applied to PDEs in [25, 14, 42, 46, 54, 75, 130].

The idea behind DG method is to flex the continuity condition of the approximate solution, that is, trial space and test space are discontinuous at the end points of each spatial interval in one-dimension or along the edges of the triangles/rectangles in two-dimension [112]. In addition, boundary conditions can be imposed weakly which is not the case for continuous finite element method. On the other hand, this method can be applied to complicated meshes with unstructured grids or hanging nodes. The order of approximation on each element might differ, so it is possible to use *hp* refinement strategies, that is, size of each element and the order of the finite element approximation can be decided adaptively on each element [25, 75]. As opposed to the finite difference or finite element methods, DG method are locally mass conservative which gains importance solving coupled flow and transport problems in porous media. On the other hand, size of the coefficient matrices is higher than the ones obtained from continuous FEM and the condition number of these matrices gets higher as we increase the order of the finite element approximation [112].

In this chapter, one-dimensional dG method is explained in Sec. 3.1. Then, two-dimensional case follows in Sec. 3.2. Associated semi-discrete variational formulations and matrix-vector systems are derived.

3.1 Discontinuous Galerkin Methods in One-Dimension

In this section, we briefly describe the interior penalty Galerkin semi-discretization in space for viscous Burgers equation:

$$\partial_t y - \epsilon y_{xx} + \frac{1}{2}(y)_x^2 = f, \quad x \in \Omega, \quad t \in I, \quad (3.1a)$$

$$y(0, t) = y(1, t) = 0, \quad t \in I, \quad (3.1b)$$

$$y(x, 0) = y_0(x), \quad x \in \Omega, \quad (3.1c)$$

where $\Omega = (0, 1)$, $I = (0, T)$. Here, $f, y^d \in L^2(0, T; L^2(\Omega))$, $y_0(x) \in L^2(\Omega)$ are given source function, desired state and initial condition with $\epsilon > 0$.

Let Ω be divided into M subintervals as $0 = x_0 < x_1 < \dots < x_M = 1$. We denote each partition by \mathcal{T}_h , each subinterval by $K := K_n = (x_n, x_{n+1})$ and the length of these subintervals by

$$h_n = x_{n+1} - x_n, \quad h_{n-1,n} = \max(h_{n-1}, h_n), \quad h = \max_{0 \leq n \leq M-1} h_n.$$

The jump and average of v can be defined for the endpoints of the subintervals as

$$[[v(x_n)]] = v(x_n^-) - v(x_n^+), \quad \{\!\!\{v(x_n)\}\!\!\} = \frac{1}{2}(v(x_n^-) + v(x_n^+)) \quad \forall n = 1, \dots, M-1,$$

where $v(x_n^+) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} v(x_n + \epsilon)$ and $v(x_n^-) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} v(x_n - \epsilon)$. Similarly, for a piecewise continuous v' , the jump and the average for the endpoints of the subintervals are given by

$$[[v'(x_n)]] = v'(x_n^-) - v'(x_n^+), \quad \{\!\!\{v'(x_n)\}\!\!\} = \frac{1}{2}(v'(x_n^-) + v'(x_n^+)).$$

These definitions can be extended to the end points of Ω as

$$[[v(x_0)]] = -v(x_0^+), \quad \{\!\!\{v(x_0)\}\!\!\} = v(x_0^+), \quad [[v(x_M)]] = v(x_M^-), \quad \{\!\!\{v(x_M)\}\!\!\} = v(x_M^-).$$

We use discontinuous piecewise finite element spaces to define the discrete test, state and control spaces

$$V_{h,p} = Y_{h,p} = U_{h,p} = \{v : v|_K \in P_p(K) \quad \forall K \in \mathcal{T}_h\}. \quad (3.2)$$

Here, $P_p(K)$ denotes the set of all polynomials of degree p on the interval K .

The diffusion term in the state equation (3.1) is discretized by the symmetric interior penalty method (SIPG) [13] and the nonlinear term is discretized by local Lax-Friedrichs flux [67, Sect. 2.3] in space for fixed control u . Then, the semi-discrete OCP is written as

$$\underset{u_h \in \tilde{V}_{h,p}}{\text{minimize}} \quad \frac{1}{2} \int_0^T \left(\sum_{K \in \mathcal{T}_h} (\|y_h - y_h^d\|_{L^2(K)}^2 + \alpha \|u_h\|_{L^2(K)}^2) \right) dt, \quad (3.3a)$$

$$\text{subject to} \quad (\partial_t y_h, v_h) + a_h^s(y_h, v_h) = (f_h + u_h, v_h), \quad \forall v_h \in V_{h,p}, \\ y_{h,0} = (y_0)_h, \quad (3.3b)$$

with the (bi-)linear forms

$$a_h(y, v) = \sum_{n=0}^{M-1} \int_{x_n}^{x_{n+1}} \epsilon y'(x, t) v'(x) dx \quad (3.4a)$$

$$- \sum_{n=0}^M \left(\left\{ \epsilon y'(x_n, t) \right\} \llbracket v(x_n) \rrbracket - \kappa \left\{ \epsilon v'(x_n) \right\} \llbracket y(x_n, t) \rrbracket - \epsilon \overbrace{\frac{\sigma}{h} \llbracket y(x_n, t) \rrbracket \llbracket v(x_n) \rrbracket}^{J_\sigma(y, v)} \right),$$

$$n_h^s(y, v) = \sum_{n=0}^{M-1} \left(- \int_{x_n}^{x_{n+1}} \frac{1}{2} y_h^2 v'(x) dx + v_{n+1} \hat{n}(y_{n+1}^-, y_{n+1}^+) - v_n \hat{n}(y_n^-, y_n^+) \right),$$

$$= \sum_{n=0}^{M-1} \left(- \int_{x_n}^{x_{n+1}} \frac{1}{2} y_h^2 v'(x) dx + \hat{n}^s(y_{n+1}^-, y_{n+1}^+) \llbracket v \rrbracket \right), \quad (3.4b)$$

$$a_h^s(y, v) = a_h(y, v) + n_h^s(y, v), \quad (3.4c)$$

$$(f, v) = \sum_{n=0}^{M-1} \int_{x_n}^{x_{n+1}} f(x, t) v(x) dx, \quad (3.4d)$$

where the term $\hat{n}(y_n^-, y_n^+)$ is the flux associated to the n th step. The penalty parameter $\sigma > 0$ should be sufficiently large to ensure the stability of the dG discretization [112, Sec. 2.7.1] with a lower bound depending only on the polynomial degree.

The semi-discrete optimality system associated to the OCP (3.3) is written as follows:

$$(\partial_t y_h, v_h) + a_h^s(y_h, v_h) = (f_h + u_h, v_h) \quad \forall v_h \in V_{h,p}, \quad t \in (0, T], \quad (3.5a)$$

$$-(\partial_t \lambda_h, q_h) + a_h^a(\lambda_h, q_h) = -(y_h - y_h^d, q_h) \quad \forall q_h \in V_{h,p}, \quad t \in (0, T], \quad (3.5b)$$

$$\alpha u_h = \lambda_h, \quad t \in (0, T], \quad (3.5c)$$

where

$$a_h^a(\lambda_\delta, \phi_\delta)$$

$$= \sum_{n=0}^{M-1} \int_{x_n}^{x_{n+1}} \epsilon y'(x, t) v'(x) dx$$

$$+ \sum_{n=0}^{M-1} \left(- \int_{x_n}^{x_{n+1}} y(x, t) \lambda'(x, t) v(x) dx + \underbrace{v_{n+1} \hat{n}^a(\lambda_{n+1}^-, \lambda_{n+1}^+) - v_n \hat{n}^a(\lambda_n^-, \lambda_n^+)}_{n_h^\lambda(\lambda_{n+1}, y_{n+1}, v_{n+1})} \right)$$

$$= \sum_{n=0}^{M-1} \left(\int_{x_n}^{x_{n+1}} \epsilon y'(x, t) v'(x) dx - \int_{x_n}^{x_{n+1}} y(x, t) \lambda'(x, t) v(x) dx + n_y^s(y_{n+1}^-, y_{n+1}^+) \llbracket \lambda \rrbracket v(x) \right).$$

The choice of the parameter κ leads to three different DG schemes:

- For $\kappa = -1$, the method is called symmetric interior penalty Galerkin (SIPG). With a large penalty term σ , the convergence is achieved.
- For $\kappa = 1$, the method is called nonsymmetric interior penalty Galerkin (NIPG). With a nonnegative penalty term σ , the convergence is achieved.
- For $\kappa = 0$, the method is called incomplete interior penalty Galerkin (IIPG). With a large penalty term σ , the convergence is achieved as in SIPG method.
- An extra stabilization term $\check{J}_\sigma(y, v)$ can be added to the bilinear form $a_h^s(y, v)$ so that the jump of the derivative is penalized.

$$\check{J}_\sigma(y, v) = \sum_{n=0}^M \frac{\check{\sigma}}{|h|^{\beta_1}} \llbracket \epsilon y'(x_n, t) \rrbracket \llbracket \epsilon v'(x_n) \rrbracket.$$

In this study, $\check{\sigma}$ has been taken zero, for simplicity.

Finite Element Discretization

To facilitate the construction of finite element matrices, all computations are carried on the reference interval \hat{K} . Then, we map the physical elements K with the endpoints x_n, x_{n+1} to the reference element \hat{K} whose endpoints are -1 and 1 .

The invertible affine map F_E can be defined from the reference element to the physical one and it is written as

$$F_E \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad x = \sum_{i=1}^2 x_i \hat{\phi}_i(\hat{x}, \hat{y}), \quad y = \sum_{i=1}^2 y_i \hat{\phi}_i(\hat{x}, \hat{y}),$$

where

$$\hat{\phi}_1(\hat{x}, \hat{y}) = \frac{1 - \hat{x}}{2}, \quad \hat{\phi}_2(\hat{x}, \hat{y}) = \frac{1 + \hat{x}}{2}.$$

For fast computation of the finite element matrices, calculations are carried on the reference interval $\hat{K} = [-1, 1]$. Therefore, we need to represent each coordinate x in the physical element K in terms of the local coordinate \hat{x} as follows:

$$x = F_K(\hat{x}) = B_K(\hat{x}) + b_K = \frac{h_n}{2} \hat{x} + \left(x_n + \frac{h_n}{2}\right).$$

We have mentioned that the test functions are discontinuous at the endpoints of each interval. Then, the support of the basis functions of $V_{h,p}$ is contained in one element. We write the DG subspace $V_{h,p}$ using basis ϕ_i^K associated to an interval K :

$$V_{h,p} = \text{span} \{ \phi_i^K : 1 \leq i \leq N_{loc}, K \in \mathcal{T}_h \}$$

with

$$\phi_i^K(x) = \begin{cases} \hat{\phi}_i \circ F_K(x), & x \in K \\ 0, & x \notin K. \end{cases} \quad (3.6)$$

Here, $\hat{\phi}_i$ represents the local basis functions and the local dimension is computed as $N_{loc} = p + 1$ where p is the degree of the finite element approximation.

The integrals obtained by the (bi)linear forms are computed on the reference element. The numerical quadrature rule can be used to approximate the integral in DG formulation

$$\int_{K_n} v \, dx = \frac{h_n}{2} \int_{\hat{K}_n} \hat{v} \, d\hat{x} \approx \frac{h_n}{2} \sum_{j=1}^{Q_D} \omega_j \hat{v}(s_j).$$

where ω_j are the weights and $s_j \in \hat{K}_n$ are the nodes.

3.1.1 Semi-discrete Optimal Control Problem

The semi-discrete solution is written using the global basis functions (3.6), in particular for the state solution, as follows:

$$y_h(x, t) = \sum_{K \in \mathcal{T}_h} \sum_{i=1}^{N_{loc}} y_{h,i}^K(t) v_i^K(x), \quad \forall x \in \Omega, \quad \forall t \in (0, T). \quad (3.7)$$

The coefficients $y_{h,i}^K$ are functions of time and the number of elements in the domain Ω are denoted by N_{el} . The basis functions and the coefficients can be written as

$$\{v_i^K(t), \quad 1 \leq i \leq N_{loc}, \quad K \in \mathcal{T}_h\}, \quad \{y_{h,i}^K(t), \quad 1 \leq i \leq N_{loc}, \quad K \in \mathcal{T}_h\}.$$

In order to write the semi-discrete OCP in matrix form, we substitute the semi-discrete solution into (3.5a) with

$$\begin{aligned} \mathbf{y}_h(t) &= (y_{h,1}^1(t), y_{h,2}^1(t), \dots, y_{h,N_{loc}}^1(t), \dots, y_{h,1}^{N_{el}}(t), y_{h,2}^{N_{el}}(t), \dots, y_{h,N_{loc}}^{N_{el}}(t)), \\ \boldsymbol{\lambda}_h(t) &= (\lambda_{h,1}^1(t), \lambda_{h,2}^1(t), \dots, \lambda_{h,N_{loc}}^1(t), \dots, \lambda_{h,1}^{N_{el}}(t), \lambda_{h,2}^{N_{el}}(t), \dots, \lambda_{h,N_{loc}}^{N_{el}}(t)), \\ \mathbf{u}(t) &= (u_{h,1}^1(t), u_{h,2}^1(t), \dots, u_{h,N_{loc}}^1(t), \dots, u_{h,1}^{N_{el}}(t), u_{h,2}^{N_{el}}(t), \dots, u_{h,N_{loc}}^{N_{el}}(t)), \end{aligned}$$

as follows:

$$\underset{u_h \in V_{h,p}}{\text{minimize}} \quad \frac{1}{2} \int_0^T ((\mathbf{y}_h(t) - \mathbf{y}_h^d(t))^T \mathbf{M} (\mathbf{y}_h(t) - \mathbf{y}_h^d(t)) + \alpha (\mathbf{u}_h(t))^T \mathbf{M} \mathbf{u}_h(t)) \, dt, \quad (3.8a)$$

$$\text{subject to} \quad \mathbf{M} \frac{d\mathbf{y}_h(t)}{dt} + \mathbf{A} \mathbf{y}_h + \mathbf{N}(\mathbf{y}_h(t)) = \mathbf{F}_h(t) + \mathbf{M} \mathbf{u}_h(t), \quad \forall v_h \in V_{h,p} \quad (3.8b)$$

$$\mathbf{M} \mathbf{y}_h(0) = \mathbf{y}_{h,0}, \quad (3.8c)$$

where

$$\begin{aligned}
\mathbf{M}_{ij} &= (v_j, v_i)_\Omega, & \mathbf{A}_{ij} &= a_h(v_j, v_i)_\Omega, & \mathbf{F}_{h,i} &= (f_h(x, t), v_i), \\
\mathbf{N}_i(\mathbf{y}(t)) &= n_h^s(y_h(x, t), v_i), & \mathbf{N}_i^\lambda(\boldsymbol{\lambda}_h(t), \mathbf{y}_h(t)) &= n_h^\lambda(\lambda_h(x, t), y_h(x, t), v_i), \\
(\mathbf{M}\mathbf{y}_h^d(t))_i &= (y_h^d(x, t), v_i), & (\mathbf{y}_{h,0})_i &= (y_{h,0}(x, t), v_i).
\end{aligned} \tag{3.9}$$

Then, the semi-discrete optimality system in matrix-vector form is read as:

$$\mathbf{M} \frac{d\mathbf{y}_h(t)}{dt} + \mathbf{A}\mathbf{y}_h + \mathbf{N}(\mathbf{y}_h(t)) = \mathbf{F}_h(t) + \mathbf{M}\mathbf{u}_h(t), \tag{3.10a}$$

$$\mathbf{M}\mathbf{y}_h(0) = \mathbf{y}_{h,0}, \tag{3.10b}$$

$$-\mathbf{M} \frac{d\boldsymbol{\lambda}_h(t)}{dt} + \mathbf{A}^T \boldsymbol{\lambda}_h - \mathbf{N}^\lambda(\boldsymbol{\lambda}_h(t), \mathbf{y}_h(t)) = -(\mathbf{M}\mathbf{y}_h(t) - \mathbf{Y}_h^d(t)), \tag{3.10c}$$

$$\mathbf{M}\boldsymbol{\lambda}_h(T) = 0, \tag{3.10d}$$

$$\alpha \mathbf{M}\mathbf{u}_h(t) = \mathbf{M}\boldsymbol{\lambda}_h(t). \tag{3.10e}$$

3.2 Discontinuous Galerkin Methods in Two-Dimension

In this section, we briefly describe the symmetric interior penalty Galerkin semi-discretization in space for time-dependent diffusion-convection-reaction equation in two-dimension:

$$\partial_t y - \epsilon \Delta y + \boldsymbol{\beta} \cdot \nabla y + r y = f, \quad (x, t) \in \Omega \times (0, T], \tag{3.11a}$$

$$y(x, t) = 0, \quad (x, t) \in \partial\Omega \times [0, T], \tag{3.11b}$$

$$y(x, 0) = y_0(x), \quad x \in \Omega. \tag{3.11c}$$

The source function is denoted by $f \in L^2(0, T; L^2(\Omega))$. The initial condition is also defined as $y_0(x) \in H_0^1(\Omega)$. The diffusion and reaction coefficients are $\epsilon > 0$ and $r \in L^\infty(\Omega)$, respectively. The velocity field $\boldsymbol{\beta} \in (W^{1,\infty}(\Omega))^2$ satisfies the incompressibility condition, i.e. $\nabla \cdot \boldsymbol{\beta} = 0$. Furthermore, we assume the existence of the constant C_0 such that $r \geq C_0$ a.e. in Ω so that the well-posedness of the problem (3.11) is guaranteed.

Let $\{\mathcal{T}_h\}_h$ be a family of shape regular meshes such that $\bar{\Omega} = \cup_{K \in \mathcal{T}_h} \bar{K}$, $K_i \cap K_j = \emptyset$ for $K_i, K_j \in \mathcal{T}_h$, $i \neq j$. The diameters of elements K are denoted by h_K . The maximum diameter is $h = \max_{K \in \mathcal{T}_h} h_K$. In addition, the length of an edge E is denoted by h_E .

We split the set of all edges \mathcal{E}_h into the set \mathcal{E}_h^0 of interior edges and the set \mathcal{E}_h^∂ of boundary edges so that $\mathcal{E}_h = \mathcal{E}_h^\partial \cup \mathcal{E}_h^0$. Let \mathbf{n} denote the unit outward normal to $\partial\Omega$. We define the inflow boundary

$$\Gamma^- = \{x \in \partial\Omega : \boldsymbol{\beta} \cdot \mathbf{n}(x) < 0\}$$

and the outflow boundary $\Gamma^+ = \partial\Omega \setminus \Gamma^-$. The boundary edges are decomposed into edges $\mathcal{E}_h^- = \{E \in \mathcal{E}_h^\partial : E \subset \Gamma^-\}$ that correspond to the inflow boundary and edges $\mathcal{E}_h^+ = \mathcal{E}_h^\partial \setminus \mathcal{E}_h^-$ that correspond to the outflow boundary. The inflow and outflow boundaries of an element $K \in \mathcal{T}_h$ are defined by

$$\partial K^- = \{x \in \partial K : \boldsymbol{\beta} \cdot \mathbf{n}_K(x) < 0\}, \quad \partial K^+ = \partial K \setminus \partial K^-,$$

where \mathbf{n}_K is the unit normal vector on the boundary ∂K of an element K .

Let the edge E be a common edge for two elements K and K^e . For a piecewise continuous scalar function y , there are two traces of y along E , denoted by $y|_E$ from interior of K and $y^e|_E$ from interior of K^e . Then, the jump and average of y across the edge E are defined by:

$$[[y]] = y|_E \mathbf{n}_K + y^e|_E \mathbf{n}_{K^e}, \quad \{\{y\}\} = \frac{1}{2}(y|_E + y^e|_E).$$

Similarly, for a piecewise continuous vector field ∇y , the jump and average across an edge E are given by

$$[[\nabla y]] = \nabla y|_E \cdot \mathbf{n}_K + \nabla y^e|_E \cdot \mathbf{n}_{K^e}, \quad \{\{\nabla y\}\} = \frac{1}{2}(\nabla y|_E + \nabla y^e|_E).$$

For a boundary edge $E \in K \cap \Gamma$, we set $\{\{\nabla y\}\} = \nabla y$ and $[[y]] = y\mathbf{n}$ where \mathbf{n} is the outward normal unit vector on Γ .

We use discontinuous piecewise finite element spaces to define the discrete test, state and control spaces

$$V_{h,p} = Y_{h,p} = U_{h,p} = \{y \in L^2(\Omega) : y|_K \in \mathbb{P}^p(K) \quad \forall K \in \mathcal{T}_h\}. \quad (3.12)$$

Here, $\mathbb{P}^p(K)$ denotes the set of all polynomials on $K \in \mathcal{T}_h$ of degree p .

The DG formulation of the state equation (3.11) upwinding the convection term for fixed control u is written as follows [13, 14]:

$$(\partial_t y_h, v_h) + a_h^s(y_h, v_h) = (f_h + u_h, v_h) \quad \forall v_h \in V_{h,p}, \quad t \in (0, T], \quad (3.13)$$

with the (bi-)linear forms

$$\begin{aligned} a^d(y, v) &= \sum_{K \in \mathcal{T}_h} \int_K \epsilon \nabla y \cdot \nabla v \, dx \\ &- \sum_{E \in \mathcal{E}_h} \int_E (\{\{\epsilon \nabla y\}\} \cdot [v] - \kappa \{\{\epsilon \nabla v\}\} \cdot [y] - \epsilon \overbrace{\frac{\sigma}{h_E} [y] \cdot [v]}^{J_\sigma(y,v)}) \, ds \end{aligned} \quad (3.14)$$

and

$$\begin{aligned} a_h^s(y, v) &= a^d(y, v) + \sum_{K \in \mathcal{T}_h} \int_K (\boldsymbol{\beta} \cdot \nabla y v + r y v) \, dx \\ &+ \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \setminus \Gamma^-} \boldsymbol{\beta} \cdot \mathbf{n} (y^e - y) v \, ds - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \boldsymbol{\beta} \cdot \mathbf{n} y v \, ds. \end{aligned} \quad (3.15)$$

The penalty parameter $\sigma > 0$ should be sufficiently large to ensure the stability of the dG discretization [112, Sec. 2.7.1] with a lower bound depending only on the polynomial degree.

The semi-discrete optimality system associated to the OCP (1.1) is written as follows:

$$(\partial_t y_h, v_h) + a_h^s(y_h, v_h) = (f_h + u_h, v_h) \quad \forall v_h \in V_{h,p}, \quad t \in (0, T], \quad (3.16a)$$

$$- (\partial_t \lambda_h, q_h) + a_h^a(\lambda_h, q_h) = -(y_h - y_h^d, q_h) \quad \forall q_h \in V_{h,p}, \quad t \in (0, T], \quad (3.16b)$$

$$\alpha u_h = \lambda_h, \quad t \in (0, T]. \quad (3.16c)$$

Here, since the problem is linear; $n_h^s(y_h, v_h) = 0$. Moreover, the bilinear form in the adjoint equation is read as

$$\begin{aligned} a_h^a(p, \psi) &= \sum_{K \in \mathcal{T}_h} \int_K \epsilon \nabla p \cdot \nabla \psi \, dx \\ &\quad - \sum_{E \in \mathcal{E}_h} \int_E (\{\{\epsilon \nabla p\}\} \cdot \llbracket \psi \rrbracket + \{\{\epsilon \nabla \psi\}\} \cdot \llbracket p \rrbracket - \frac{\sigma \epsilon}{h_E} \llbracket p \rrbracket \cdot \llbracket \psi \rrbracket) \, ds \\ &\quad + \sum_{K \in \mathcal{T}_h} \int_K (-\beta \cdot \nabla p \psi + r p \psi) \, dx \\ &\quad - \sum_{K \in \mathcal{T}_h} \int_{\partial K^+ \setminus \Gamma^+} \beta \cdot \mathbf{n} (p^e - p) \psi \, ds + \sum_{K \in \mathcal{T}_h} \int_{\partial K^+ \cap \Gamma^+} \beta \cdot \mathbf{n} p \psi \, ds. \end{aligned}$$

Finite Element Discretization

We mention the discontinuous piecewise finite element spaces

$$V_{h,p} = \{y \in L^2(\Omega) : y|_K \in \mathbb{P}^p(K) \quad \forall K \in \mathcal{T}_h\}.$$

which is a subspace of $H^s(\xi_h)$ for $s > 3/2$. The test functions are discontinuous along the edges of each finite element.

To facilitate the construction of finite element matrices, all computations are carried on the reference element \hat{E} . Then, we map the physical elements E with vertices $V_i(x_i, y_i)$ for $i = 1, 2, 3$ to the reference element \hat{E} whose vertices are $\hat{V}_1(0, 0)$, $\hat{V}_2(1, 0)$ and $\hat{V}_3(0, 1)$.

The invertible affine map F_E can be defined from the reference element to the physical one and it is written as

$$F_E \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad x = \sum_{i=1}^3 x_i \hat{\phi}_i(\hat{x}, \hat{y}), \quad y = \sum_{i=1}^3 y_i \hat{\phi}_i(\hat{x}, \hat{y}),$$

where

$$\hat{\phi}_1(\hat{x}, \hat{y}) = 1 - \hat{x} - \hat{y}, \quad \hat{\phi}_2(\hat{x}, \hat{y}) = \hat{x}, \quad \hat{\phi}_3(\hat{x}, \hat{y}) = \hat{y}.$$

This map can be written using an invertible matrix B_E and a translation vector b_E as follows

$$\begin{pmatrix} x \\ y \end{pmatrix} = F_E \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = B_E \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} + b_E$$

where

$$B_E = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}, \quad b_E = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}.$$

Passing to the reference element from the physical elements via the mapping F_E can be seen as a change of variables. By $\hat{v} = v \circ F_E$, we have $\hat{v}(\hat{x}, \hat{y}) = v(x, y)$. In addition, the gradient is mapped as $\hat{\nabla} \hat{v} = B_E^T \nabla v \circ F_E$.

We have mentioned that the test functions are discontinuous along the edges. Then, the support of the basis functions of $V_{h,p}$ is contained in one element. We write the DG subspace $V_{h,p}$ using basis ϕ_i^E associated to an element E :

$$V_{h,p} = \text{span} \{ \phi_i^E : 1 \leq i \leq N_{loc}, E \in \xi_h \}$$

with

$$\phi_i^E(x) = \begin{cases} \hat{\phi}_i \circ F_E(x), & x \in E, \\ 0, & x \notin E. \end{cases} \quad (3.17)$$

Here, $\hat{\phi}_i$ represents the local basis functions and the local dimension is computed as $N_{loc} = \frac{(k+1)(k+2)}{2}$ where k is the degree of the finite element approximation.

As we have mentioned before, the integrals obtained by the (bi)linear forms are computed on the reference element. The numerical quadrature rule can be used to approximate the integral in DG formulation

$$\int_{\hat{E}} \hat{v} dx \approx \sum_{j=1}^{Q_D} \omega_j \hat{v}(s_{x,j}, s_{y,j}),$$

where ω_j are the weights and $(s_{x,j}, s_{y,j}) \in \hat{E}$ are the nodes.

By the following equality, we can observe how the map F_E is used to transfer the physical element to the reference element:

$$\int_E v dx = \int_{\hat{E}} v \circ F_E \det(B_E) dx = 2|E| \int_{\hat{E}} \hat{v} dx \approx 2|E| \sum_{j=1}^{Q_D} \omega_j \hat{v}(s_{x,j}, s_{y,j}).$$

3.2.1 Semi-discrete Optimal Control Problem

The semi-discrete solution is written using the global basis functions (3.17), in particular for the state solution, as follows:

$$y_h(x, t) = \sum_{K \in \mathcal{T}_h} \sum_{i=1}^{N_{loc}} y_{h,i}^K(t) v_i^K(x), \quad \forall x \in \Omega, \quad \forall t \in (0, T). \quad (3.18)$$

In order to write the semi-discrete OCP in matrix form, we substitute the semi-discrete solution into (3.5a) with

$$\begin{aligned} \mathbf{y}_h(t) &= (y_{h,1}^1(t), y_{h,2}^1(t), \dots, y_{h,N_{loc}}^1(t), \dots, y_{h,1}^{N_{el}}(t), y_{h,2}^{N_{el}}(t), \dots, y_{h,N_{loc}}^{N_{el}}(t)), \\ \boldsymbol{\lambda}_h(t) &= (\lambda_{h,1}^1(t), \lambda_{h,2}^1(t), \dots, \lambda_{h,N_{loc}}^1(t), \dots, \lambda_{h,1}^{N_{el}}(t), \lambda_{h,2}^{N_{el}}(t), \dots, \lambda_{h,N_{loc}}^{N_{el}}(t)), \\ \mathbf{u}_h(t) &= (u_{h,1}^1(t), u_{h,2}^1(t), \dots, u_{h,N_{loc}}^1(t), \dots, u_{h,1}^{N_{el}}(t), u_{h,2}^{N_{el}}(t), \dots, u_{h,N_{loc}}^{N_{el}}(t)), \end{aligned}$$

as follows:

$$\underset{u_h \in V_{h,p}}{\text{minimize}} \quad \frac{1}{2} \int_0^T ((\mathbf{y}_h(t) - \mathbf{y}_h^d(t))^T \mathbf{M} (\mathbf{y}_h(t) - \mathbf{y}_h^d(t)) + \alpha (\mathbf{u}_h(t))^T \mathbf{M} \mathbf{u}_h(t)) dt, \quad (3.19a)$$

$$\text{subject to} \quad \mathbf{M} \frac{d\mathbf{y}_h(t)}{dt} + \mathbf{A} \mathbf{y}_h = \mathbf{F}_h(t) + \mathbf{M} \mathbf{u}_h(t), \quad \forall v_h \in V_{h,p}, \quad (3.19b)$$

$$\mathbf{M} \mathbf{y}_h(0) = \mathbf{y}_{h,0}. \quad (3.19c)$$

Then, the semi-discrete optimality system in weak form is read as in (3.10) with $\mathbf{N}^s(\mathbf{y}_h(t)) = 0$ and $\mathbf{N}^\lambda(\boldsymbol{\lambda}_h(t), \mathbf{y}_h(t)) = 0$.

CHAPTER 4

VARIATIONAL TIME DISCRETIZATION METHODS

Variational time discretization method appeared first in [77] and developed by various papers. For continuous Galerkin-Petrov (cGP($q+1$)) method, solution space consists of continuous piecewise polynomials of degree $q + 1$ and test functions are piecewise discontinuous polynomials of degree q . For discontinuous Galerkin (dG(q)) methods, both of the test and the trial spaces are piecewise discontinuous polynomials of degree q [120, Chap. 12]. Advantages of variational time discretization are stability, nodal superconvergence, and applicability of space-time adaptivity. Both continuous and discontinuous Galerkin methods are A -stable; discontinuous Galerkin methods are even L -stable (strongly stable). Convergence order of cGP($q+1$) methods are of one order higher than the dG(q) methods. A priori error estimates of optimal order can be obtained with respect to the size of time steps and the regularity requirements of the solutions [49], whereas dG methods require less regular solutions than the cG methods. The dG(q) methods are superconvergent at the nodal points of order $2q + 1$ when the order of the method is q and the solution of the problem is sufficiently regular. For dG method, long-time integration is achievable without accumulating the error on the current time interval [49]. Time-space adaptivity can be easily implemented because time and space discretizations are treated similarly [25]. Using a posteriori error estimates, adaptive hp time stepping and dynamic meshes (the use of different spatial discretization for each time step) can be directly incorporated in the discrete formulation [111, 114].

DG methods have one more advantage in terms of discretization of the Lagrangian of the optimization problem. As we mentioned before, there exist two different approaches for solving OCPs: *optimize-then-discretize* (OD) and *discretize-then-optimize* (DO). In the OD approach, first infinite dimensional optimality system is derived containing the state and the adjoint equations and the optimality condition. Then, optimality system is discretized by using a suitable discretization method in space and time. In DO approach, infinite dimensional OCP is discretized and then finite-dimensional optimality system is derived. DO and OD approaches do not commute in general for OCPs governed by diffusion-convection-reaction equation [34]. However, commutativity is achieved in the case of SIPG discretization for steady state problems [91]. For discontinuous Galerkin time discretization, where both trial and test spaces are discontinuous, we show that OD and DO approaches commute [3], i.e. the adjoint state is discretized as we do for the state variable. For continuous Galerkin time discretization, where the trial spaces are continuous and the test spaces are discontinuous, OD

and DO approaches do not commute.

In the literature, there are several studies concerning this method for a single PDE. For example, space-time dG method is applied to time-dependent domains in [127] and stability and error estimates are proven. Compressible Navier-Stokes equation is solved numerically and local grid adaptation is utilised in [81]. For evolutionary DCR equation, this method is employed separately in space and time using nonconforming meshes for each time step [52]. A quasi-linear boundary value problem is solved in [26]. In the studies [26, 42, 44, 52], problems with nonlinear convective term are investigated numerically and a priori error estimates are derived. On the other hand, adaptive time refinement strategies are developed for linear problems [48, 49] and for nonlinear problems [47]. Continuous and discontinuous Galerkin methods are compared and dynamic mesh refinement techniques is employed [114]. In [136], hp refinement both in spatial and temporal domain is presented.

Although the number of studies concerning a single PDE is numerous, variational time discretization methods have not been discussed much for optimal control problems when compared to finite difference approximation in time. In [16], continuous and discontinuous Galerkin method is applied to continuous OCPs and optimization techniques for an efficient solution of parabolic OCPs are compared. A posteriori error estimates are given in [100] for parabolic optimization problems using space-time FEM. In [101, 102, 117], a priori error estimates are derived for this problem for unconstrained and control-constrained problems. Discontinuous in time and conforming elements in space are used for discretization of the parabolic OCPs in [29, 30] and theoretical estimates are proven. A velocity tracking problem for Navier-Stokes equation is discussed in [45] using space-time FEM. In the studies above, spatial discretization is performed using continuous finite element method. In our previous study [3], variational time discretization method combined with discontinuous Galerkin spatial discretization is discussed for optimal control of DCR equation, and error estimates and numerical findings are presented.

In this chapter, we discuss continuous Galerkin-Petrov and discontinuous Galerkin methods in Sec. 4.1 and Sec. 4.2, respectively, by discussing the commutativity approaches of these methods. In addition, we derive the associated weak forms, gradient and Hessian-times-vector computations. Then, we mention some auxiliary results and derive a priori error estimates for OCP of DCR equation in Sec. 4.3 and explain the derivation of the associated linear system of equations in Sec. 4.4. Then, numerical results follow in Sec. 4.4.1. In Sec. 4.5, a priori error estimates for OCP of Burgers equation are proven and associated weak forms in matrix-vector form for dG time discretization is given in Sec. 4.6. Then, numerical results are presented in Sec. 4.6.1.

Let $0 = t_0 < t_1 < \dots < t_{N_T} = T$ be a subdivision of $I = (0, T]$ with time intervals $I_m = (t_{m-1}, t_m]$ and time steps $k_m = t_m - t_{m-1}$ for $m = 1, \dots, N_T$ and $k = \max_{1 \leq m \leq N_T} k_m$. We note that the same mesh is used at each time level t_m for $m = 0, \dots, N_T$. Let f_δ and y_δ^d be approximations of the source function f and the desired state function y^d on each interval I_m .

The fully-discrete solution is written, in particular for the state solution, as follows:

$$y_\delta(x, t) = \sum_{i=1}^{N_{loc}N_{el}} y_{\delta,i}(t)v_i(x), \quad \forall x \in \Omega, \quad \forall K \in \mathcal{T}, \quad \forall t \in (0, T).$$

We proceed with the fully-discrete OCP and associated weak forms.

4.1 Continuous Galerkin-Petrov (cGP(q+1)) Method

We define the discontinuous test space

$$V_{h,p}^{k,q} = \{v \in L^2(I; V_{h,p}) : v|_{I_m} \in P_q(I_m, V_{h,p}), m = 1, \dots, N_T, v_m(0) \in L^2(\Omega)\}, \quad (4.1)$$

and the continuous trial space as follows

$$\tilde{V}_{h,p}^{k,q+1} = \{v \in C(\bar{I}; V_{h,p}) : v|_{I_m} \in P_{q+1}(I_m, V_{h,p}), m = 1, \dots, N_T\},$$

where $P_q(I_m, V_{h,p})$ denotes the space of polynomials of degree q defined on I_m with values on $V_{h,p}$. Then, the fully-discrete optimal control problem is written as

$$\underset{u_\delta \in \tilde{V}_{h,p}^{k,q+1}}{\text{minimize}} \quad \frac{1}{2} \int_0^T \left(\sum_{K \in \mathcal{T}_h} (\|y_\delta - y_\delta^d\|_{L^2(K)}^2 + \alpha \|u_\delta\|_{L^2(K)}^2) \right) dt, \quad (4.2a)$$

$$\begin{aligned} \text{subject to} \quad & \int_0^T ((\partial_t y_\delta, v_\delta) + a_h^s(y_\delta, v_\delta)) dt = \int_0^T (f_\delta + u_\delta, v_\delta) dt, \quad \forall v_\delta \in V_{h,p}^{k,q}, \\ & y_{\delta,0} = (y_0)_\delta. \end{aligned} \quad (4.2b)$$

The OCP (4.2) has a unique solution (y_δ, u_δ) and that pair $(y_\delta, u_\delta) \in \tilde{V}_{h,p}^{k,q+1} \times \tilde{V}_{h,p}^{k,q+1}$ is the solution of (4.2) if and only if there is an adjoint $\lambda_\delta \in \tilde{V}_{h,p}^{k,q+1}$ such that $(y_\delta, u_\delta, \lambda_\delta) \in \tilde{V}_{h,p}^{k,q+1} \times \tilde{V}_{h,p}^{k,q+1} \times \tilde{V}_{h,p}^{k,q+1}$ is the unique solution of the fully-discrete optimality system [122]

$$\begin{aligned} \int_0^T ((\partial_t y_\delta, v_\delta) + a_h^s(y_\delta, v_\delta)) dt &= \int_0^T (f_\delta + u_\delta, v_\delta) dt, \quad \forall v_\delta \in V_{h,p}^{k,q}, \\ y_{\delta,0} &= (y_0)_\delta, \end{aligned} \quad (4.3a)$$

$$\begin{aligned} \int_0^T (-\partial_t \lambda_\delta, \phi_\delta) + a_h^a(\lambda_\delta, \phi_\delta) dt &= - \int_0^T (y_\delta - y_\delta^d, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}, \\ \lambda_{\delta,N} &= 0, \end{aligned} \quad (4.3b)$$

$$\alpha u_\delta = \lambda_\delta. \quad (4.3c)$$

We note that the term $a_h^a(\lambda_\delta, \phi_\delta)$ in the variational formulation (4.3b) corresponds to $a_h^s(y_\delta, \lambda_\delta)$ if the PDE constraint $c(y, u) = 0$ models the diffusion-convection-reaction equation.

We write Algorithm 2.2 in terms of the variational forms arising due to (cGP(q+1)) method.

Algorithm 4.1 Gradient Computation in Algorithm 2.2 Using (cGP(q+1)) Method

1: Given u_δ , solve the state equation

$$\int_0^T ((\partial_t y_\delta, v_\delta) + a_h^s(y_\delta, v_\delta)) dt = \int_0^T (f_\delta + u_\delta, v_\delta) dt, \quad \forall v_\delta \in V_{h,p}^{k,q},$$

$$y_{\delta,0} = (y_0)_\delta, \quad (4.4)$$

for $y_\delta := y_\delta(u_\delta)$.

2: Solve the adjoint equation

$$\int_0^T (-(\partial_t \lambda_\delta, \phi_\delta) + a_h^a(\lambda_\delta, \phi_\delta)) dt = - \int_0^T (y_\delta - y_\delta^d, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q},$$

$$\lambda_{\delta,N} = 0, \quad (4.5)$$

for $\lambda_\delta := \lambda_\delta(u_\delta)$.

3: Compute

$$\nabla \widehat{J}(u) = \int_0^T (\alpha u_\delta - \lambda_\delta, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}.$$

Commutativity Properties

We derive the optimality system arising from DO approach, that also appears in [16, Sec.3], and compare it with the optimality system (4.3). For simplicity, we take $c(y, u) = 0$ as the diffusion-convection-reaction equation. We consider the discrete Lagrangian defined on $\tilde{V}_{h,p}^{k,q+1} \times \tilde{V}_{h,p}^{k,q+1} \times V_{h,p}^{k,q}$ as follows

$$\begin{aligned} \mathcal{L}(y_\delta, u_\delta, \lambda_\delta) &= \frac{1}{2} \int_0^T \left(\sum_{K \in \mathcal{T}_h} (\|y_\delta - y_\delta^d\|_{L^2(K)}^2 + \alpha \|u_\delta\|_{L^2(K)}^2) \right) dt \\ &+ \sum_{m=1}^{N_T} \left\{ \int_{I_m} ((\partial_t y_\delta, \lambda_\delta) + a_h^s(y_\delta, \lambda_\delta)) dt - \int_{I_m} (f_\delta + u_\delta, \lambda_\delta) dt \right\} \\ &+ ((y_0)_\delta - y_{\delta,0}, \lambda_{\delta,0}). \end{aligned}$$

We differentiate \mathcal{L} with respect to y_δ , apply integration by parts. We add and subtract $(\phi_{\delta, N_T}, \lambda_{\delta, N_T}^+)$. Then, on each subinterval I_m , the adjoint equation reads as

$$\begin{aligned} &\int_{I_m} (-(\partial_t \lambda_\delta, \phi_\delta) + a_h^s(\phi_\delta, \lambda_\delta)) dt - ([\lambda_\delta]_m, \phi_\delta^m) \\ &= - \int_{I_m} (y_\delta - y_\delta^d, \phi_\delta) dt, \quad \forall \lambda_\delta \in V_{h,p}^{k,q}, \phi_\delta \in \tilde{V}_{h,p}^{k,q+1}, \end{aligned} \quad (4.10)$$

Algorithm 4.2 Hessian-Times-Vector Computation in Algorithm 2.3 Using (cGP(q+1)) Method

1: Given u_δ , solve the state equation

$$\int_0^T ((\partial_t y_\delta, v_\delta) + a_h^s(y_\delta, v_\delta)) dt = \int_0^T (f_\delta + u_\delta, v_\delta) dt, \quad \forall v_\delta \in V_{h,p}^{k,q},$$

$$y_{\delta,0} = (y_0)_\delta, \quad (4.6)$$

for $y_\delta := y_\delta(u_\delta)$.

2: Solve the adjoint equation

$$\int_0^T (-(\partial_t \lambda_\delta, \phi_\delta) + a_h^a(\lambda_\delta, \phi_\delta)) dt = - \int_0^T (y_\delta - y_\delta^d, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q},$$

$$\lambda_{\delta,N} = 0, \quad (4.7)$$

for $\lambda_\delta := \lambda_\delta(u_\delta)$.

3: Solve the equation

$$\int_0^T ((\partial_t w_\delta, \phi_\delta) + a_h^s(w_\delta, \phi_\delta)) dt = \int_0^T (u, \nu_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q},$$

$$w_{\delta,0} = 0. \quad (4.8)$$

4: Solve the equation

$$\int_0^T (-(\partial_t p_\delta, \phi_\delta) + a_h^a(p_\delta, \phi_\delta)) dt = - \int_0^T (w, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q},$$

$$p_{\delta,N} = 0. \quad (4.9)$$

5: Compute

$$\nabla^2 \widehat{J}(u) \nu_\delta = \int_0^T (\alpha \nu_\delta - p, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}.$$

with $\lambda_{\delta,N}^+ = 0$. We observe that the temporal jump terms appear in (4.10), while it is not the case for the adjoint equation in (4.3b). In addition, in OD approach $\lambda_\delta \in \tilde{V}_{h,p}^{k,q+1}$, while in DO approach $\lambda_\delta \in V_{h,p}^{k,q}$. Therefore, OD and DO approaches do not commute for cGP-method.

4.2 Discontinuous Galerkin (dG(q)) Methods

We use (4.1) as discontinuous test and trial space. We define the temporal jump of $v \in V_{h,p}^{k,q}$ as $[v]_m = v_+^m - v_-^m$, where $w_\pm^m = \lim_{\varepsilon \rightarrow 0^\pm} v(t_m + \varepsilon)$. Then, the fully-discrete optimal control problem is written as

$$\underset{u_\delta \in V_{h,p}^{k,q}}{\text{minimize}} \quad \frac{1}{2} \int_0^T \left(\sum_{K \in \mathcal{T}_h} (\|y_\delta - y_\delta^d\|_{L^2(K)}^2 + \alpha \|u_\delta\|_{L^2(K)}^2) \right) dt, \quad (4.11a)$$

$$\begin{aligned} \text{subject to} \quad & \sum_{m=1}^{N_T} \int_{I_m} (\partial_t y_\delta, v_\delta) dt + \int_0^T a_h^s(y_\delta, v_\delta) dt + \sum_{m=1}^{N_T} ([y_\delta]_{m-1}, v_{\delta,+}^{m-1}) \\ & = \int_0^T (f_\delta + u_\delta, v_\delta) dt, \quad \forall v_\delta \in V_{h,p}^{k,q}, \quad y_{\delta,0}^- = (y_0)_\delta. \end{aligned} \quad (4.11b)$$

The OCP (4.11) has a unique solution (y_δ, u_δ) and that pair $(y_\delta, u_\delta) \in V_{h,p}^{k,q} \times V_{h,p}^{k,q}$ is the solution of (4.11) if and only if there is an adjoint $\lambda_\delta \in V_{h,p}^{k,q}$ such that $(y_\delta, u_\delta, \lambda_\delta) \in V_{h,p}^{k,q} \times V_{h,p}^{k,q} \times V_{h,p}^{k,q}$ is the unique solution of the fully-discrete optimality system [122]

$$\begin{aligned} & \sum_{m=1}^{N_T} \int_{I_m} (\partial_t y_\delta, v_\delta) dt + \int_0^T a_h^s(y_\delta, v_\delta) dt + \sum_{m=1}^{N_T} ([y_\delta]_{m-1}, v_{\delta,+}^{m-1}) \\ & = \int_0^T (f_\delta + u_\delta, v_\delta) dt, \quad \forall v_\delta \in V_{h,p}^{k,q}, \quad y_{\delta,0}^- = (y_0)_\delta, \end{aligned} \quad (4.12a)$$

$$\begin{aligned} & \sum_{m=1}^{N_T} \int_{I_m} (-\partial_t \lambda_\delta, \phi_\delta) dt + \int_0^T a_h^a(\lambda_\delta, \phi_\delta) dt - \sum_{m=1}^{N_T} ([\lambda_\delta]_m, \phi_{\delta,-}^m) \\ & = - \int_0^T (y_\delta - y_\delta^d, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}, \quad \lambda_{\delta,N}^+ = 0, \end{aligned} \quad (4.12b)$$

$$\alpha u_\delta = \lambda_\delta. \quad (4.12c)$$

We write Algorithm 2.2 in terms of the variational forms arising due to (dG(q)) method.

Commutativity Properties

We derive the optimality system arising from DO approach and compare it with the optimality system (4.12). For simplicity, we take $c(y, u) = 0$ as the diffusion-convection-

Algorithm 4.3 Gradient Computation in Algorithm 2.2 Using (dG(q)) Method

1: Given u_δ , solve the state equation

$$\begin{aligned}
 & \sum_{m=1}^{N_T} \int_{I_m} (\partial_t y_\delta, v_\delta) dt + \int_0^T a_h^s(y_\delta, v_\delta) dt + \sum_{m=1}^{N_T} ([y_\delta]_{m-1}, v_{\delta,+}^{m-1}) \\
 & = \int_0^T (f_\delta + u_\delta, v_\delta) dt, \quad \forall v_\delta \in V_{h,p}^{k,q}, \quad y_{\delta,0}^- = (y_0)_\delta, \quad (4.13)
 \end{aligned}$$

for $y_\delta := y_\delta(u_\delta)$.

2: Solve the adjoint equation

$$\begin{aligned}
 & \sum_{m=1}^{N_T} \int_{I_m} (-\partial_t \lambda_\delta, \phi_\delta) dt + \int_0^T a_h^a(\lambda_\delta, \phi_\delta) dt - \sum_{m=1}^{N_T} ([\lambda_\delta]_m, \phi_{\delta,-}^m) \\
 & = - \int_0^T (y_\delta - y_\delta^d, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}, \quad \lambda_{\delta,N}^+ = 0, \quad (4.14)
 \end{aligned}$$

for $\lambda_\delta := \lambda_\delta(u_\delta)$.

3: Compute

$$\nabla \widehat{J}(u) = \int_0^T (\alpha u_\delta - \lambda_\delta, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}.$$

Algorithm 4.4 Hessian-Times-Vector Computation in Algorithm 2.3 Using (dG(q)) Method

1: Given u_δ , solve the state equation

$$\begin{aligned} & \sum_{m=1}^{N_T} \int_{I_m} (\partial_t y_\delta, v_\delta) dt + \int_0^T a_h^s(y_\delta, v_\delta) dt + \sum_{m=1}^{N_T} ([y_\delta]_{m-1}, v_{\delta,+}^{m-1}) \\ &= \int_0^T (f_\delta + u_\delta, v_\delta) dt, \quad \forall v_\delta \in V_{h,p}^{k,q}, \quad y_{\delta,0}^- = (y_0)_\delta, \end{aligned} \quad (4.15)$$

for $y_\delta := y_\delta(u_\delta)$.

2: Solve the adjoint equation

$$\begin{aligned} & \sum_{m=1}^{N_T} \int_{I_m} (-\partial_t \lambda_\delta, \phi_\delta) dt + \int_0^T a_h^a(\lambda_\delta, \phi_\delta) dt - \sum_{m=1}^{N_T} ([\lambda_\delta]_m, \phi_{\delta,-}^m) \\ &= - \int_0^T (y_\delta - y_\delta^d, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}, \quad \lambda_{\delta,N}^+ = 0, \end{aligned} \quad (4.16)$$

for $\lambda_\delta := \lambda_\delta(u_\delta)$.

3: Solve the equation

$$\begin{aligned} & \sum_{m=1}^{N_T} \int_{I_m} (\partial_t w_\delta, \phi_\delta) dt + \int_0^T a_h^s(w_\delta, \phi_\delta) dt + \sum_{m=1}^{N_T} ([w_\delta]_{m-1}, \phi_{\delta,+}^{m-1}) \\ &= \int_0^T (u, \nu_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}, \quad w_{\delta,0}^- = 0. \end{aligned} \quad (4.17)$$

4: Solve the equation

$$\begin{aligned} & \sum_{m=1}^{N_T} \int_{I_m} -(\partial_t p_\delta, \phi_\delta) dt + \int_0^T a_h^a(p_\delta, \phi_\delta) dt - \sum_{m=1}^{N_T} ([p_\delta]_m, \phi_{\delta,-}^m) \\ &= - \int_0^T (w, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}, \quad p_{\delta,N}^+ = 0. \end{aligned} \quad (4.18)$$

5: Compute

$$\nabla^2 \hat{J}(u) \nu_\delta = \int_0^T (\alpha \nu_\delta - p, \phi_\delta) dt, \quad \forall \phi_\delta \in V_{h,p}^{k,q}.$$

reaction equation. We construct the discrete Lagrangian defined on $V_{h,p}^{k,q} \times V_{h,p}^{k,q} \times V_{h,p}^{k,q}$ as follows

$$\begin{aligned} & \mathcal{L}(y_\delta, u_\delta, \lambda_\delta) \\ &= \frac{1}{2} \int_0^T \left(\sum_{K \in \mathcal{T}_h} (\|y_\delta - y_\delta^d\|_{L^2(K)}^2 + \alpha \|u_\delta\|_{L^2(K)}^2) \right) dt \\ &+ \sum_{m=1}^{N_T} \left\{ \int_{I_m} ((\partial_t y_\delta, \lambda_\delta) + a_h^s(y_\delta, \lambda_\delta)) dt + ([y_\delta]_{m-1}, \lambda_{\delta,+}^{m-1}) - \int_{I_m} (f_\delta + u_\delta, \lambda_\delta) dt \right\} \\ &+ ((y_0)_\delta - y_{\delta,0}^-, \lambda_{\delta,0}^-). \end{aligned}$$

Differentiating \mathcal{L} with respect to y_δ and applying the same technique in (4.10), we obtain the adjoint equation on each time interval I_m

$$\begin{aligned} & \int_{I_m} (-\partial_t \lambda_\delta, \phi_\delta) + a_h^s(\phi_\delta, \lambda_\delta) dt - ([\lambda_\delta]_m, \phi_{\delta,-}^m) \\ &= - \int_{I_m} (y_\delta - y_\delta^d, \phi_\delta) dt, \quad \forall \lambda_\delta, \phi_\delta \in V_{h,p}^{k,q}. \end{aligned}$$

Similar to cGP(q+1)-method, $\lambda_{\delta,N}^+ = 0$. Now, we use commutativity of DG bilinear form provided in [91], i.e., $a_h^s(\phi_\delta, \lambda_\delta) = a_h^a(\lambda_\delta, \phi_\delta)$. Thus, we arrive at (4.12b). Therefore, OD and DO approaches commute.

For Burgers equation, the difference might arise from the nonlinear term. Thus, we consider the nonlinear term in the Lagrangian

$$\begin{aligned} & \int_{I_m} n_h^s(y, v) \lambda \\ &= \sum_{n=0}^{M-1} \left(- \int_{x_n}^{x_{n+1}} \frac{1}{2} y_h^2 v'(x) dx + v_{n+1} \hat{n}(y_{n+1}^-, y_{n+1}^+) - v_n \hat{n}(y_n^-, y_n^+) \right) \lambda. \end{aligned} \quad (4.19)$$

Differentiating (4.19) with respect to y and applying integration by parts, we obtain

$$(n_h^s)_y(\lambda, y, v) = \sum_{n=0}^{M-1} \left(- \int_{x_n}^{x_{n+1}} y_h \lambda'(x) v(x) dx + v_{n+1} \hat{n}(\lambda_{n+1}^-, \lambda_{n+1}^+) - v_n \hat{n}(\lambda_n^-, \lambda_n^+) \right). \quad (4.20)$$

For the adjoint equation in OD approach, if we apply integration by parts on each element and use upwinding for this term, we arrive at the same expression. Thus, OD and DO approaches commute.

4.3 A Priori Error Estimates for Optimal Control of Diffusion-Convection-Reaction Equation

In this section, firstly, we give some a priori error estimates in the literature and then we present the discrete characteristic function that provides error estimates at arbitrary

time points. Then, we prove some useful lemmas and state the main estimate of this study, particularly for discontinuous Galerkin time discretization.

4.3.1 Some Auxiliary Results

We proceed with the standard estimates derived for finite element methods [32]. Consider the L^2 -projection $\Pi_h : L^2(\Omega) \rightarrow V_{h,p}$ so that

$$\|\Pi_h v - v\|_{L^2(K)} \leq C_\Pi h^{p+1} |v|_{H^{p+1}(K)}, \quad |\Pi_h v - v|_{H^1(K)} \leq C_\Pi h^p |v|_{H^{p+1}(K)} \quad (4.21)$$

for all $v \in H^{p+1}(K)$, $K \in \mathcal{T}_h$ where C_Π is a positive constant and independent of v and h . In addition, as suggested in [130, Section 4], using the study [44], the following estimate holds for all $v \in H^{p+1}(\Omega, \mathcal{T}_h)$

$$|||\Pi_h v - v|||_{DG} \leq (2C_M + 1)C_\Pi h^p |v|_{H^{p+1}(\Omega, \mathcal{T}_h)}, \quad (4.22)$$

where C_M and C_Π are positive constants from (2.3) and (4.21), respectively. In the following we introduce the parabolic projection for $m = 0, \dots, N_T$ and mention the properties given in [130]. Suppose that $X \subset L^2(\Omega)$ is a Hilbert space. Let us denote the space of polynomial functions depending on time as follows:

$$P^\alpha(I_m, X) = \left\{ v \in L^2(0, T; L^2(\Omega)) : v = \sum_{s=0}^{\alpha} t^s \phi_{s,m}, t \in I_m, \phi_{s,m} \in X \right\}.$$

A space-time projection π of $y \in C(0, T; H^1(\Omega))$ into $V_{h,p}^{k,q}$ is employed for the convergence estimates. Time projection P of $y \in C(0, T; H^1(\Omega))$ is defined as

$$\begin{aligned} Py &\in \{v \in L^2(Q_T) : v|_{I_m} \in P^q(I_m, L^2(\Omega))\}, \\ \int_{I_m} (Py - y, t^j v) dt &= 0, \quad \forall v \in L^2(\Omega), j = 0, \dots, q-1, \\ (Py)_-^m &= y(t^m). \end{aligned}$$

In addition, for $m = 0, \dots, N_T$, with $y \in C(0, T; H^1(\Omega))$, $\pi y \in V_{h,p}^{k,q}$ is defined as

$$\begin{aligned} \pi y &= \Pi_h(Py) \iff ((\pi y)(t), v) = ((Py)(t), v), \quad \forall v \in V_{h,p}, \forall t \in I_m, \\ \int_{I_m} (\pi y - y, v) dt &= \int_{I_m} ((Py, v) - (y, v)) dt = 0, \quad \forall v \in V_{h,p}^{k,q-1}, \quad (4.23) \\ ((\pi y)_-^m - y(t^m), v) &= (((Py)_-^m, v) - (y(t^m), v)) = 0, \quad \forall v \in V_{h,p}. \end{aligned}$$

We note that the definition of the projection π is likewise in the study [115].

We give some estimates from [130, Lemma 4.3, 4.5], which we need in the proofs.

Lemma 4.1. *Suppose that $y \in W^{q+1, \infty}(I_m, H^1(\Omega))$ such that $y = 0$ on $\partial\Omega$. Then,*

$$\begin{aligned} \|y(t) - Py(t)\| &\leq C_P k_m^{q+1} |y|_{W^{q+1, \infty}(I_m, L^2(\Omega))} \quad \forall t \in I_m, \\ |y(t) - Py(t)|_{H^1(\Omega)} &\leq C_P k_m^{q+1} |y|_{W^{q+1, \infty}(I_m, H^1(\Omega))} \quad \forall t \in I_m, \quad (4.24) \\ |||y(t) - Py(t)|||_{DG} &\leq C_P k_m^{q+1} |y|_{W^{q+1, \infty}(I_m, H^1(\Omega))} \quad \forall t \in I_m. \end{aligned}$$

Lemma 4.2. *Suppose that $y \in W^{q+1,\infty}(I_m, H^1(\Omega)) \cap L^\infty(I_m, H^{p+1}(\Omega))$ such that $y = 0$ on $\partial\Omega$. Then,*

$$\begin{aligned} \|y(t) - \pi y(t)\| &\leq C_\pi (h^{p+1} + k_m^{q+1}) \|y\|_R \quad \forall t \in I_m, \\ \| \|y(t) - \pi y(t)\| \|_{DG} &\leq C_\pi (h^p + k_m^{q+1}) \|y\|_R \quad \forall t \in I_m, \end{aligned} \quad (4.25)$$

where $\|y\|_R = \max(|y|_{W^{q+1,\infty}(I_m, H^1(\Omega))}, |y|_{L^\infty(I_m, H^{p+1}(\Omega))})$ and C_π is a positive constant independent of h, k_m, m and y .

Lemma 4.3. *There exists a positive constant C_A which is independent of h, v_h, w_h, ϵ such that*

$$\begin{aligned} a^d(y(t) - \Pi_h y(t), v_h) &\leq C_A \epsilon h^p \|y(t)\|_{H^{p+1}(\Omega)} \| \|v_h\| \|_{DG}, \\ &\text{a.e. } t \in (0, T), y \in L^2(0, T; H^{p+1}(\Omega)), v_h \in V_{h,p}, \\ a^d(v_h, w_h) &\leq C_A \epsilon \| \|v_h\| \|_{DG} \| \|w_h\| \|_{DG}, \quad v_h, w_h \in V_{h,p}. \end{aligned} \quad (4.26)$$

Proof. The proof in [41, Lemma 3.8] is adapted to the bilinear form (3.14) using the estimate (4.22). \square

Remark 4.1. A similar estimate for the bilinear form arising from the non-symmetric interior penalty Galerkin method can be found in [130, Lemma 4.2].

Lemma 4.4. *The bilinear form $a^d(\cdot, \cdot)$ satisfies the coercivity inequality*

$$a^d(v_h, v_h) \geq \frac{\epsilon}{2} \| \|v_h\| \|_{DG}^2, \quad \forall v_h \in V_{h,p}. \quad (4.27)$$

Proof. The proof in [41, Corollary 3.10] is adopted to the bilinear form (3.14) using the norm (2.1). \square

4.3.2 Discrete Characteristic Function

We use the discrete characteristic function in order to provide error estimates at arbitrary time points as suggested in [31]. We can work on $[0, k)$ instead of I_m , since the construction of the discrete characteristic function is invariant under translation. We consider polynomials $s \in \mathcal{P}_q(0, k)$ and the discrete approximation of $\chi_{[0,t)} s$ of s which is a polynomial

$$\tilde{s} \in \{\tilde{s} \in \mathcal{P}_q(0, k) : \tilde{s}(0) = s(0)\} \text{ such that } \int_0^k \tilde{s} z = \int_0^t s z, \quad \forall z \in \mathcal{P}_{q-1}(0, k).$$

This definition can be extended from $\mathcal{P}_q(0, k)$ to $V_{h,p}^{k,q}$. The discrete approximation of $\chi_{[0,t)} v$ for $v \in V_{h,p}^{k,q}$ is written as $\tilde{v} = \sum_{i=0}^q \tilde{s}_i(t) v_i$. On account of these inequalities, the following estimate is given in [130]

$$\int_{I_m} \| \|\tilde{w}\| \|_{DG}^2 dt \leq C_D \int_{I_m} \| \|w\| \|_{DG}^2 dt, \quad C_D = C_D(q). \quad (4.28)$$

A suitable discrete approximation $\chi_{(t,t^n]}v_h$ must be constructed for the adjoint problem, as it is noted in the proof of [29, Theorem 3.8]. The discrete approximation of $\chi_{(t,t^{N_T}]}s$ is a polynomial

$$\tilde{s} \in \{\tilde{s} \in \mathcal{P}_q(t^{N_T-1}, t^{N_T}) : \tilde{s}(t^{N_T}) = s(t^{N_T})\} \text{ such that } \int_{t^{N_T-1}}^{t^{N_T}} \tilde{s}z = \int_t^{t^{N_T}} sz,$$

$\forall z \in \mathcal{P}_{q-1}(t^{N_T-1}, t^{N_T})$. This definition can be extended from $\mathcal{P}_q(t^{N_T-1}, t^{N_T})$ to $V_{h,p}^{k,q}$ and the estimates above can be modified for the adjoint [29, Theorem 3.8].

4.3.3 Main Result

We proceed with the derivation of convergence estimates for the optimality system and its space-time dG approximation. Firstly, we take $c(y, u) = 0$ as the diffusion-convection-reaction equation. We define the auxiliary state and adjoint equation which are needed for a priori error analysis

$$\sum_{m=1}^{N_T} \int_{I_m} (\partial_t y_\delta^u, v_\delta) dt + \int_0^T a_h^s(y_\delta^u, v_\delta) dt + \sum_{m=1}^{N_T} ([y_\delta^u]_{m-1}, v_{\delta,+}^{m-1}) = \int_0^T (f_\delta + u, v_\delta) dt, \quad (4.29a)$$

$$y_{\delta,-}^{u,0} = (y_0)_\delta,$$

$$\sum_{m=1}^{N_T} \int_{I_m} (-\partial_t \lambda_\delta^u, \phi_\delta) dt + \int_0^T a_h^a(\lambda_\delta^u, \phi_\delta) dt - \sum_{m=1}^{N_T} ([\lambda_\delta^u]_m, \phi_{\delta,-}^m) = - \int_0^T (y_\delta^u - y_\delta^d, \phi_\delta) dt, \quad (4.29b)$$

$$\lambda_{\delta,+}^{u,N} = 0.$$

Following [54], we assume that the reaction term satisfies $|r| \leq C_r$ a.e. in Ω ; the velocity field is bounded by a constant C_β a.e. in Ω .

We shall prove some useful lemmas before stating the main theorem of this study.

Lemma 4.5. *Let $(y_\delta, \lambda_\delta)$ and $(y_\delta^u, \lambda_\delta^u)$ be the solutions of (4.12) and (4.29), respectively. Then, there exists a constant C independent of h and k such that*

$$\sup_{t \in I_n} \|y_\delta^u(t) - y_\delta(t)\| + \sup_{t \in I_n} \|\lambda_\delta^u(t) - \lambda_\delta(t)\| \leq C \int_0^{t_n} \|u - u_\delta\| dt. \quad (4.30)$$

Proof. Firstly, we shall study the fully discrete state equation on each subinterval I_m . We subtract (4.12a) from (4.29a) to obtain

$$\int_{I_m} (\partial_t \theta, v_\delta) dt + ([\theta]_{m-1}, v_{\delta,+}^{m-1}) + \int_{I_m} a_h^s(\theta, v_\delta) dt = \int_{I_m} (u - u_\delta, v_\delta) dt, \quad (4.31)$$

where $\theta = y_\delta^u - y_\delta$. We substitute $v_\delta = 2\theta$ in (4.31). Then,

$$\int_{I_m} 2(\partial_t \theta, \theta) dt + 2([\theta]_{m-1}, \theta_+^{m-1}) = \|\theta_-^m\|^2 - \|\theta_-^{m-1}\|^2 + \|[\theta]_{m-1}\|^2, \quad (4.32)$$

is achieved. For the right-hand side, we employ Cauchy-Schwarz, Young inequalities, Poincaré inequality (2.4) and the definition of dG norm (2.1). For the left-hand side, we use (4.27) for diffusion term and follow the technique in (see [54, Theorem 5.1]) for convection and reaction terms. Then, we derive the following estimate in the middle of (4.33)

$$\begin{aligned}
& \|\theta_-^m\|^2 - \|\theta_-^{m-1}\|^2 + \frac{\epsilon}{2} \int_{I_m} \|\theta\|_{DG}^2 dt + 2C_0 \int_{I_m} \|\theta\|^2 dt \\
& + \frac{\epsilon}{2} \int_{I_m} \left(\sum_{K \in \mathcal{T}_h} \left(\|\theta\|_{\partial K^- \cap \Gamma^-}^2 + \|[\theta]\|_{\partial K^- \setminus \Gamma^-}^2 + \|\theta\|_{\partial K^+ \cap \Gamma^+}^2 \right) \right) dt \\
\leq & \|\theta_-^m\|^2 - \|\theta_-^{m-1}\|^2 + \frac{\epsilon}{2} \int_{I_m} \|\theta\|_{DG}^2 dt + 2C_0 \int_{I_m} \|\theta\|^2 dt \\
& + \int_{I_m} \left(\sum_{K \in \mathcal{T}_h} \left(\|\theta\|_{\partial K^- \cap \Gamma^-}^2 + \|[\theta]\|_{\partial K^- \setminus \Gamma^-}^2 + \|\theta\|_{\partial K^+ \cap \Gamma^+}^2 \right) \right) dt \\
\leq & C \int_{I_m} \|u - u_\delta\|^2 dt. \tag{4.33}
\end{aligned}$$

We note that the lower bound on the left-hand side of (4.33) has been added after deriving the estimate in the middle for the clearance of the proof and will be used later. Now, we proceed by substituting $v_\delta = 2\tilde{\theta}$ into (4.31). We employ the discrete characteristic function as in the proof of [130, Theorem 5.2] to obtain an estimate at arbitrary points and use the properties given there. With $z = \arg \sup_{\bar{I}_m} \|\theta(t)\|$, the discrete characteristic function defined in Section 4.3.2 leads to

$$\int_{I_m} (\partial_t \theta, \tilde{\theta}) dt = \int_{t_{m-1}}^z (\partial_t \theta, \theta) dt, \quad \tilde{\theta}_+^{m-1} = \theta_+^{m-1}, \quad [\tilde{\theta}]_{m-1} = [\theta]_{m-1}, \tag{4.34}$$

$$\int_{I_m} 2(\partial_t \theta, \tilde{\theta}) dt + 2([\theta]_{m-1}, \tilde{\theta}_+^{m-1}) = \|\theta(z)\|^2 - \|\theta_-^{m-1}\|^2 + \|[\theta]_{m-1}\|^2. \tag{4.35}$$

We use (4.34)-(4.35) and the inequality $\|\theta_-^{m-1}\| \leq \sup_{t \in I_{m-1}} \|\theta(t)\|$ to bound the terms arising in the time derivative. We proceed by moving $2 \int_{I_m} a_h(\theta, \tilde{\theta}) dt$ to the right-hand side. We employ (4.26) for the diffusion term, the proof of [54, Theorem 5.1] for the convection term. The reaction term and the control on the right-hand side is bounded by using Cauchy-Schwarz and Young inequalities (2.4) and (2.1) such that $\|\cdot\|^2 \leq C \|\cdot\|_{DG}^2$ is satisfied for a positive constant C . We eliminate the term $\|\tilde{\theta}\|_{DG}^2$ on the right-hand side by using (4.28). Then, we obtain the following inequality

$$\begin{aligned}
& \sup_{t \in I_m} \|\theta(t)\|^2 - \sup_{t \in I_{m-1}} \|\theta(t)\|^2 \\
& \leq C_b \int_{I_m} \|\theta\|_{DG}^2 dt + \int_{I_m} \sum_{K \in \mathcal{T}_h} \left(\|\theta\|_{\partial K^+ \cap \Gamma^+}^2 + \|[\theta]\|_{\partial K^- \setminus \Gamma^-}^2 \right) dt + C \int_{I_m} \|u - u_\delta\|^2 dt \\
& \leq C'_b \int_{I_m} \left(\|\theta\|_{DG}^2 + \sum_{K \in \mathcal{T}_h} \left(\|\theta\|_{\partial K^+ \cap \Gamma^+}^2 + \|[\theta]\|_{\partial K^- \setminus \Gamma^-}^2 \right) \right) dt + C \int_{I_m} \|u - u_\delta\|^2 dt, \tag{4.36}
\end{aligned}$$

where $C_b = C(1 + C_D)(\epsilon C_A + C_S(C_r + C_\beta))$, $C'_b = \max\{1, C_b\}$. In order to eliminate the terms θ on the right-hand side of (4.36), we use (4.33) multiplying it by $C''_b = \frac{2}{\epsilon} C'_b$. By adding these inequalities and denoting $\Theta_m = \sup_{t \in I_m} \|\theta(t)\|^2 + C''_b \|\theta^m\|^2$, we arrive at

$$\Theta_m - \Theta_{m-1} \leq C(1 + C''_b) \int_{I_m} \|u - u_\delta\|^2 dt. \quad (4.37)$$

We sum (4.37) over $m = 1, \dots, n \leq N_T$ and use $\theta = 0$ at $t = 0$ to derive the estimate

$$\sup_{t \in I_n} \|\theta(t)\|^2 = \sup_{t \in I_n} \|y_\delta^u(t) - y_\delta(t)\|^2 \leq C \int_0^{t_n} \|u - u_\delta\|^2 dt. \quad (4.38)$$

Secondly, we proceed with the adjoint equation subtracting (4.12b) from (4.29b) and using $\zeta = \lambda_\delta^u - \lambda_\delta$. A discrete approximation to $\chi_{(t, t_m]} v_h$ specified for the adjoint problem must be used, as we discussed in Section 4.3.2. Then, this leads to

$$\int_{I_m} 2(-\partial_t \zeta, \tilde{\zeta}) dt - 2([\zeta]_m, \tilde{\zeta}^m) = \|\zeta(z)\|^2 - \|\zeta^m\|^2 + \|[\zeta]_m\|^2, \quad (4.39)$$

where $z = \arg \sup_{I_m} \|\zeta(t)\|$. In addition, the inequalities $\|\zeta^m\|^2 \leq \sup_{I_{N_T - m + 2}} \|\zeta(t)\|^2$ and $\|\zeta(z)\|^2 = \sup_{I_{N_T - m + 1}} \|\zeta(t)\|^2$ are needed. Then, we follow the same idea used to derive (4.38) to reach the inequality

$$\sup_{t \in I_{N_T - m + 1}} \|\zeta(t)\|^2 - \sup_{t \in I_{N_T - m + 2}} \|\zeta(t)\|^2 \leq C k_m \int_{t \in I_m} \|u - u_\delta\|^2 dt. \quad (4.40)$$

We shall sum (4.40) over $m = N_T, \dots, n \geq 1$ and use $\zeta = 0$ at $t = t_{N_T}$. The final result (4.30) follows from standard algebra, (4.38) and (4.40). \square

We shall proceed with the estimate between the exact and the approximate control.

Lemma 4.6. *Let (y, λ, u) and $(y_\delta, \lambda_\delta, u_\delta)$ be the solutions of (2.6) and (4.12), respectively. Then, we have*

$$\|u - u_\delta\|_{L^2(0, T; L^2(\Omega))} \leq \frac{1}{\alpha} \|\lambda - \lambda_\delta^u\|_{L^2(0, T; L^2(\Omega))}. \quad (4.41)$$

Proof. We apply the technique used for the steady-state optimal control problem in [91, Section 4.2]. We start using the continuous and fully-discrete optimality conditions (2.6c)-(4.12c) to obtain the following equation

$$\begin{aligned} \alpha \|u - u_\delta\|_{L^2(0, T; L^2(\Omega))}^2 &= \alpha \int_0^T (u - u_\delta, u - u_\delta) dt \\ &= \int_0^T (\alpha u - \lambda, u - u_\delta) dt - \int_0^T (\alpha u_\delta - \lambda_\delta, u - u_\delta) dt + \int_0^T (\lambda - \lambda_\delta, u - u_\delta) dt \\ &= \int_0^T (\lambda - \lambda_\delta^u, u - u_\delta) dt + \int_0^T (\lambda_\delta^u - \lambda_\delta, u - u_\delta) dt = J_1 + J_2. \end{aligned} \quad (4.42)$$

We use Cauchy-Schwarz and Young inequalities to show that

$$0 \leq J_1 \leq \frac{1}{2\alpha} \|\lambda - \lambda_\delta^u\|_{L^2(0,T;L^2(\Omega))}^2 + \frac{\alpha}{2} \|u - u_\delta\|_{L^2(0,T;L^2(\Omega))}^2. \quad (4.43)$$

We proceed with J_2 and use the auxiliary state equation (4.29) to obtain

$$\begin{aligned} J_2 &= \int_0^T (\lambda_\delta^u - \lambda_\delta, u - u_\delta) dt \\ &= \sum_{m=1}^{N_T} \int_{I_m} (\partial_t(y_\delta^u - y_\delta), \lambda_\delta^u - \lambda_\delta) dt + \int_0^T a_h^s(y_\delta^u - y_\delta, \lambda_\delta^u - \lambda_\delta) dt \\ &\quad + \sum_{m=1}^N ([y_\delta^u - y_\delta]_{m-1}, (\lambda_\delta^u - \lambda_\delta)_+^{m-1}). \end{aligned}$$

We proceed applying integration by parts in time and use the auxiliary adjoint equation (4.29) to arrive at

$$\begin{aligned} J_2 &= - \sum_{m=1}^{N_T} \int_{I_m} (\lambda_\delta^u - \lambda_\delta, \partial_t(y_\delta^u - y_\delta)) dt + \sum_{m=1}^N (y_\delta^u - y_\delta, \lambda_\delta^u - \lambda_\delta)|_{t_{m-1}}^{t_m} \\ &\quad + \int_0^T a_h^s(y_\delta^u - y_\delta, \lambda_\delta^u - \lambda_\delta) dt + \sum_{m=1}^N ([y_\delta^u - y_\delta]_{m-1}, (\lambda_\delta^u - \lambda_\delta)_+^{m-1}) \\ &= - \sum_{m=1}^{N_T} \int_{I_m} (\lambda_\delta^u - \lambda_\delta, \partial_t(y_\delta^u - y_\delta)) dt + \int_0^T a_h^s(y_\delta^u - y_\delta, \lambda_\delta^u - \lambda_\delta) dt \\ &\quad - \sum_{m=1}^N ((y_\delta^u - y_\delta)_-^m, [\lambda_\delta^u - \lambda_\delta]_m) \\ &= - \int_0^T (y_\delta^u - y_\delta, y_\delta^u - y_\delta) dt \leq 0. \end{aligned} \quad (4.44)$$

Then, using (4.42)-(4.44), we derive the final result (4.41). \square

Lemma 4.7. *Let (y, λ) and $(y_\delta^u, \lambda_\delta^u)$ be the solutions of (2.6) and (4.29), respectively. Assume that $y, \lambda \in W^{q+1,\infty}(0, T; H^1(\Omega)) \cap L^\infty(0, T; H^{p+1}(\Omega))$. Then, there exists a constant C independent of h and k such that*

$$\sup_{t \in I_n} \|y - y_\delta^u\| + \sup_{t \in I_n} \|\lambda - \lambda_\delta^u\| \leq \mathcal{O}(h^p + k^{q+1}). \quad (4.45)$$

Proof. Firstly, we integrate (2.6a) over I_m and subtract the result from (4.29a) in order to obtain the following equation

$$\begin{aligned} & \int_{I_m} (\partial_t \xi, v_\delta) dt + ([\xi]_{m-1}, v_{\delta,+}^{m-1}) + \int_{I_m} a_h^s(\xi, v_\delta) dt \\ &= - \left(\int_{I_m} (\partial_t \eta, v_\delta) dt + ([\eta]_{m-1}, v_{\delta,+}^{m-1}) \right) - \int_{I_m} a_h(\eta, v_\delta) dt, \end{aligned} \quad (4.46)$$

where $y - y_\delta^u = (y - \pi y) + (\pi y - y_\delta^u) = \eta + \xi$.

Since we use the same mesh on each time interval, (4.23) leads to the following identity.

$$\int_{I_m} (\partial_t \eta, v_\delta) dt + ([\eta]_{m-1}, v_{\delta,+}^{m-1}) = 0, \quad \forall v_\delta \in V_h^{k,q}. \quad (4.47)$$

We proceed as in the proof of Lemma 4.5 and the proof of [54, Theorem 5.1] by inserting the estimate (4.25) to obtain

$$\begin{aligned} & \int_{I_m} (\partial_t \xi, v_\delta) dt + ([\xi]_{m-1}, v_{\delta,+}^{m-1}) + \int_{I_m} a_h^s(\xi, v_\delta) dt \\ & \leq \frac{\epsilon}{4} \int_{I_m} \|v_\delta\|_{DG}^2 dt + \frac{C_0}{2} \int_{I_m} \|v_\delta\|^2 dt + \frac{1}{2} \int_{I_m} \sum_{K \in \mathcal{T}_h} (\|v_\delta\|_{\partial K^+ \cap \Gamma^+}^2 + \|[\![v_\delta]\!] \|_{\partial K^- \setminus \Gamma^-}^2) dt \\ & + k_m C_A C_\pi (h^{2p} + k^{2q+2}) |y|_R^2 + k_m 2C_\beta C_\pi C_M (h^{2p+1} + k^{2q+2}) |y|_R^2 \\ & + k_m C_\pi \frac{C_\beta C_r}{C_0} (h^{2p+2} + k^{2q+2}) |y|_R^2, \end{aligned} \quad (4.48)$$

where $|y|_R = \max(|y|_{W^{q+1,\infty}(I_m; H^1(\Omega))}, |y|_{L^\infty(I_m; H^{p+1}(\Omega))})$.

Firstly, we shall substitute $v_\delta = 2\xi$ into (4.48) to obtain

$$\begin{aligned} & \|\xi_-^m\|^2 - \|\xi_-^{m-1}\|^2 + \frac{\epsilon}{2} \int_{I_m} \|\xi\|_{DG}^2 dt + C_0 \int_{I_m} \|\xi\|^2 dt \\ & + \int_{I_m} \sum_{K \in \mathcal{T}_h} \left(\|\xi\|_{\partial K^- \cap \Gamma^-}^2 + \frac{1}{2} \|[\![\xi]\!] \|_{\partial K^- \setminus \Gamma^-}^2 + \frac{1}{2} \|\xi\|_{\partial K^+ \cap \Gamma^+}^2 \right) dt \\ & \leq k_m C_b (h^{2p} + h^{2p+1} + h^{2p+2} + k^{2q+2}) |y|_R^2, \end{aligned} \quad (4.49)$$

where $C_b = \max\{C_A C_\pi, 2C_\beta C_\pi C_M, C_\pi \frac{C_\beta C_r}{C_0}\}$.

Secondly, we substitute $v_\delta = 2\tilde{\xi}$ into (4.48) to obtain

$$\begin{aligned}
& \sup_{t \in I_m} \|\xi(t)\|^2 - \sup_{t \in I_{m-1}} \|\xi(t)\|^2 \\
& \leq C'_b \int_{I_m} \|\xi\|_{DG}^2 dt + \int_{I_m} \sum_{K \in \mathcal{T}_h} \left(\|\llbracket \xi \rrbracket\|_{\partial K^- \setminus \Gamma^-}^2 + \|\xi\|_{\partial K^+ \cap \Gamma^+}^2 \right) dt \\
& \quad + k_m C_b (h^{2p} + h^{2p+1} + h^{2p+2} + k^{2q+2}) |y|_R^2 \\
& \leq C''_b \int_{I_m} \left(\|\xi\|_{DG}^2 + \sum_{K \in \mathcal{T}_h} \left(\|\llbracket \xi \rrbracket\|_{\partial K^- \setminus \Gamma^-}^2 + \|\xi\|_{\partial K^+ \cap \Gamma^+}^2 \right) \right) dt \\
& \quad + k_m C_b (h^{2p} + h^{2p+1} + h^{2p+2} + k^{2q+2}) |y|_R^2, \tag{4.50}
\end{aligned}$$

where $C'_b = C(1 + C_D)(\epsilon C_A + C_S(C_\beta + C_\tau))$, $C''_b = \max\{1, C'_b\}$. Now, we proceed as in the proof of Lemma 4.5. We multiply (4.49) by $C'''_b = \frac{2}{\epsilon} C''_b$ in order to eliminate the terms ξ on the right-hand side of (4.50). Then, we add it to (4.50) and denote $\Theta_m = \sup_{t \in I_m} \|\xi(t)\|^2 + C'''_b \|\xi_m^-\|^2$ in order to obtain

$$\Theta_m - \Theta_{m-1} \leq k_m 2C'''_b (h^{2p} + h^{2p+1} + h^{2p+2} + k^{2q+2}) |y|_R^2. \tag{4.51}$$

We sum (4.51) over $m = 1, \dots, n \leq N_T$ to obtain

$$\sup_{t \in I_n} \|\xi(t)\|^2 \leq \mathcal{O}(h^{2p} + k^{2q+2}). \tag{4.52}$$

Thirdly, we integrate (2.6b) over I_m and subtract it from (4.29b) and denote $\lambda - \lambda_\delta^u = (\lambda - \pi\lambda) + (\pi\lambda - \lambda_\delta^u) = \varphi + \mu$. Then, we use the idea in the proof of (4.52) in order to derive

$$\sup_{t \in I_{N-m+1}} \|\mu(t)\|^2 - \sup_{t \in I_{N-m+2}} \|\mu(t)\|^2 \leq C k_m \sup_{t \in I_m} \|\xi(t)\|^2 dt + \mathcal{O}(h^{2p} + k^{2q+2}), \tag{4.53}$$

for $C > 0$. The resulting inequality is summed over $m = N_T, \dots, n \geq 1$. Then, it is combined with (4.52) to derive the final result (4.45). \square

Remark 4.2. For guaranteeing the assumptions on the exact solution, it is necessary to require a higher regularity of the data of the problem.

We state the main estimate of this study by combining Lemmas 4.5-4.7.

Theorem 4.8. *Suppose that (y, λ, u) and $(y_\delta, \lambda_\delta, u_\delta)$ are the solutions of (2.6) and (4.12), respectively. We assume that all conditions of Lemmas 4.5-4.7 are satisfied. Then, there exists a constant C independent of h and k such that*

$$\|y - y_\delta\|_{L^\infty(0, T; L^2(\Omega))} + \|\lambda - \lambda_\delta\|_{L^\infty(0, T; L^2(\Omega))} + \|u - u_\delta\|_{L^2(0, T; L^2(\Omega))} \leq C (h^p + k^{q+1}). \tag{4.54}$$

In Theorem 4.8, the error in the state and control is measured with respect to the norm $L^\infty(0, T; L^2(\Omega))$ and $L^2(0, T; L^2(\Omega))$, respectively. The same norms are used, for example, in [56], too. The former norm is due to the discrete characteristic function

which is used to provide error estimates at arbitrary time points. The latter norm arises from the optimality condition which is shown in Lemma 4.6. On the other hand, we observe that Theorem 4.8 is optimal in time, suboptimal in space in the $L^\infty(0, T; L^2(\Omega))$ norm for the state and $L^2(0, T; L^2(\Omega))$ for the control, i.e. $\mathcal{O}(h^p, k^{q+1})$, using p -degree spatial, q -degree temporal polynomial approximation. However, for example, optimal spatial convergence rate for SIPG discretization combined with backward Euler is achieved using an elliptic projection in [5]. The first reason behind the order reduction in this study is the estimate (4.25) for the space-time projection which is employed to bound the continuity estimate of the bilinear form in Lemma 4.3. The convection term also has an influence on the spatial order reduction since we follow the proof of [54, Theorem 5.1]. In order to improve this suboptimal estimate, the effect of the space-time projector in the bilinear form of the diffusion term must be eliminated.

4.4 Computational Aspects

In this section, we apply cGP(2) and dG(1) methods [99, 135] to the OCP (2.6) and derive the fully discrete formulations of the adjoint equation. cGP(2) and dG(1) methods are convergent of order 3 and 2, respectively. cGP(2) method is super-convergent of order 4, and dG(1) is of order 3 at the nodal points, i.e. at the endpoints of the time intervals [120, 99].

cGP(2)-method

In cG(2)-method, the state y_δ and the control u_δ are approximated in the time interval $I_m = (t_{m-1}, t_m]$ by

$$y_\delta := y_\delta(t) = \sum_{i=0}^q Y_m^i \phi_{m,i}(t), \quad u_\delta := u_\delta(t) = \sum_{i=0}^q U_m^i \phi_{m,i}(t), \quad \forall t \in I_m.$$

Equivalently, we have

$$\begin{aligned} y_\delta &= Y_m^0 \phi_{m,0}(t) + Y_m^1 \phi_{m,1}(t) + Y_m^2 \phi_{m,2}(t), \quad \forall t \in I_m, \\ u_\delta &= U_m^0 \phi_{m,0}(t) + U_m^1 \phi_{m,1}(t) + U_m^2 \phi_{m,2}(t), \quad \forall t \in I_m, \end{aligned}$$

where $\phi_{m,j} \in \mathbb{P}_2(I_m)$ are the orthogonal quadratic Lagrange basis functions on \bar{I}_m .

Discrete time steps on \bar{I}_m are chosen according to the 3-point Gauss-Lobatto rule on I_m ,

$$t_{m,0} = t_{m-1}, \quad t_{m,1} = (t_{m-1} + t_m)/2, \quad t_{m,2} = t_m,$$

with the reference weights $\hat{w}_0 = \hat{w}_2 = 1/3$, $\hat{w}_1 = 4/3$.

The test functions are chosen as $\psi_{m,i}^s \in \mathbb{P}_1(I_m)$ such that

$$\hat{\psi}_i^s(\hat{t}_\mu) = (\hat{w}_\mu)^{-1} \delta_{i,\mu} \quad i, \mu = 1, 2, \quad \iff \quad \hat{\psi}_1^s = \frac{3}{4}(1 - \hat{t}), \quad \hat{\psi}_2^s = 3\hat{t}.$$

Using the transformation $T_n : \hat{I} \rightarrow I_m$ where $\hat{I} = (-1, 1]$

$$t = T_n(\hat{t}) = \frac{t_{m-1} + t_m}{2} + \frac{k_m}{2}\hat{t} \in I_m, \quad \forall \hat{t} \in \hat{I}, \quad m = 1, \dots, N_T,$$

they are transformed to the interval I_m .

The initial conditions on each time interval I_m are

$$Y_m^0 = y_\delta(t_{m-1}) \quad \text{if } m \geq 2 \quad \text{or} \quad Y_m^0 = y_0 \quad \text{if } m = 0.$$

For the state equation, on each time interval I_m , the following linear system has to be solved for Y_m^1, Y_m^2

$$\begin{pmatrix} \mathbf{M} + \frac{k_m}{2}\mathbf{A} & \frac{1}{4}\mathbf{M} \\ -4\mathbf{M} & 2\mathbf{M} + \frac{k_m}{2}\mathbf{A} \end{pmatrix} \begin{pmatrix} Y_m^1 \\ Y_m^2 \end{pmatrix} = \begin{pmatrix} (\frac{5}{4}\mathbf{M} - \frac{k_m}{4}\mathbf{A})Y_m^0 + \frac{k_m}{2}\mathbf{M}(\frac{1}{2}U_m^0 + U_m^1) + \frac{k_m}{2}(\frac{1}{2}\mathbf{F}_h(t_{m,0}) + \mathbf{F}_h(t_{m,1})) \\ (-2\mathbf{M} + \frac{k_m}{2}\mathbf{A})Y_m^0 + \frac{k_m}{2}\mathbf{M}(U_m^2 - U_m^0) + \frac{k_m}{2}(\mathbf{F}_h(t_{m,2}) - \mathbf{F}_h(t_{m,0})) \end{pmatrix}. \quad (4.55)$$

The solution of state at the $t = t_m$ is given by $y_{h,m} = Y_m^2$.

cGP(2) method for solving the adjoint equation is constructed in a similar way as for the state equation, but by integrating backwards.

$$\lambda_\delta := \lambda_\delta(t) = \sum_{i=0}^q \lambda_m^i \phi_{m,i}(t), \quad \forall t \in I_m.$$

Equivalently, we have

$$\lambda_\delta = \lambda_m^0 \phi_{m,0}(t) + \lambda_m^1 \phi_{m,1}(t) + \lambda_m^2 \phi_{m,2}(t), \quad \forall t \in I_m.$$

Discrete time steps on \bar{I}_m are chosen according to the 3-point Gauss-Lobatto rule on I_m ,

$$t_{m,0} = t_{m-1}, \quad t_{m,1} = (t_{m-1} + t_m)/2, \quad t_{m,2} = t_m,$$

with the reference weights $\hat{w}_0 = \hat{w}_2 = 1/3$, $\hat{w}_1 = 4/3$.

The test functions are chosen as $\psi_{m,i}^a \in \mathbb{P}_1(I_m)$ such that

$$\hat{\psi}_i^a(\hat{t}_\mu) = (\hat{w}_\mu)^{-1} \delta_{i,\mu} \quad i, \mu = 0, 1, \quad \iff \quad \hat{\psi}_1^a = -3\hat{t}, \quad \hat{\psi}_2^a = \frac{3}{4}(1 + \hat{t}).$$

The initial conditions on each time interval I_m are

$$\lambda_m^2 = \lambda_\delta|_{I_{m-1}}(t_m) \quad \text{if } m \leq 2N_T - 1 \quad \text{or} \quad P_m^2 = 0 \quad \text{if } m = 2N_T - 1.$$

The system of equations to be solved for the adjoint equation becomes

$$\begin{pmatrix} \mathbf{M} + \frac{k_m}{2}\mathbf{A} & \frac{1}{4}\mathbf{M} \\ -4\mathbf{M} & 2\mathbf{M} + \frac{k_m}{2}\mathbf{A} \end{pmatrix} \begin{pmatrix} \lambda_m^1 \\ \lambda_m^0 \end{pmatrix} \\ = \begin{pmatrix} (\frac{5}{4}\mathbf{M} - \frac{k_m}{4}\mathbf{A})\lambda_m^2 - \frac{k_m}{2}\mathbf{M}(\frac{1}{2}Y_m^2 + Y_m^1) + \frac{k_m}{2}(\frac{1}{2}\mathbf{Y}_h^d(t_{m,2}) + \mathbf{Y}_h^d(t_{m,1})) \\ (-2\mathbf{M} + \frac{k_m}{2}\mathbf{A})\lambda_m^2 - \frac{k_m}{2}\mathbf{M}(Y_m^0 - Y_m^2) + \frac{k_m}{2}(\mathbf{Y}_h^d(t_{m,0}) - \mathbf{Y}_h^d(t_{m,2})) \end{pmatrix}. \quad (4.56)$$

The solution at the $t = t_{m-1}$ is given as $\lambda_{h,m-1} = \lambda_m^0$.

dG(1)-method

For dG(1)-method, the state y_δ and the control u_δ are approximated in the time interval $I_m = (t_{m-1}, t_m]$ by

$$y_\delta := y_\delta(t) = \sum_{i=1}^q Y_m^i \phi_{m,i}^s(t), \quad u_\delta := u_\delta(t) = \sum_{i=0}^{q-1} U_m^i \phi_{m,i}^a(t), \quad \forall t \in I_m.$$

Equivalently, we have

$$\begin{aligned} y_\delta &= Y_m^1 \phi_{m,1}^s(t) + Y_m^2 \phi_{m,2}^s(t), \quad \forall t \in I_m, \\ u_\delta &= U_m^0 \phi_{m,0}^a(t) + U_m^1 \phi_{m,1}^a(t), \quad \forall t \in I_m, \end{aligned}$$

using Gauss-Radau quadrature rule, where $\phi_{m,j}^s \in \mathbb{P}_1(I_m)$ are the linear Lagrange basis functions on \bar{I}_m .

Discrete time steps on \bar{I}_m are chosen according to the right-handed 2-point Gauss-Radau rule on I_m

$$t_{m,1}^s = t_{m-1} + \frac{k_m}{3}, \quad t_{m,2}^s = t_m,$$

we use the reference weights

$$\hat{w}_1^s = 3/2, \quad \hat{w}_2^s = 1/2.$$

The test functions are chosen as $\psi_{m,i}^s \in \mathbb{P}_1(I_m)$ such that

$$\hat{\psi}_i^s(\hat{t}_\mu) = (\hat{w}_\mu^s)^{-1} \delta_{i,\mu} \quad i, \mu = 1, 2, \quad \iff \quad \hat{\psi}_1^s = \frac{1-\hat{t}}{2}, \quad \hat{\psi}_2^s = \frac{3\hat{t}+1}{2}.$$

The initial conditions for state equation on each time interval I_m are

$$Y_m^0 = (y_\delta)_{m-1}^- \quad \text{if } m \geq 2 \quad \text{or} \quad (y_\delta)_{m-1}^- = y_0^- \quad \text{if } m = 0.$$

On each time interval $I_m = (t_{m-1}, t_m]$, we solve the following linear system for Y_m^1, Y_m^2

$$\begin{pmatrix} \frac{3}{4}\mathbf{M} + \frac{k_m}{2}\mathbf{A} & \frac{1}{4}\mathbf{M} \\ -\frac{9}{4}\mathbf{M} & \frac{5}{4}\mathbf{M} + \frac{k_m}{2}\mathbf{A} \end{pmatrix} \begin{pmatrix} Y_m^1 \\ Y_m^2 \end{pmatrix} = \begin{pmatrix} \mathbf{M}Y_m^0 + \frac{k_m}{2}\mathbf{F}_h(t_{m,1}) + \frac{k_m}{4}\mathbf{M}(U_m^0 + U_m^1) \\ -\mathbf{M}Y_m^0 + \frac{k_m}{2}\mathbf{F}_h(t_{m,2}) + \frac{k_m}{4}\mathbf{M}(3U_m^1 - U_m^0) \end{pmatrix}. \quad (4.57)$$

Again the discrete state at $t = t_m$ is given as $y_{h,t_m} = Y_m^2$.

Let λ_δ be the approximate solution of the adjoint

$$\lambda_\delta := \lambda_\delta(t) = \sum_{i=0}^{q-1} U_m^i \phi_{m,i}^a(t), \quad \forall t \in I_m.$$

Equivalently, we have

$$\lambda_\delta = \lambda_m^0 \phi_{m,0}^a(t) + \lambda_m^1 \phi_{m,1}^a(t), \quad \forall t \in I_m,$$

using linear orthogonal Lagrange functions and Gauss-Radau points. We note that the integrals in the variational formulation of the state equation are approximated using the right-handed 2-point Gauss-Radau rule whereas left-handed Gauss-Radau points are used for the adjoint equation.

Discrete time steps on \bar{I}_m are chosen according to the left-handed 2-point Gauss-Radau rule

$$t_{m,0}^a = t_{m-1}, \quad t_{m,1}^a = t_{m-1} + \frac{k_m}{3},$$

we use the corresponding reference weights

$$\hat{w}_0^a = 1/2, \quad \hat{w}_1^a = 3/2.$$

The test functions are chosen $\psi_{m,i}^a \in \mathbb{P}_1(I_m)$ such that

$$\hat{\psi}_0^a = \frac{1-3\hat{t}}{2}, \quad \hat{\psi}_1^a = \frac{1+\hat{t}}{2}.$$

The initial conditions for the adjoint equation on each time interval I_m are

$$\lambda_m^2 = (\lambda_\delta)_m^+ \quad \text{if } m \leq N_T \quad \text{or} \quad \lambda_m^2 = 0 \quad \text{if } m = N_T.$$

On each time interval $I_m = (t_{m-1}, t_m]$, we solve the following linear system for λ_m^1, λ_m^0

$$\begin{pmatrix} \frac{3}{4}\mathbf{M} + \frac{k_m}{2}\mathbf{A} & \frac{1}{4}\mathbf{M} \\ -\frac{9}{4}\mathbf{M} & \frac{5}{4}\mathbf{M} + \frac{k_m}{2}\mathbf{A} \end{pmatrix} \begin{pmatrix} \lambda_m^1 \\ \lambda_m^0 \end{pmatrix} = \begin{pmatrix} \mathbf{M}\lambda_m^2 + \frac{k_m}{2}\mathbf{Y}_h^d(t_{m,1}) - \frac{k_m}{4}\mathbf{M}(Y_m^1 + Y_m^2) \\ -\mathbf{M}\lambda_m^2 + \frac{k_m}{2}\mathbf{Y}_h^d(t_{m,0}) - \frac{k_m}{4}\mathbf{M}(3Y_m^1 - Y_m^2) \end{pmatrix}. \quad (4.58)$$

Similar to the cGP(q+1) method, the discrete adjoint at t_{m-1} is given as $\lambda_{h,m-1} = \lambda_m^0$.

The main drawback of dG time discretization is the solution of large coupled linear systems in block form. Because we are using constant time steps, the coupled matrices on the right-hand sides of (4.55)-(4.58) have to be decomposed (LU block factorization) at the beginning of the integration. Then, the state and the adjoint equations are solved at each time step by forward elimination and back substitution using the block factorized matrices. Advantage of the variational time integration methods above is that only one of the variables is needed in the coupled system of equations (4.55)-(4.58) to determine the discrete state and the adjoint. The form of the linear systems (4.55)-(4.58) are the same, which is not the case when arbitrary test functions ψ^s and ψ^a are used. Using this technique, the preconditioner given in [135] can be applied both to the state and the adjoint equations. Additionally, the orthogonal test functions lead to sparse matrices with approximately half of the non-zero entries than of variational time discretization methods with nodal basis test functions [50].

4.4.1 Numerical Results

In this section, we present some numerical results, which are presented in our work [3]. We measure the error in the state and control approximation in terms of $L^\infty(0, 1; L^2(\Omega))$ and $L^2(0, 1; L^2(\Omega))$ norm, respectively. We have used discontinuous piecewise linear polynomials in space. In all numerical examples, we have taken $h = \mathcal{O}(k)$. We follow *optimize-then-discretize* approach.

Example 1: The first test problem is a convection dominated problem with smooth solutions depending implicitly on the diffusion term

$$Q = (0, 1] \times \Omega, \quad \Omega = (0, 1)^2, \quad \epsilon = 10^{-5}, \quad \beta = (1, 1)^T, \quad r = 1 \quad \text{and} \quad \alpha = 1.$$

The source function f , the desired state y_d and the initial condition y_0 are computed from (2.6) using the following exact solutions of the state and control, respectively,

$$\begin{aligned} y(x, t) &= 50 \exp(-t)xy(x-1)(y-1) \cos(5x-5y-5), \\ u(x, t) &= 100 \exp(-t)(1-t)xy(x-1)(y-1) \sin(5x+5y-5). \end{aligned}$$

In Figures (4.1-4.2), we present the numerical solutions and the errors for the state and the control at $t = 0.5$. We observe that the problem is approximated well and the error is equally-distributed over the whole domain.

In Table 4.1, we give the errors for cGP(2) and dG(1) methods; cGP(2) method yields smaller error than dG(1) method. For cGP(2) method, theoretical convergence rate $\mathcal{O}(h^2, k^3)$ leads to $\mathcal{O}(h^2)$ with $h = k$. We achieve this rate numerically. For dG(1) method, the numerical results indicate a higher order experimental order of convergence, i.e. $\mathcal{O}(h^2)$, than the one shown in Theorem 4.8, which is $\mathcal{O}(h)$ with $h = k$.

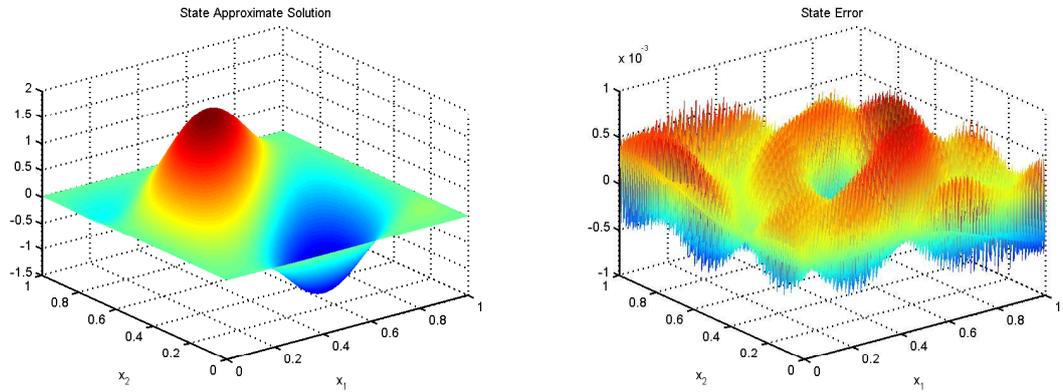


Figure 4.1: DCR Eqn: Numerical solution of the state (*left*), error (*right*) for Example 1 at $t=0.5$ with $h = k = 1/80$ by dG(1) method

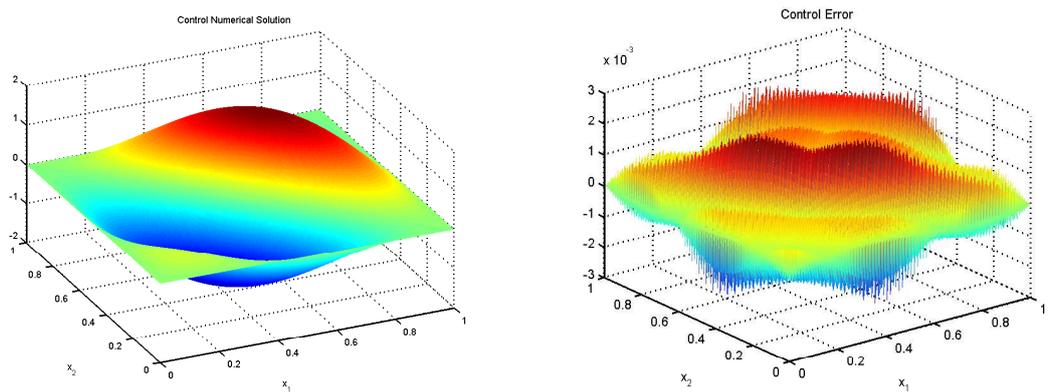


Figure 4.2: DCR Eqn: Numerical solution of the control (*left*), error (*right*) for Example 1 at $t=0.5$ with $h = k = 1/80$ by dG(1) method

Table 4.1: DCR Eqn: Example 1 by cGP(2) and dG(1) (*in parenthesis*) method.

$h = k$	$\ y - y_\delta\ $	Rate	$\ u - u_\delta\ $	Rate
$\frac{1}{5}$	7.43e-2(8.01e-2)	-(-)	1.59e-1(2.66e-1)	- (-)
$\frac{1}{10}$	1.98e-2(2.16e-2)	1.91(1.89)	4.84e-2(6.21e-2)	1.71(2.10)
$\frac{1}{20}$	5.18e-3(5.63e-3)	1.93(1.94)	1.31e-2(1.57e-2)	1.88(1.98)
$\frac{1}{40}$	1.35e-3(1.43e-3)	1.94(1.98)	3.41e-3(3.63e-3)	1.94(2.11)
$\frac{1}{80}$	3.43e-4(3.61e-4)	1.97(1.99)	8.90e-4(9.18e-4)	1.94(1.98)

Example 2: This example is a convection dominated OCP constructed from Example 2 in [56] by adding the reaction term.

$$Q = (0, 1] \times \Omega, \quad \Omega = (0, 1)^2, \quad \epsilon = 10^{-5}, \quad \beta = (0.5, 0.5)^T, \quad r = 3, \quad \alpha = 1.$$

The source function f , the desired state y_d and the initial condition y_0 are computed from (2.6) using the following exact solutions of the control and state, respectively,

$$u(x_1, x_2, t) = \sin(\pi t) \sin(2\pi x_1) \sin(2\pi x_2) \exp\left(\frac{-1 + \cos(t_x)}{\sqrt{\epsilon}}\right),$$

$$y(x_1, x_2, t) = u \left(\frac{1}{2\sqrt{\epsilon}} \sin(t_x) + 8\epsilon\pi^2 + \frac{\sqrt{\epsilon}}{2} \cos(t_x) - \frac{1}{2} \sin^2(t_x) \right) - \pi \cos(\pi t) \sin(2\pi x_1) \sin(2\pi x_2) \exp\left(\frac{-1 + \cos(t_x)}{\sqrt{\epsilon}}\right),$$

where $t_x = t - 0.5(x_1 + x_2)$. As opposed to the previous example, the exact solution of PDE constrained depends on the diffusion explicitly and the problem is highly convection dominated.

In Figures (4.3-4.4), we present the numerical solution and the error between the exact and the numerical solution for state and control at $t = 0.5$. We observe that the problem is approximated well. As expected, the error is more prominent on the regions where the gradient of the solution is higher.

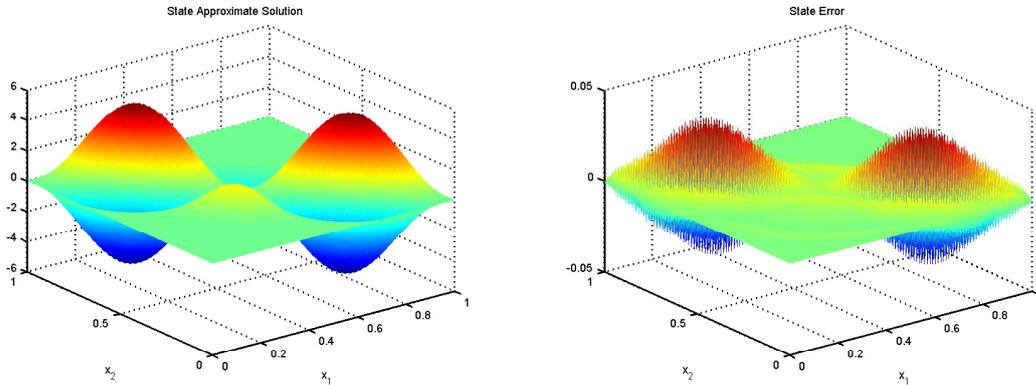


Figure 4.3: DCR Eqn: Numerical solution of the state (*left*), error (*right*) for Example 2 at $t=0.5$ with $h = k = 1/80$ by dG(1) method

In Table 4.2, we present the error for cGP(2) and dG(1) methods. cGP(2) method yields smaller error than dG(1) method because its rate of convergence is cubic in time. The second order convergence rates are achieved with both methods. For dG(1) method, the numerical results indicate a higher order experimental order of convergence than the one shown in Theorem 4.8; it is even higher than quadratic for the control.

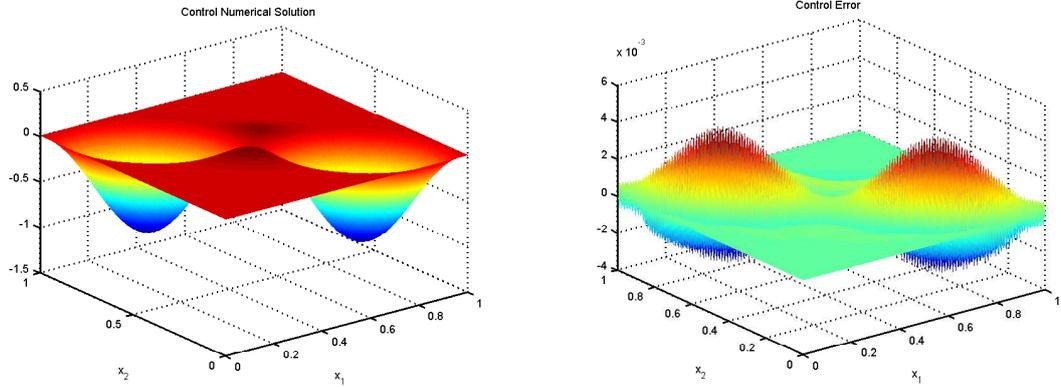


Figure 4.4: DCR Eqn: Numerical solution of the control (*left*), error (*right*) for Example 2 at $t=0.5$ with $h = k = 1/80$ by dG(1) method

Table 4.2: DCR Eqn: Example 2 by cGP(2) and dG(1) (*in parenthesis*) method.

$h = k$	$\ y - y_\delta\ $	Rate	$\ u - u_\delta\ $	Rate
$\frac{1}{5}$	8.08e-1(1.18)	-(-)	7.44e-2(1.04e-1)	- (-)
$\frac{1}{10}$	2.96e-1(4.08e-1)	1.45(1.54)	2.46e-2(3.50e-2)	1.60(1.58)
$\frac{1}{20}$	8.95e-2(1.10e-1)	1.72(1.89)	6.90e-3(1.12e-2)	1.83(1.65)
$\frac{1}{40}$	2.39e-2(2.60e-2)	1.91(2.08)	1.41e-3(2.05e-3)	2.29(2.44)
$\frac{1}{80}$	6.09e-3(6.25e-3)	1.97(2.06)	3.07e-4(3.67e-4)	2.20(2.48)

4.5 A Priori Error Estimates for Optimal Control of Burgers Equation

In this section, we derive error estimates for optimal control of Burgers equation, i.e. $c(y, u) = 0$ denotes the Burgers equation. We follow the approach in [29] and [130]. We write the error in the state and the adjoint as

$$\begin{aligned} e &= y - y_\delta = (\pi y - y_\delta) + (y - \pi y) = \xi + \eta, \\ r &= \lambda - \lambda_\delta = (\pi \lambda - \lambda_\delta) + (\lambda - \pi \lambda) = \varphi + \phi, \end{aligned}$$

where π denotes the space-time projection which is explained in Sec.4.3.

4.5.1 Some Auxiliary Results

We assume that the numerical flux $\hat{n}^s(y_{n+1}^-, y_{n+1}^+)$ is consistent such that $\hat{n}^s(y, y) = n(y)$, conservative and locally Lipschitz continuous [94, Chap. 5].

We make the following assumptions using the study [53, Sec.4.2]:

$$\begin{aligned} &|n_h^s(y, v) - n_h^s(y_\delta, v)| \tag{4.59} \\ &\leq \frac{\epsilon}{8} \|v\|_{DG}^2 + C(\|y_\delta - \pi y\|^2 + h^{2p}|y|_{L^2(I_j; H^{p+1}(\Omega))}^2 + k^{2q+2}|y|_{H^{q+1}(I_j; H^1(\Omega))}^2). \end{aligned}$$

Moreover, we mention some estimates given in [53].

$$\int_{I_m} \|\xi\|^2 dt \leq C k_m \left(\|\xi_{m-1}^-\|^2 + \|\eta_{m-1}^-\|^2 + \int_{I_m} R_m(\eta) dt \right), \tag{4.60a}$$

$$\|\eta_{m-1}^-\|^2 \leq C h^{2p} |y(t_m)|_{H^{p+1}(\Omega)}^2, \tag{4.60b}$$

$$\sum_{j=1}^m \|\eta_j^-\|^2 \leq C \sum_{j=1}^m k_j h^{2p} |y(t_j)|_{H^{p+1}(\Omega)}^2 \leq C T h^{2p} |y|_{C([0, T]; H^{p+1}(\Omega))}^2, \tag{4.60c}$$

$$\int_{I_m} R_m(\eta) dt \leq C \left(\epsilon + \frac{1}{\epsilon} \right) \int_{I_m} \left(h^{2p} |y|_{L^2((I_m); H^{p+1}(I_n))}^2 + k^{2q+2} |y|_{H^{q+1}((I_m), H^1(I_n))}^2 \right) dt. \tag{4.60d}$$

For the first and the second parts of convective term of the adjoint equation, using the definition of the DG-norm and the boundedness of the exact and the discrete solution, we show that there is a constant C independent of h and k such that the following inequalities hold with $v = \frac{1}{\alpha} \phi$:

$$\begin{aligned} &\left| (y \lambda_x, \frac{1}{\alpha} \phi) - (y_\delta (\lambda_\delta)_x, \frac{1}{\alpha} \phi) \right| \\ &\leq \left| ((y - \pi y) \lambda_x, \frac{1}{\alpha} \phi) + ((\pi y) \lambda_x, \frac{1}{\alpha} \phi) - (y_\delta (\lambda_\delta - \pi \lambda)_x, \frac{1}{\alpha} \phi) + (y_\delta (\pi \lambda)_x, \frac{1}{\alpha} \phi) \right| \\ &\leq \frac{\epsilon}{8\alpha} \| \lambda_\delta - \pi \lambda \|_{DG}^2 + C(\|y - \pi y\|^2 + \|\phi\|^2), \end{aligned} \tag{4.61}$$

$$\begin{aligned}
& \left| n_{h,y}^s(y_{n+1}^-, y_{n+1}^+) \llbracket \lambda \rrbracket \frac{1}{\alpha} \phi - n_{h,y}^s(y_{\delta,n+1}^-, y_{\delta,n+1}^+) \llbracket \lambda_\delta \rrbracket \frac{1}{\alpha} \phi \right| \\
& \leq \left| n_{h,y}^s(y_{\delta,n+1}^-, y_{\delta,n+1}^+) \llbracket \lambda_\delta \rrbracket \frac{1}{\alpha} \phi \right| \\
& \leq \left| n_{h,y}^s(y_{\delta,n+1}^-, y_{\delta,n+1}^+) \llbracket \lambda_\delta - \pi \lambda \rrbracket \frac{1}{\alpha} \phi \right| + \left| n_{h,y}^s(y_{\delta,n+1}^-, y_{\delta,n+1}^+) \llbracket \pi \lambda \rrbracket \frac{1}{\alpha} \phi \right| \\
& \leq \frac{\epsilon}{8\alpha} \|\lambda_\delta - \pi \lambda\|_{DG}^2 + C \frac{2}{\epsilon} \|\phi\|^2.
\end{aligned} \tag{4.62}$$

Therefore, we obtain the following estimate

$$\left| n_h^a(\lambda, y, \frac{1}{\alpha} \phi) - n_h^a(\lambda_\delta, y_\delta, \frac{1}{\alpha} \phi) \right| \leq \frac{\epsilon}{4\alpha} \|\lambda_\delta - \pi \lambda\|_{DG}^2 + C(\|y - \pi y\|^2 + (1 + \frac{2}{\epsilon}) \|\phi\|^2). \tag{4.63}$$

4.5.2 Main Result

Lemma 4.9. *Let (y, λ) and $(y_\delta, \lambda_\delta)$ be the solutions of (2.8) and (4.12), respectively. Assume that $y, \lambda \in H^{q+1}(0, T; H^1(\Omega)) \cap L^2(0, T; H^{p+1}(\Omega))$. Then, there is a constant C independent of h and k such that the following inequalities hold:*

$$\|e^n\|^2 + \frac{\epsilon}{2} \int_0^{t_n} \|\xi\|_{DG}^2 dt \leq C \left(\|e_0^-\|^2 - \frac{2}{\alpha} \sum_{m=1}^n \int_{I_m} (\lambda - \lambda_\delta, \xi) dt \right) + \mathcal{O}(h^{2p}, k^{2q+2}), \tag{4.64a}$$

$$\|r^n\|^2 + \frac{\epsilon}{2} \int_0^{t_n} \|\varphi\|_{DG}^2 dt \leq C \left(\frac{2}{\alpha} \sum_{m=1}^n \int_{I_m} (y - y_\delta, \varphi) dt \right) + \mathcal{O}(h^{2p}, k^{2q+2}). \tag{4.64b}$$

Proof. The local error equation for the state equation can be written as follows:

$$\begin{aligned}
& \int_{I_m} ((\partial_t \xi, v) + a_h^s(\xi, v)) dt + ([\xi]_{m-1}, v_+^{m-1}) \\
& = - \left(\int_{I_m} (\partial_t \eta, v) dt + ([\eta]_{m-1}, v_+^{m-1}) \right) - \int_{I_m} a_h^s(\eta, v) dt \\
& + \left(\int_{I_n} n_h^s(y, v) dt - \int_{I_n} n_h^s(y_\delta, v) dt \right) - \frac{1}{\alpha} \int_{I_m} (\lambda - \lambda_\delta, v) dt, \quad \forall v \in V_{h,p}^{k,q}.
\end{aligned} \tag{4.65}$$

We choose $v = 2\xi$ in (4.65). Then, we use coercivity of DG-bilinear form (4.27), the

equality (4.32) and (4.47) and we arrive at

$$\begin{aligned}
& \|\xi_-^m\|^2 - \|\xi_-^{m-1}\|^2 + \epsilon \int_{I_m} \|\xi\|_{DG}^2 dt \\
& \leq \frac{\epsilon}{2} \int_{I_m} \|\xi\|_{DG}^2 dt + \left(\left(1 + \frac{C}{\epsilon} k_m\right) \|\xi_{m-1}^-\|^2 + C \|\eta_{m-1}^-\|^2 + C \int_{I_m} R_m(\eta) dt \right) \\
& + \mathcal{O}(h^{2p}, k^{2q+2}) - \frac{2}{\alpha} \int_{I_m} (\lambda - \lambda_\delta, \xi) dt. \tag{4.66}
\end{aligned}$$

We proceed with the estimates (4.59-4.60) and simplify the common terms. Then, we sum from $m = 1$ to n to derive the following inequality

$$\begin{aligned}
& \|\xi_-^m\|^2 + \frac{\epsilon}{2} \sum_{m=1}^n \int_{I_m} \|\xi\|_{DG}^2 dt \tag{4.67} \\
& \leq C \left(\|\xi_0^-\|^2 + \sum_{m=1}^n \frac{C}{\epsilon} k_m \|\xi_-^{m-1}\|^2 - \frac{2}{\alpha} \sum_{m=1}^n \int_{I_m} (\lambda - \lambda_\delta, \xi) dt \right) + \mathcal{O}(h^{2p}, k^{2q+2}).
\end{aligned}$$

By applying the discrete Gronwall's lemma, we obtain the desired result (4.64a).

We proceed with the local error equation for the adjoint equation.

$$\begin{aligned}
& \int_{I_m} (-\partial_t \varphi, q) + a_h^a(\varphi, q) dt - ([\varphi]_m, q_-^m) \\
& = - \int_{I_m} (\partial_t \phi, q) dt - ([\phi]_m, q_-^m) - \int_{I_m} a_h^a(\varphi, q) dt \\
& - \left(\int_{I_m} n_h^a(\lambda, q) dt - \int_{I_m} n_h^a(\lambda_\delta, q) dt + \int_{I_m} (y - y_\delta, q) dt, \quad \forall q \in V_{h,p}^{k,q}. \tag{4.68}
\end{aligned}$$

We choose $v = \frac{2}{\alpha} \xi$ in (4.68). Similar to the state equation, we use coercivity of DG-bilinear form (4.27), the equality (4.32) and (4.47). Then, we sum from $m = n$ to $m = 1$ to arrive at

$$\begin{aligned}
& \frac{1}{\alpha} \|\varphi^n\|^2 + \frac{\epsilon}{2\alpha} \sum_{m=1}^n \int_{I_m} \|\varphi\|_{DG}^2 dt \\
& \leq C \frac{1}{\alpha} \left(\sum_{m=1}^n \|\varphi^m\|^2 \right) + \mathcal{O}(h^{2p}, k^{2q+2}) + \frac{2}{\alpha} \sum_{m=1}^n \int_{I_m} (y - y_\delta, \xi) dt. \tag{4.69}
\end{aligned}$$

By applying the discrete Gronwall's lemma and using the properties of the space-time projection in Sec.4.3.1, we obtain the desired result (4.64b).

□

Now, we proceed with the estimates at arbitrary time steps using the previous estimate and the discrete characteristic function (4.3.2).

Lemma 4.10. *Let $(y_\delta, \lambda_\delta)$ be the solution of (4.12). Assume that $y, \lambda \in H^{q+1}(0, T; H^1(\Omega)) \cap L^2(0, T; H^{p+1}(\Omega))$. Then, there is a constant C independent of h and k such that the following inequalities hold:*

$$\sup_{t \in I_m} \|\xi(t)\|^2 + \sup_{t \in I_m} \|\varphi(t)\|^2 \leq C \|e_0^-\|^2 + \mathcal{O}(h^{2p}, k^{2q+2}). \quad (4.70)$$

Proof. The proof is similar to the case in Lemma 4.9. To do this, for any fixed $t \in [t^{n-1}, t^n)$, we set $v_\delta = \tilde{e}_h$ in (4.65).

$$\begin{aligned} & \sup_{t \in I_m} \|\xi(t)\|^2 - \|\xi_-^{m-1}\|^2 \\ & \leq \frac{\epsilon}{2} \int_{I_m} \|\tilde{\xi}\|_{DG}^2 dt + \frac{C}{\epsilon} \int_{I_m} \|y_\delta - \pi y\|^2 dt - \frac{2}{\alpha} \int_{I_m} (\lambda - \lambda_\delta, \tilde{\xi}) dt + \mathcal{O}(h^{2p}, k^{2q+2}). \end{aligned} \quad (4.71)$$

Here, we use the properties of the discrete characteristic functions which has been explained in Sec.4.3.2. We choose $t \in (t^{n-1}, t^n]$ such that

$$\|e(t)\|^2 = \sup_{t^{n-1} \leq s \leq t^n} \|e_h\|^2.$$

We note that the term

$$\|e_h^{n-1}\|^2 \leq \sup_{t^{n-2} \leq s \leq t^{n-1}} \|e(s)\|^2.$$

Then, we have

$$\int_{I_n} \|e_h\|^2 dt \leq k_n \sup_{t^{n-1} \leq s \leq t^n} \|e_h\|^2 \equiv k_n \|e(t)\|^2.$$

For the adjoint equation, for any fixed $t \in (t^{N-1}, t^N]$, we set $v_\delta = \frac{1}{\alpha} \tilde{r}_h$ in (4.68) and proceed similarly.

$$\begin{aligned} & \sup_{t \in I_m} \|\varphi(t)\|^2 - \|\varphi^m\|^2 \\ & \leq \frac{\epsilon}{2} \int_{I_m} \|\tilde{\varphi}\|_{DG}^2 dt + \frac{C}{\epsilon} \int_{I_m} \|y - \pi y\|^2 dt + \frac{2}{\alpha} \int_{I_m} (y - y_\delta, \tilde{\xi}) dt + \mathcal{O}(h^{2p}, k^{2q+2}). \end{aligned} \quad (4.72)$$

In order to eliminate the terms $p - p_\delta$ and $y - y_\delta$ in the estimates 4.71 and 4.72, we follow the approach in [29] and write the coupling terms as follows:

$$\begin{aligned} & -\frac{1}{\alpha} \int_{I_m} (r, \tilde{\xi}) dt + \frac{1}{\alpha} \int_{I_m} (e, \tilde{\varphi}) dt \\ & \leq C \int_{I_m} \left(\frac{2}{\epsilon \alpha^2} \|\phi\|^2 + \frac{2}{\epsilon \alpha} \|\eta\|^2 \right) dt + \int_{I_m} \left(\frac{\epsilon}{8} \|\tilde{\xi}\|_{DG}^2 + \frac{\epsilon}{8\alpha} \|\tilde{\varphi}\|_{DG}^2 \right) dt, \\ & \leq \mathcal{O}(h^{2p+2}, k^{2q+2}) + C_D \int_{I_m} \left(\frac{\epsilon}{8} \|\xi\|_{DG}^2 + \frac{\epsilon}{8\alpha} \|\varphi\|_{DG}^2 \right) dt. \end{aligned} \quad (4.73)$$

We add the resulting inequalities for the state and the adjoint equation and sum from $m = 1$ to n . Then, we apply discrete Gronwall's inequality to arrive at the estimate (4.70).

□

Using the estimates in Lemma 4.9-4.10, we state the main estimate for optimal control of Burgers equation.

Theorem 4.11. *Let (y, λ, u) and $(y_\delta, \lambda_\delta, u_\delta)$ be the solutions of (2.8) and (4.12), respectively. Assume that all the conditions in Lemma 4.9-4.10 hold. Then, there exists a constant C independent of h and k such that*

$$\begin{aligned} & \|y - y_\delta\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda - \lambda_\delta\|_{L^\infty(0,T;L^2(\Omega))} + \|u - u_\delta\|_{L^\infty(0,T;L^2(\Omega))} \\ & \leq C (h^{p+1}\|y_0\| + \mathcal{O}(h^p, k^{q+1})). \end{aligned} \quad (4.74)$$

4.6 Computational Aspects

In Section 4.4, the derivation of the linear systems associated to the state and the adjoint equation are explained. Now, we derive the corresponding systems optimal control of Burgers equation.

dG(1)-method

For the dG(1)-method, the state y_δ and the control u_δ are approximated in the time interval $I_m = (t_{m-1}, t_m]$ by

$$\begin{aligned} y_\delta &= Y_m^1 \phi_{m,1}^s(t) + Y_m^2 \phi_{m,2}^s(t), \quad \forall t \in I_m, \\ u_\delta &= U_m^0 \phi_{m,0}^a(t) + U_m^1 \phi_{m,1}^a(t), \quad \forall t \in I_m, \end{aligned}$$

using the Gauss-Radau quadrature rule, where $\phi_{m,j}^s \in \mathbb{P}_1(I_m)$ are the linear Lagrange basis functions on \bar{I}_m .

The discrete time steps on \bar{I}_m are chosen according to the right-handed 2-point Gauss-Radau rule on I_m

$$t_{m,1}^s = t_{m-1} + \frac{k_m}{3}, \quad t_{m,2}^s = t_m,$$

we use the reference weights

$$\begin{aligned} \hat{w}_1^s &= 3/2, & \hat{w}_2^s &= 1/2, \\ \hat{\psi}_1^s &= \frac{1 - \hat{t}}{2}, & \hat{\psi}_2^s &= \frac{3\hat{t} + 1}{2}. \end{aligned}$$

The initial conditions for state equation on each time interval I_m are

$$Y_m^0 = (y_\delta)_{m-1}^- \quad \text{if } m \geq 2 \quad \text{or} \quad (y_\delta)_{m-1}^- = y_0^- \quad \text{if } m = 0.$$

For the state equation, on each time interval I_m , we have the following systems of equations

$$\begin{aligned} \left(\frac{3}{4}\mathbf{M} + \frac{k}{2}\mathbf{A} \right) Y_m^1 + \frac{k}{2}\mathbf{N}^s((Y_m^1)^2) + \frac{1}{4}\mathbf{M}Y_m^2 \\ = \mathbf{M}Y_m^0 + \frac{k}{2}\mathbf{F}_h(t_{m,1}) + \frac{k}{4}\mathbf{M}(U_m^0 + U_m^1), \end{aligned} \quad (4.75a)$$

$$\begin{aligned} -\frac{9}{4}\mathbf{M}Y_m^1 + \left(\frac{5}{4}\mathbf{M} + \frac{k}{2}\mathbf{M} \right) Y_m^2 + \frac{k}{2}\mathbf{M}^s((Y_m^2)^2) \\ = -\mathbf{M}Y_m^0 + \frac{k}{2}\mathbf{M}_h(t_{m,2}) + \frac{k}{4}\mathbf{M}(3U_m^1 - U_m^0). \end{aligned} \quad (4.75b)$$

This system can be linearised using Newton's method for $Y_m^{1,\alpha}, Y_m^{2,\alpha}$ and solved for $\delta Y_m^1, \delta Y_m^2$ on each time interval $I_m = (t_{m-1}, t_m]$.

$$\begin{aligned} \left(\begin{array}{cc} \frac{3}{4}\mathbf{M} + \frac{k_m}{2}(\mathbf{A}^s + k\mathbf{N}_y^s(Y_m^{1,\alpha})) & \frac{1}{4}\mathbf{M} \\ -\frac{9}{4}\mathbf{M} & \frac{5}{4}\mathbf{M} + \frac{k_m}{2}(\mathbf{A}^s + k\mathbf{N}_y^s(Y_m^{2,\alpha})) \end{array} \right) \begin{pmatrix} \delta Y_m^1 \\ \delta Y_m^2 \end{pmatrix} \\ = \begin{pmatrix} \left(\frac{3}{4}\mathbf{M} + \frac{k}{2}\mathbf{A} \right) Y_m^1 + \frac{k}{2}\mathbf{N}^s((Y_m^1)^2) + \frac{1}{4}\mathbf{M}Y_m^2 - \mathbf{M}Y_m^0 - \frac{k}{2}\mathbf{F}_h(t_{m,1}) - \frac{k}{4}\mathbf{M}(U_m^0 + U_m^1) \\ -\frac{9}{4}\mathbf{M}Y_m^1 + \left(\frac{5}{4}\mathbf{M} + \frac{k}{2}\mathbf{A} \right) Y_m^2 + \frac{k}{2}\mathbf{N}^s((Y_m^2)^2) + \mathbf{M}Y_m^0 - \frac{k}{2}\mathbf{F}_h(t_{m,2}) - \frac{k}{4}\mathbf{M}(3U_m^1 - U_m^0) \end{pmatrix}. \end{aligned} \quad (4.76)$$

The solution is updated as

$$Y_m^{1,\alpha+1} = Y_m^{1,\alpha} - \delta Y_m^1$$

and

$$Y_m^{2,\alpha+1} = Y_m^{2,\alpha} - \delta Y_m^2$$

up to a given tolerance. Then, we denote $Y_m^1 := Y_m^{1,\alpha+1}$ and $Y_m^2 := Y_m^{2,\alpha+1}$.

Let λ_δ be the approximate solution of the adjoint

$$\lambda_\delta = \lambda_m^0 \phi_{m,0}^a(t) + \lambda_m^1 \phi_{m,1}^a(t), \quad \forall t \in I_m,$$

using linear orthogonal Lagrange functions and Gauss-Radau points.

The discrete time steps on \bar{I}_m are chosen according to the left-handed 2-point Gauss-Radau rule

$$t_{m,0}^a = t_{m-1}, \quad t_{m,1}^a = t_{m-1} + \frac{k_m}{3},$$

we use the corresponding reference weights and test functions

$$\begin{aligned} \hat{w}_0^a = 1/2, \quad \hat{w}_1^a = 3/2, \\ \hat{\psi}_0^a = \frac{1-3\hat{t}}{2}, \quad \hat{\psi}_1^a = \frac{1+\hat{t}}{2}. \end{aligned}$$

The initial conditions for the adjoint equation on each time interval I_m are

$$\lambda_m^2 = (\lambda_\delta)_m^+ \quad \text{if } m \leq N_T \quad \text{or} \quad \lambda_m^2 = 0 \quad \text{if } m = N_T.$$

On each time interval $I_m = (t_{m-1}, t_m]$, we solve the following linear system for λ_m^1, λ_m^0

$$\begin{pmatrix} \frac{3}{4}\mathbf{M} + \frac{k_m}{2}(\mathbf{A}^a - \mathbf{N}^a(\frac{Y_m^1 + Y_m^2}{2})) & \frac{1}{4}\mathbf{M} \\ -\frac{9}{4}\mathbf{M} & \frac{5}{4}\mathbf{M} + \frac{k_m}{2}(\mathbf{A}^a - \mathbf{N}^a(3\frac{Y_m^1 - Y_m^2}{2})) \end{pmatrix} \begin{pmatrix} \lambda_m^1 \\ \lambda_m^0 \end{pmatrix} \\ = \begin{pmatrix} \mathbf{M}\lambda_m^2 + \frac{k_m}{2}\mathbf{Y}_h^d(t_{m,1}) - \frac{k_m}{4}\mathbf{M}(Y_m^1 + Y_m^2) \\ -\mathbf{M}\lambda_m^2 + \frac{k_m}{2}\mathbf{Y}_h^d(t_{m,0}) - \frac{k_m}{4}\mathbf{M}(3Y_m^1 - Y_m^2) \end{pmatrix}. \quad (4.77)$$

Similar to the cGP(q+1) method, the discrete adjoint at t_{m-1} is given as $\lambda_{h,m-1} = \lambda_m^0$.

4.6.1 Numerical Results

In this section, we present some numerical results. We measure the error in the approximate cost functional at two successive space-time meshes. We have used discontinuous piecewise linear polynomials in space and in time. In all numerical examples, we have taken $h = \mathcal{O}(k)$.

We consider the optimal control problem in [82] with

$$Q = (0, 1] \times \Omega, \quad \Omega = (0, 1), \quad \epsilon = 10^{-2}, \quad \text{and } \alpha = 0.05.$$

We take the source function $f = 0$, the desired state y_d and the initial condition y_0 are defined as

$$y_d(x, t) = \begin{cases} 1 & \text{in } (0, 1/2], \\ 0 & \text{otherwise.} \end{cases}$$

In Figure 4.5, we presented the numerical results for the state, the adjoint and the control. It is observed that the solutions are similar the ones in [65]. Finding a suitable control, the discontinuity at $x = 1/2$ is smoothed and the state is pushed left as time passes so that the difference between the state and the desired state becomes small. The relation (2.8c) between the adjoint and the control is realised in the figures.

In Table 4.3, we give the errors for dG(1) method. We observe that numerical results indicate better convergence rates, that is around 1.6, than the theoretical one which is $\mathcal{O}(h)$ with $h = k$.

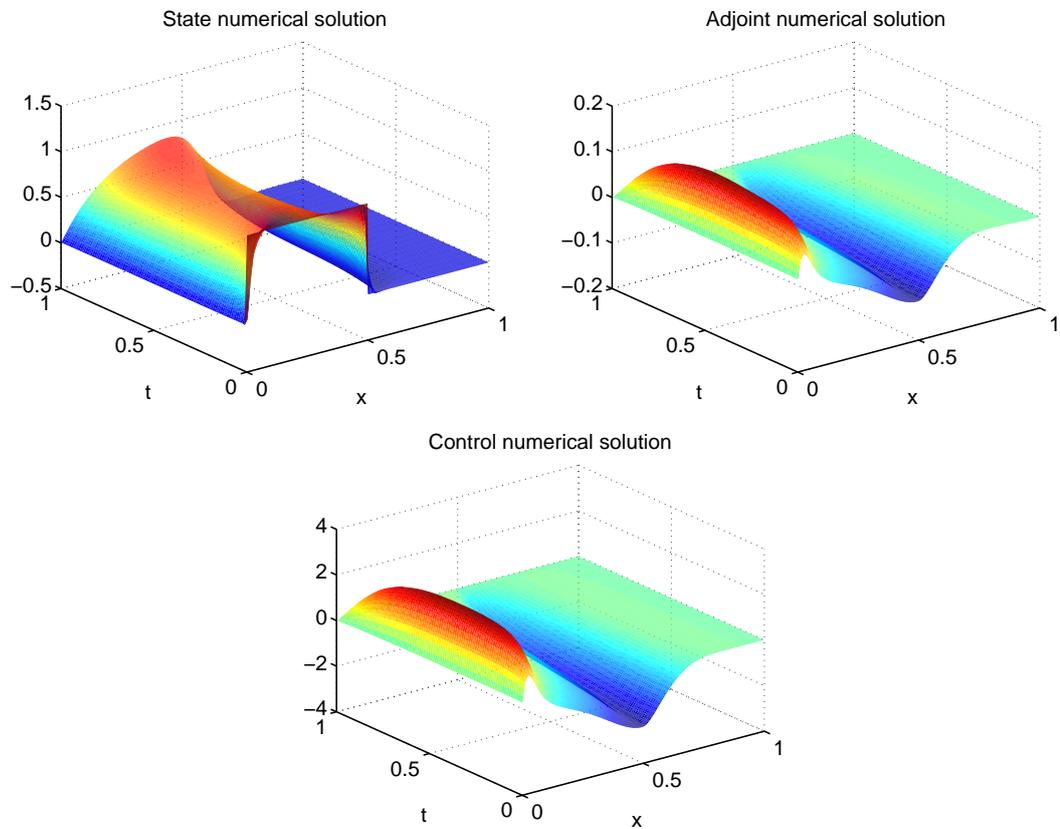


Figure 4.5: Burgers Eqn: Numerical solution of the state (*top-left*), the adjoint (*top-right*) and the control (*bottom*) for Example 1 with $h = k = 1/128$ by dG(1) method

Table 4.3: Burgers Eqn: Example 1 by dG(1) method.

$h = k$	$J(y_\delta, u_\delta)$	$J(y_{2\delta}, u_{2\delta}) - J(y_\delta, u_\delta)$	Rate
$\frac{1}{8}$	0.1781574	-	-
$\frac{1}{16}$	0.1678330	1.0324e-2	-
$\frac{1}{32}$	0.1611209	6.7121e-3	0.62
$\frac{1}{64}$	0.1584206	2.7003e-3	1.31
$\frac{1}{128}$	0.1574704	9.5020e-4	1.51
$\frac{1}{256}$	0.1571551	3.1530e-4	1.59

CHAPTER 5

MODEL REDUCTION USING PROPER ORTHOGONAL DECOMPOSITION METHOD

Spatial discretization techniques, for example finite difference, discontinuous Galerkin, finite volume or spectral method, are used to convert the continuous problem to a linear/nonlinear ordinary differential equation. Then, temporal discretization follows. Resulting fully-discrete problem is called as full-order model (FOM), which is discussed in Sec. 4. In order to achieve an accurate and stable numerical solution, one should use a fine grid/mesh, which increases the dimension of the discrete problem, so the computational time. In case of perturbations in the data, FOM is required to be solved for each new parameter in the data set or in case of optimization problems, FOM associated to the differential equations in the Algorithm 4.4 must be resolved after updating the control. Therefore, a method, that eliminates the necessity of the solution of the FOM for each parameter or that enables us to solve the problem in a fast way, is required. Here, model-order reduction (MOR) techniques are used to replace the FOM with the reduced-order model (ROM).

In the literature, there are several MOR techniques and each of them has different advantages and disadvantages depending on the problem at hand. For an overview of the methods, we refer the reader to the study [19]. For example, balanced truncation method is employed for linear time invariant systems with many input or many output [11], because the error between the input-output map associated to full-order and reduced-order systems can be estimated in case of a homogeneous initial condition. On the other hand, in order to apply the balanced truncation method, the system must be balanced. In other words, two Lyapunov equations corresponding to the controllability and observability Gramian of the system must be solved [61]. Unfortunately, this method is not efficient for large-scale problems since dense matrix factorizations must be obtained for exact balanced truncation; so, approximate balanced truncation methods are preferred [60]. For inhomogeneous initial condition, the study [66] is suggested. This method is advantageous because ROM is kept to be asymptotically stable. Although there are some promising results [18], there are still open issues in order to apply this method to time invariant or nonlinear systems.

One of the most popular MOR techniques is Krylov subspace method or moment matching approximation [60]. This time, FOM is projected on to Krylov subspace where Arnoldi or Lanczos process are used to derive the Krylov subspace. Efficiency of this method lies on the fact that the matrix-vector multiplication is sufficient

throughout the procedure [55].

In addition, reduced-basis method is a rapid and reliable technique based on Galerkin projection and it is especially preferable for parametrized PDEs [107]. To apply this method, FOM is projected onto a low-ordered subspace which is constructed using precomputed finite element solutions. Offline-online decomposition of the model is utilized where parameter independent but expensive parts of the model are computed offline, while parameter-dependent and inexpensive parts are taken into account in the online step in order to accelerate the solution procedure [89]. Basis selection is performed using a posteriori error estimation. After deriving the low-dimensional subspace, an approximate solution of the problem depending on another set of parameters, which can model geometrical or physical properties of the system, is derived. In addition, one can check the accuracy of the reduced-order solution using the mentioned a posteriori error estimator [103, 109, 89]. This method has been applied to optimal control of elliptic problems in [103] and parabolic problems in [36, 37, 79]. Moreover, reduced-basis methods on space-time domain have started to gain interest, too [138].

MOR techniques can be formulated as a goal-oriented scheme where the aim is to choose the reduced basis so that the difference between FOM and the output associated to ROM is minimized over a set of input and a time interval [22]. Quality of this approach, of course, depends on the information contained in the snapshot ensemble and the method can be applied to optimal control, optimal design and inverse problems. On the other hand, domain decomposition methods can be employed as a MOR technique [15] where this idea has been applied to PDE constrained optimization problems in [9, 10].

Apart from these methods, one of the most popular and successful MOR techniques is the proper orthogonal decomposition (POD) method, also known as Karhunen-Loève decomposition or principal component analysis [71, 83, 133]. The idea behind the POD method is to derive a new basis spanning the space whose dimension is lower than the finite element space. Then, FOM is projected onto the low-dimensional subspace using the new optimal POD basis as we proceed in Galerkin projection. POD basis is computed using the full-order solution at discrete time steps, which are called as the snapshots. Because the snapshots might be linearly dependent or almost linearly dependent, they cannot be directly used as a basis. Therefore, we construct a snapshot matrix and employ singular value decomposition or eigenvalue decomposition to derive the POD basis, which is explained in Sec. 5.2. Quality of the reduced-order solution depends on how much information is contained in the snapshot matrix [84] or the location of each snapshot [85]. Order of the POD basis must be sufficiently large to approximate the problem accurately. On the other hand, it shouldn't be too large, for example not larger than the full-order space, in order not to lose the efficiency.

In the literature, there are several studies concerning the reduced solution of the PDE-constrained optimization based on POD. However, studies on the optimal control of diffusion-convection-reaction equation, compared to parabolic PDEs, or Burgers equation are not so popular cases. For example, adaptive time step method is applied to optimal control of one-dimensional diffusion-convection equation in [6]. Another adaptive scheme is employed for reacting flows in order to derive the local information

and to reduce the cost of the solution process in [116]. A comparison of two different space discretization methods are given in [95] to compute the POD basis functions for optimal control of diffusion-convection-reaction processes. For optimal control of Burgers equation, we refer the reader to the study [82] where open-loop and closed-loop OCPs are solved using POD. In [86, 87], optimal boundary feedback control of Burgers equation is discussed by utilizing the Hamilton-Jacobi-Bellman equation for infinite horizon OCPs. In the presence of control constraints, a posteriori error analysis is utilized in order to measure the difference between the suboptimal control computed from the POD basis and the optimal control; then, the number of POD basis is decided [124]. The state-constrained case is discussed in [59].

In this study, we use POD method as a MOR technique where the fully-discrete snapshots are derived by space-time dG method. To the best of our knowledge, this is the first study where a priori error analysis based on POD for space-time dG discretization is presented.

In this chapter, firstly, continuous POD method is presented, namely the trajectories are continuous in time and space (see, Sec. 5.1). Then, in Sec. 5.2, discrete version of the POD is explained for space-time dG method. In Sec. 5.2.4, optimality system based on POD and the associated linear systems are derived. Then, numerical results are given in Sec. 5.3.3. A priori error estimates based on POD follows in Sec. 5.3 for OCP of DCR eqn. In Sec. 5.4, discrete empirical interpolation method, which is used to increase the efficiency of POD in case of nonlinear state equation, is explained. Error estimates for OCP of Burgers equation is proven in Sec. 5.5. Then, theoretical analysis is compared with the numerical results in Sec. 5.5.3.

5.1 Continuous POD Method

In this section, we briefly explain the POD method following [133] and its generalization to Galerkin type spatial discretization [51, Sec. 3]. The associated error bound and the semi-discrete optimality system based on POD are presented additionally.

Let $y = y(u)$ denote the weak solution of the state associated the control u . We assume the data is given as

$$\mathcal{V} = \{y(t) : t \in [0, T]\} \subseteq V,$$

with $d := \dim \mathcal{V} \leq \infty$.

Definition 5.1. A POD basis $\{\psi_j\}_{j=1}^l$ of rank l is computed by solving the following optimization problem [82]

$$\min_{\{\psi_j\}_{j=1}^l} \int_0^T \left\| y(t) - \sum_{j=1}^l (y(t), \psi_j) \psi_j \right\|^2 dt \quad \text{s.t.} \quad (\psi_j, \psi_k) = \delta_{jk} \text{ for } 1 \leq j, k \leq l, \quad (5.1)$$

where (ϕ, ψ) denotes the $L^2(\Omega)$ inner product; δ_{jk} is the Kronecker symbol, that is, $\delta_{jk} = 0$ for $i \neq j$ and $\delta_{jj} = 1$.

The minimization problem in (5.1) can be solved by constructing the Lagrangian and deriving the first and the second order optimality conditions. Then, we define the operator $\mathcal{R} := V \rightarrow \mathcal{V}$ such that

$$\mathcal{R}\psi = \int_0^T (y(t), \psi)y(t) dt \quad \text{for } \psi \in V. \quad (5.2)$$

The operator \mathcal{R} is self-adjoint, linear, bounded and nonnegative [133]. The solution of the problem (5.1) is the eigenfunctions $\{\psi_i\}_{i=1}^l$ associated to the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l \geq 0$ such that

$$\mathcal{R}\psi_i = \int_0^T (y(t), \psi_i)y(t) dt = \lambda_i\psi_i, \quad (\psi_i, \psi_j) = \delta_{ij}, \quad i, j = 1, \dots, l. \quad (5.3)$$

We define the operator $\mathcal{K} := L^2(0, T) \rightarrow L^2(0, T)$ such that

$$\mathcal{K}\phi = \int_0^T (y(t), y(\cdot))\phi(t) dt \quad \text{for } \phi \in L^2(0, T). \quad (5.4)$$

We note that the equation (5.4) is equivalent to (5.2) and they have the same eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$. Then, an alternative formula to (5.3), which is called as method of snapshots, can be stated as follows:

$$\mathcal{K}\phi_i = \int_0^T (y(t), y(\cdot))\phi_i(t) dt = \lambda_i\phi_i, \quad (\psi_i, \psi_j) = \delta_{ij}, \quad i, j = 1, \dots, l. \quad (5.5)$$

The following approximation error and a detailed proof can be found in [133].

Proposition 5.1. *Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d > 0$ denote the positive eigenvalues of \mathcal{K} with $\lambda_i \neq 0$. Then, the POD basis of rank $l \leq d$ is given by*

$$\psi_j = \frac{1}{\sqrt{\lambda_j}} \int_0^T \phi_j(t)y(t) dt \quad \text{for } j = 1, \dots, l. \quad (5.6)$$

The POD approximation error is given by

$$\int_0^T \left\| y(t) - \sum_{j=1}^l (y(t), \psi_j)\psi_j \right\|^2 dt = \sum_{j=l+1}^d \lambda_j. \quad (5.7)$$

After computing the orthonormal POD basis of rank $l < d$, namely $\{\psi_1, \dots, \psi_l\}$, we can write the state, the adjoint and the control trajectories for $t \in [0, T]$ as

$$y(t) = \sum_{i=1}^d (y(t), \psi_i)\psi_i, \quad \lambda(t) = \sum_{i=1}^d (\lambda(t), \psi_i)\psi_i, \quad u(t) = \sum_{i=1}^d (u(t), \psi_i)\psi_i. \quad (5.8)$$

Then, an approximation to (5.8) with $l < d$ for $t \in [0, T]$ is written as

$$y^l(t) = \sum_{i=1}^l \underbrace{(y(t), \psi_i)}_{:=y_i^l(t)} \psi_i, \quad \lambda^l(t) = \sum_{i=1}^l \underbrace{(\lambda(t), \psi_i)}_{:=\lambda_i^l(t)} \psi_i, \quad u^l(t) = \sum_{i=1}^l \underbrace{(u(t), \psi_i)}_{:=u_i^l(t)} \psi_i. \quad (5.9)$$

After deriving the POD basis, low-dimensional optimality system corresponding to (1.1) or (1.2) is derived by applying Galerkin projection. In addition, the initial condition y_0 , the desired state y_d and the source function f are also projected onto the low-dimensional subspace. Then, POD based optimality system associated to (2.6) or (2.8) is written as

$$(\partial_t y^l, \psi_j) + a(y^l, \psi_j) + n(y^l, \psi_j) = (f^l + u^l, \psi_j), \quad \forall \psi_j \in \mathcal{V}^l, \quad t \in (0, T], \quad (5.10a)$$

$$-(\partial_t \lambda^l, \psi_j) + a^a(\lambda^l, \psi_j) - n^\lambda(\lambda^l, \psi_j) = -(y^l - y^{d,l}, \psi_j), \quad \forall \psi_j \in \mathcal{V}^l, \quad t \in (0, T], \quad (5.10b)$$

$$\alpha u^l = \lambda^l, \quad t \in (0, T]. \quad (5.10c)$$

We note that for optimal control of DCR equation (1.1), the terms $n(y^l, \psi_j)$ and $n^\lambda(\lambda^l, \psi_j)$ are zero.

5.2 Discrete POD Method

In this section, we explain the discrete POD starting with the spatial discretization. Then, the temporal discretization based on space-time dG method follows.

5.2.1 Spatial Discretization of POD Method

We suppose that the semi-discrete trajectories, namely $y_h(t)$, are at hand. Then, using the trajectories, POD basis coefficients are derived as follows:

Definition 5.2. POD basis coefficients $\{\Psi_{:,j}\}_{j=1}^l$ associated to the POD basis of rank l are computed by solving the following optimization problem [82]

$$\begin{aligned} \min_{\{\Psi_{:,j}\}_{j=1}^l} \int_0^T \left\| y_h(t) - \sum_{j=1}^l (y_h(t), \Psi_{:,j})_{\mathbf{M}} \Psi_{:,j} \right\|_{\mathbf{M}}^2 dt, \quad (5.11) \\ \text{s.t.} \quad (\Psi_{:,j}, \Psi_{:,k})_{\mathbf{M}} = \delta_{jk} \text{ for } 1 \leq j, k \leq l. \end{aligned}$$

where $(\Phi, \Psi)_{\mathbf{M}} = \Phi^T \mathbf{M} \Psi$ with the finite element mass matrix \mathbf{M} and δ_{jk} is the Kronecker symbol, that is, $\delta_{jk} = 0$ for $i \neq j$ and $\delta_{jj} = 1$.

After deriving the coefficients Ψ , l -many POD basis functions are expressed as a linear combination of the discontinuous finite element basis functions $\varphi_i(x)$ as follows

$$\psi_j(x) = \sum_{i=1}^{\text{dof}} \Psi_{ij} \varphi_i(x), \quad j = 1, \dots, l. \quad (5.12)$$

The solution of the problem (5.11) is the eigenvectors $\{\Psi_{:,i}\}_{i=1}^l$ associated to the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l \geq 0$ such that

$$\mathcal{R}_h \Psi_{:,i} = \int_0^T (y_h(t), \Psi_{:,i}) y_h(t)_{\mathbf{M}} dt = \lambda_i \Psi_{:,i}, \quad \text{with } (\Psi_{:,i}, \Psi_{:,j})_{\mathbf{M}} = \delta_{ij}, \quad i, j = 1, \dots, l. \quad (5.13)$$

Proposition 5.2. *The POD approximation error associated to the POD basis of rank $l \leq d$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d > 0$ is given by*

$$\int_0^T \left\| y_h(t) - \sum_{j=1}^l (y_h(t), \Psi_{:,j})_{\mathbf{M}} \Psi_{:,j} \right\|_{\mathbf{M}}^2 dt = \sum_{j=l+1}^d \lambda_j. \quad (5.14)$$

An approximation to (5.8) with $l < d$ for $t \in [0, T]$ is written as

$$y_h^l(t) = \sum_{i=1}^l \underbrace{(y_h(t), \psi_i)}_{:= \mathbf{y}_{h,i}^l(t)} \psi_i, \quad \lambda_h^l(t) = \sum_{i=1}^l \underbrace{(\lambda_h(t), \psi_i)}_{:= \lambda_{h,i}^l(t)} \psi_i, \quad u_h^l(t) = \sum_{i=1}^l \underbrace{(u_h(t), \psi_i)}_{:= \mathbf{u}_{h,i}^l(t)} \psi_i. \quad (5.15)$$

Remark 5.1. After computing the POD basis coefficients, i.e., the matrix $\Psi \in \mathbb{R}^{\text{dof} \times l}$, the POD approximation of the state, the adjoint and the control in (5.9) can be written as [88]

$$y_h^l(t) = \sum_{i=1}^l \mathbf{y}_{h,i}^l(t) \psi_i = \sum_{i=1}^l \mathbf{y}_{h,i}^l(t) \sum_{j=1}^{\text{dof}} \Psi_{ji} \varphi_j(x) = \sum_{j=1}^{\text{dof}} (\Psi \mathbf{y}_h^l(t))_j \varphi_j(x),$$

$$\lambda_h^l(t) = \sum_{j=1}^{\text{dof}} (\Psi \lambda_h^l(t))_j \varphi_j(x), \quad u_h^l(t) = \sum_{j=1}^{\text{dof}} (\Psi \mathbf{u}_h^l(t))_j \varphi_j(x).$$

5.2.2 Semi-discrete Optimality System Based on POD

The semi-discrete optimality system associated to (5.10) is written in matrix vector-form as

$$\mathbf{M}^l \frac{d\mathbf{y}_h^l(t)}{dt} + \mathbf{A}^l \mathbf{y}_h^l + \mathbf{N}^l(\mathbf{y}_h^l(t)) = \mathbf{F}_h^l(t) + \mathbf{M}^l \mathbf{u}_h^l(t), \quad (5.16a)$$

$$\mathbf{M}^l \mathbf{y}_h^l(0) = \mathbf{y}_{l,0}, \quad (5.16b)$$

$$-\mathbf{M}^l \frac{d\boldsymbol{\lambda}_h^l(t)}{dt} + (\mathbf{A}^l)^T \boldsymbol{\lambda}_h^l - \mathbf{N}^{\lambda,l}(\boldsymbol{\lambda}_h^l(t)) = -(\mathbf{M}^l \mathbf{y}_h^l(t) - \mathbf{Y}_h^{d,l}(t)), \quad (5.16c)$$

$$\mathbf{M}^l \boldsymbol{\lambda}_h^l(T) = 0, \quad (5.16d)$$

$$\int_0^T (\alpha \mathbf{M}^l \mathbf{u}_h^l(t) - \mathbf{M}^l \boldsymbol{\lambda}_h^l(t), w_h) = 0. \quad (5.16e)$$

where

$$\begin{aligned} \mathbf{M}_{ij}^l &= (\psi_j, \psi_i)_\Omega, \quad \mathbf{A}_{ij}^l = a_h(\psi_j, \psi_i)_\Omega, \quad \mathbf{F}_i^l = (f_h(x, t), \psi_i), \\ \mathbf{N}_i^l &= n_h(y^l(x, t), \psi_i), \quad \mathbf{N}_i^{\lambda,l} = n_h^\lambda(\lambda^l(x, t), \psi_i), \\ (\mathbf{Y}^{d,l}(t))_i &= (y_h^d(x, t), \psi_i), \quad (\mathbf{y}_{l,0})_i = (y_{h,0}(x, t), \psi_i), \\ &\text{for } 1 \leq i, j \leq l, \quad t \in (0, T]. \end{aligned} \quad (5.17)$$

Then, the matrices in (5.17) can be computed as follows

$$\begin{aligned} \mathbf{M}^l &= \Psi^T \mathbf{M} \Psi, \quad \mathbf{A}^l = \Psi^T \mathbf{A} \Psi, \quad \mathbf{F}^l = \Psi^T \mathbf{F}, \quad \mathbf{Y}^{d,l} = \Psi^T \mathbf{Y}^d, \\ \mathbf{N}^l &= \Psi^T \mathbf{N}(\Psi^T \mathbf{y}^l), \quad \mathbf{N}^{\lambda,l} = \Psi^T \mathbf{N}(\Psi^T \boldsymbol{\lambda}^l). \end{aligned} \quad (5.18)$$

We note that the terms $\mathbf{N}^l(\mathbf{y}_h^l(t))$ and $\mathbf{N}^{\lambda,l}(\boldsymbol{\lambda}_h^l(t))$ in (5.16) are zero for DCR equation.

5.2.3 Temporal Discretization POD Method

In order to solve the optimization problem (5.1) numerically, we need the discrete solution corresponding to $y(t)$ at discrete time steps. We suppose that the snapshots $y_{\delta,i} := y_\delta|_{I_i}$ on each time interval I_i , arising from space-time dG discretization are available. We consider the following snapshot ensemble $Y_l = \{y_{\delta,1}, \dots, y_{\delta,N}\}$. We assume that at least one of the snapshots in Y_l is non-zero. For an orthonormal basis of Y_l , namely $\{\psi_j\}_{j=1}^d$, with $d = \dim Y_l$; each member in the snapshot ensemble Y_l can be written as

$$y_{\delta,i} = \underbrace{\sum_{j=1}^d (y_{\delta,i}, \psi_j)_{\mathbf{M}} \psi_j}_{:= \mathbf{y}_{\delta,j}} = \sum_{j=1}^d \mathbf{y}_{\delta,i} \psi_j \quad \text{for } i = 1, \dots, N.$$

On the other hand, POD approximation for the state, the adjoint and the control for $t \in I_i$ are written as

$$y_{\delta,i}^l = \sum_{j=1}^l \mathbf{y}_{\delta,i}^l \psi_j, \quad \lambda_{\delta,i}^l = \sum_{j=1}^l \boldsymbol{\lambda}_{\delta,i}^l \psi_j, \quad u_{\delta,i}^l = \sum_{j=1}^l \mathbf{u}_{\delta,i}^l \psi_j, \quad \text{for } i = 1, \dots, N.$$

where $l \ll d$.

Definition 5.3. POD basis coefficients associated to the POD basis of rank l , which represents the snapshot ensemble as well as possible, is the solution of the following optimization problem [82]

$$\min_{\{\Psi_{:,j}\}_{j=1}^l} \sum_{i=1}^N \int_{I_i} \left\| y_{\delta,i} - \sum_{j=1}^l (y_{\delta,i}, \Psi_{:,j})_{\mathbf{M}} \Psi_{:,j} \right\|_{\mathbf{M}}^2 dt \quad \text{s.t.} \quad (\Psi_{:,j}, \Psi_{:,k})_{\mathbf{M}} = \delta_{jk}, \text{ for } 1 \leq j, k \leq l, \quad (5.19)$$

where $(\Phi, \Psi)_{\mathbf{M}} = \Phi^T \mathbf{M} \Psi$ with the finite element mass matrix \mathbf{M} .

In order to compute the POD basis, or equivalently to solve (5.19), there are three ways of which advantages depends on the snapshot matrix.

We define the matrix $\bar{Y}_l = \mathbf{M}^{1/2} Y_l \in \mathbb{R}^{\text{dof} \times N}$ and Cholesky decomposition of the symmetric positive definite mass matrix gives $\mathbf{M} = \mathbf{M}^{1/2} (\mathbf{M}^{1/2})^T$. Then,

- POD basis can be computed using the singular value decomposition (SVD). To do this, SVD of \bar{Y}_l is written as $\bar{Y}_l = U \Sigma V^T$ where $U \in \mathbb{R}^{\text{dof} \times \text{dof}}$, $V \in \mathbb{R}^{N \times N}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{\text{dof} \times N}$ collects the singular values $\sigma_i^2 = \lambda_i$ on its diagonal in a descending order. Then, the i -th column of Ψ is given by

$$\Psi_{:,i} = \mathbf{M}^{-1/2} u_i, \quad (5.20)$$

where u_i is the i -th column of U .

- If $N < \text{dof}$, then we define the correlation matrix $\bar{K} \in \mathbb{R}^{N \times N}$ associated to the snapshot ensemble Y_l as $\bar{K} = \bar{Y}_l^T \bar{Y}_l$. We solve the symmetric eigenvalue problem

$$\bar{K} v_i = \lambda_i v_i, \quad 1 \leq i \leq l, \quad v_i^T v_j = \delta_{ij}, \quad 1 \leq i, j \leq l.$$

Then, the i -th column of Ψ is given by

$$\Psi_{:,i} = Y_l v_i / \sqrt{\lambda_i},$$

where v_i is the i -th column of V .

- If $\text{dof} < N$, then we define the matrix $\bar{R} \in \mathbb{R}^{\text{dof} \times \text{dof}}$ associated to the snapshot ensemble Y_l as $\bar{R} = \bar{Y}_l \bar{Y}_l^T$. We solve the symmetric eigenvalue problem

$$\bar{R} u_i = \lambda_i u_i, \quad 1 \leq i \leq l, \quad u_i^T u_j = \delta_{ij}, \quad 1 \leq i, j \leq l.$$

Then, the i -th column of Ψ is given by

$$\Psi_{:,i} = \mathbf{M}^{-1/2} u_i / \sqrt{\lambda_i},$$

where u_i is the i -th column of U .

For the solution of (5.19), we refer the reader to the studies [74, 83]. We mention the following proposition [83].

Proposition 5.3. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d > 0$ denote the positive eigenvalues of \bar{K} . Then, the approximation error associated to the POD basis of rank $l \leq d$ is given by

$$\sum_{i=1}^N \int_{I_i} \left\| y_{\delta,i} - \sum_{j=1}^l (y_{\delta,i}, \Psi_{:,j})_{\mathbf{M}} \Psi_{:,j} \right\|_{\mathbf{M}}^2 dt = \sum_{j=l+1}^d \lambda_j. \quad (5.21)$$

Accuracy of the reduced solution depends on how much information about the full-order solution is contained in this basis. More accurate solutions can be found by increasing the number of POD basis functions. However, this increases the computational cost. Therefore, basis dimension is decided by balancing the accuracy and the number of the truncated POD basis. Specifically, the most energetic POD modes are chosen by measuring the ratio between the eigenvalues of the retained POD basis and the sum of the whole eigenvalues. To do this, the number of POD basis functions l is decided according to the ratio between the modelled and the total energy,

$$\mathcal{E}(l) = \sum_{i=1}^l \sigma_i^2 / \sum_{i=1}^d \sigma_i^2, \quad (5.22)$$

where σ_i 's denote the singular values of the data matrix \tilde{Y}_l and $d = \text{rank}(\tilde{Y}_l)$.

Remark 5.2. We note that each column of POD basis coefficients, namely $\Psi_{:,1}, \dots, \Psi_{:,l}$ are orthonormal with respect to the mass matrix \mathbf{M} ; while the POD basis functions are orthonormal in $L^2(\Omega)$ norm.

5.2.4 Fully-discrete Optimality System Based on POD

We define the discontinuous test space for the reduced-order model:

$$V_{h,l}^{k,q} = \{v \in L^2(I; V_{h,l}) : v|_{I_m} \in P_q(I_m, V_{h,l}), m = 1, \dots, N_T, v_m(0) \in L^2(\Omega)\}. \quad (5.23)$$

We discretize the semi-discrete reduced-order optimality system in (5.16) in time using dG method. Then, POD Galerkin model of the optimality system associated to DCR

equation is written as follows:

$$\begin{aligned} & \sum_{m=1}^{N_T} \int_{I_m} (\partial_t y_\delta^l, \psi) dt + \int_0^T a_h^s(y_\delta^l, \psi) dt + \sum_{m=1}^{N_T} ([y_\delta^l]_{m-1}, \psi_+^{m-1}) \\ &= \int_0^T (f_\delta^l + u_\delta^l, \psi) dt, \quad \forall \psi \in V_{h,l}^{k,q}, \quad y_{\delta,0}^{l,-} = (y_0)_\delta^l, \end{aligned} \quad (5.24a)$$

$$\begin{aligned} & \sum_{m=1}^{N_T} \int_{I_m} (-\partial_t \lambda_\delta^l, \psi) dt + \int_0^T (a_{h,y}^s(y_\delta^l, \psi))^T \lambda_\delta dt - \sum_{m=1}^{N_T} ([\lambda_\delta^l]_m, \psi_-^m) \\ &= - \int_0^T (y_\delta^l - y_\delta^{l,d}, \psi) dt, \quad \forall \psi \in V_{h,l}^{k,q}, \quad \lambda_{\delta,N}^{l,+} = 0, \end{aligned} \quad (5.24b)$$

$$\int_0^T (\alpha u_\delta^l - \lambda_\delta^l, \psi - u_\delta^l) dt = 0, \quad \forall \psi \in V_{h,l}^{k,q}. \quad (5.24c)$$

We proceed by projecting the linear systems given in Sec.4.4 onto the low-dimensional subspace for DCR equation. We note that the initial conditions for state equation on each time interval I_m must be projected on to the low-dimensional space as follows:

$$Y_m^{l,0} = \Psi^T (y_\delta)_{m-1}^- \quad \text{if } m \geq 2 \quad \text{or} \quad (y_\delta)_{m-1}^{l,-} = \Psi^T y_0^- \quad \text{if } m = 0.$$

Then, for dG(1)-method, the reduced-order system associated to state equation is written as:

$$\begin{aligned} & \begin{pmatrix} \frac{3}{4} \mathbf{M}^l + \frac{k_m}{2} \mathbf{A}^l & \frac{1}{4} \mathbf{M}^l \\ -\frac{9}{4} \mathbf{M}^l & \frac{5}{4} \mathbf{M}^l + \frac{k_m}{2} \mathbf{A}^l \end{pmatrix} \begin{pmatrix} Y_m^{l,1} \\ Y_m^{l,2} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{M}^l Y_m^{l,0} + \frac{k_m}{2} \mathbf{F}_\delta^l(t_{m,1}) + \frac{k_m}{4} \mathbf{M}^l (U_m^{l,0} + U_m^{l,1}) \\ -\mathbf{M}^l Y_m^{l,0} + \frac{k_m}{2} \mathbf{F}_\delta^l(t_{m,2}) + \frac{k_m}{4} \mathbf{M}^l (3U_m^{l,1} - U_m^{l,0}) \end{pmatrix}, \end{aligned} \quad (5.25)$$

on each time interval $I_m = (t_{m-1}, t_m]$. Again the discrete state at $t = t_m$ is given as $y_{l,t_m} = Y_m^{l,2}$.

To derive the reduced-order system associated to adjoint equation, the initial conditions for the adjoint equation on each time interval I_m must be projected onto the low-dimensional space as:

$$\lambda_m^{l,2} = \Psi^T (\lambda_\delta)_m^+ \quad \text{if } m \leq N_T \quad \text{or} \quad \lambda_m^{l,2} = 0 \quad \text{if } m = N_T.$$

On each time interval $I_m = (t_{m-1}, t_m]$, the reduced-order system for the adjoint equation is read as

$$\begin{aligned} & \begin{pmatrix} \frac{3}{4} \mathbf{M}^l + \frac{k_m}{2} \mathbf{A}^l & \frac{1}{4} \mathbf{M}^l \\ -\frac{9}{4} \mathbf{M}^l & \frac{5}{4} \mathbf{M}^l + \frac{k_m}{2} \mathbf{A}^l \end{pmatrix} \begin{pmatrix} \lambda_m^{l,1} \\ \lambda_m^{l,0} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{M}^l \lambda_m^{l,2} + \frac{k_m}{2} \mathbf{Y}_\delta^{l,d}(t_{m,1}) - \frac{k_m}{4} \mathbf{M}^l (Y_m^{l,1} + Y_m^{l,2}) \\ -\mathbf{M}^l \lambda_m^{l,2} + \frac{k_m}{2} \mathbf{Y}_\delta^{l,d}(t_{m,0}) - \frac{k_m}{4} \mathbf{M}^l (3Y_m^{l,1} - Y_m^{l,2}) \end{pmatrix}. \end{aligned} \quad (5.26)$$

In Alg. 5.1, we summarize the derivation of the reduced-order solution of the OCP based on POD step by step.

Algorithm 5.1 Reduced-order approximation for the OCP based on POD

- 1: **if** Snapshots of the state and the adjoint are not given **then**
 - 2: Solve the full-order problem and keep the solutions of the state and the adjoint equation.
 - 3: **end if**
 - 4: Construct the snapshot matrix W (using the state $W := Y$; the adjoint $W := P$ or a combination of them $W := Y \cup P$.)
 - 5: Compute the POD basis $\psi(x)$ using the snapshot matrix W .
 - 6: **if** The state equation is nonlinear. **then**
 - 7: Compute a new basis for the nonlinear term using DEIM.
 - 8: **end if**
 - 9: Solve the reduced-order model of the OCP.
 - 10: Increase the number of POD basis function measuring (5.22).
-

5.3 A Priori Error Estimates for Optimal Control of Diffusion-Convection-Reaction Equation

In this section, we prove some useful lemmas and state the main estimate of this study for DCR equation, particularly for dG time discretization. Firstly, in Sec. 5.3.1, we prove some useful lemmas. Then, in Sec. 5.3.2, we derive the main estimate of this study for POD method.

5.3.1 Auxiliary Results

We use the mass matrix as a weight matrix in the computation of POD basis functions, so \mathcal{M}^l is the identity matrix of size $l \times l$. We define the reduced stiffness matrix \mathcal{S}^l as

$$\mathbf{S}_{ij}^l = \sum_{K \in \mathcal{T}_h} (\nabla \psi_j, \nabla \psi_i)_K.$$

We mention the following bounds given in [83, Lemma. 2]:

$$\sum_{K \in \mathcal{T}_h} |y|_{H^1(K)}^2 \leq \|\mathcal{S}^l\|_2 \|y\|^2, \quad \forall y \in V_{k,q}^{h,l}, \quad (5.27)$$

where $\|\cdot\|_2$ denotes the spectral norm for symmetric matrices.

Lemma 5.4. *For all $y \in V_{h,l}^{k,q}$, we have $\|y\|_{\partial K}^2 \leq \|y\|_K^2$.*

Proof. The proof is motivated by the study [134, Thm. 3]. For any element $y \in V_{h,l}^{k,q}$, we have

$$y|_{\partial K} = \sum_{j=1}^d (y|_{\partial K}, (\psi_j)|_{\partial K}) (\psi_j)|_{\partial K} := \sum_{j=1}^d x_j (\psi_j)|_{\partial K} = \mathbf{x}^T \mathcal{M}_{|\partial K}^l \mathbf{x},$$

where $\mathcal{M}_{|\partial K}^l$ is the reduced-order edge mass matrix. Suppose that we have the snapshots of $y_{|\partial K}$, namely the solution along the edge ∂K on the time interval I_n for $n = 1, \dots, N$. Then, POD basis $(\psi_j)_{|\partial K}$ of order \tilde{l} (with respect to L^2 -norm) can be computed in order to write a reduced-order approximation of $y_{|\partial K}$. Then, the reduced-order edge mass matrix, namely the identity matrix of size \tilde{l} , can be found. Thus, $\|y\|_{\partial K}^2 \leq \mathbf{x}^T \mathbf{x} = \|y\|_K^2$, where the last inequality follows from Parseval's inequality. \square

Lemma 5.5. *The following inequalities are satisfied:*

$$\sum_{K \in \mathcal{T}_h} \|v_h(\boldsymbol{\beta} \cdot \nabla w_h)\|_{L^2(K)} \leq C_\beta \sqrt{\|\mathcal{S}^l\|_2} \|v_h\| \|w_h\|, \quad v_h, w_h \in V_{h,l}, \quad (5.28a)$$

$$\sum_{K \in \mathcal{T}_h} \left(\|v_h\|_{\partial K^+ \cap \partial \Omega^+}^2 + \|v_h^e\|_{\partial K^- \setminus \partial \Omega^-}^2 \right) \leq C_\beta \|v_h\|^2, \quad v_h \in V_{h,l}, \quad (5.28b)$$

where $\|\boldsymbol{\beta}\|_2 \leq C_\beta$ and \mathcal{S}^l is the reduced stiffness matrix.

Proof. To prove (5.28a), we write the following upper bound

$$\sum_{K \in \mathcal{T}_h} \|v_h(\boldsymbol{\beta} \cdot \nabla w_h)\|_{L^2(K)} \leq \sum_{K \in \mathcal{T}_h} C_\beta \|v_h\|_{L^2(K)} |w_h|_{H^1(K)}.$$

The rest of the proof follows from (5.27). For the inequality (5.28b), we have

$$\|v_h\|_{\partial K^+ \cap \partial \Omega^+}^2 + \|v_h^e\|_{\partial K^- \setminus \partial \Omega^-}^2 \stackrel{[54, Eqn. (5.21)]}{\leq} C_\beta \|y\|_{\partial K}^2.$$

Then, Lemma (5.4) is used. \square

Lemma 5.6. *The bilinear form $a^d(\cdot, \cdot)$ satisfies the following inequality.*

$$\begin{aligned} & a^{cr}(v_h, w_h) \\ & \stackrel{[54, Thm. 5.1]}{\leq} r \|v_h\| \|w_h\| + \frac{1}{4} \sum_{K \in \mathcal{T}_h} \left(\|w_h\|_{\partial K^+ \cap \Gamma^+}^2 + \|w_h\|_{\partial K^- \setminus \Gamma^-}^2 \right) \\ & + \sum_{K \in \mathcal{T}_h} \left(\|v_h\|_{\partial K^+ \cap \Gamma^+}^2 + \|v_h^e\|_{\partial K^- \setminus \Gamma^-}^2 \right) + \sum_{K \in \mathcal{T}_h} \|\boldsymbol{\beta} - \tilde{\pi} \boldsymbol{\beta}\|_{L^\infty(K)} \|v_h\|_{L^2(K)} |w_h|_{H^1(K)} \\ & \stackrel{(5.28)}{\leq} r \|v_h\| \|w_h\| + \frac{1}{4} \sum_{K \in \mathcal{T}_h} \left(\|w_h\|_{\partial K^+ \cap \Gamma^+}^2 + \|w_h\|_{\partial K^- \setminus \Gamma^-}^2 \right) \\ & + C_\beta \|v_h\|^2 + C_\beta \sqrt{\|\mathcal{S}^l\|_2} \|v_h\| \|w_h\|, \quad v_h, w_h \in V_{h,l}. \end{aligned} \quad (5.29)$$

Proof. The proof is adapted from [54, Thm. 5.1] using the inequalities (5.28). \square

In order to approximate $\|y^{n-1} - P^l y^{n-1}\|$, we make the following definition.

Definition 5.4. We define the following minimization problem associated to y_δ^i ,

$$\min_{\{\psi_j\}_{j=1}^l} \sum_{i=1}^N \left\| y_\delta^i - \sum_{j=1}^l (y_\delta^i, \psi_j)_{\mathbf{M}} \psi_j \right\|_{\mathbf{M}}^2 \quad \text{s.t.} \quad (\psi_j, \psi_k)_{\mathbf{M}} = \delta_{jk}, \text{ for } 1 \leq j, k \leq l, \quad (5.30)$$

such that the following error bound holds

$$\sum_{i=1}^N \left\| y_\delta^i - \sum_{j=1}^l (y_\delta^i, \psi_j)_{\mathbf{M}} \psi_j \right\|_{\mathbf{M}}^2 = \sum_{j=l+1}^d \lambda_j. \quad (5.31)$$

L^2 -norm of the POD projection error is given by (5.21). DG energy norm of the POD projection error is proved in the following lemma where the spatial discretization, namely SIPG method, leads to the term $\|\psi_j\|_{DG}^2$ on the right-hand side of the estimate.

Lemma 5.7. *POD projection error in dG energy norm satisfies*

$$\int_0^T \left\| y_h - \sum_{j=1}^l (y_h, \psi_j)_{\mathbf{M}} \psi_j \right\|_{DG}^2 dt \leq 2C_A \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2, \quad (5.32a)$$

$$\sum_{i=1}^N \int_{I_i} \left\| y_\delta - \sum_{j=1}^l (y_\delta, \psi_j)_{\mathbf{M}} \psi_j \right\|_{DG}^2 dt \leq 2C_A \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2, \quad (5.32b)$$

$$\sum_{i=1}^N \left\| y_\delta^i - \sum_{j=1}^l (y_\delta^i, \psi_j)_{\mathbf{M}} \psi_j \right\|_{DG}^2 \leq 2C_A \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2. \quad (5.32c)$$

Proof. We use the idea in the proof of [76, Lemma 3.2]. In particular, we give the proof of (5.32b). Other proofs follow the same idea. Let $Y_l = [y_\delta|_{I_1}, \dots, y_\delta|_{I_N}]$ be the $M \times N$ snapshot matrix. The POD basis can be found by solving the following eigenvalue problem

$$Y_l Y_l^T \psi_j = \lambda_j \psi_j \in \mathbb{R}^M, \quad j = 1, \dots, l, \quad (5.33)$$

with the snapshot matrix Y . The POD projection error in dG norm satisfies

$$\begin{aligned}
& \sum_{i=1}^N \int_{I_i} \left\| \left\| y_\delta|_{I_i} - \sum_{j=1}^l (y_\delta|_{I_i}, \psi_j)_{\mathbf{M}} \psi_j \right\| \right\|_{DG}^2 dt = \sum_{i=1}^N \int_{I_i} \left\| \left\| \sum_{j=l+1}^d (y_\delta|_{I_i}, \psi_j)_{\mathbf{M}} \psi_j \right\| \right\|_{DG}^2 dt \\
& \stackrel{(4.27)}{\leq} \frac{2}{\epsilon} \sum_{i=1}^N \int_{I_i} a^d \left(\sum_{j=l+1}^d (y_\delta|_{I_i}, \psi_j)_{\mathbf{M}} \psi_j, \sum_{k=l+1}^d (y_\delta|_{I_i}, \psi_k)_{\mathbf{M}} \psi_k \right) dt \\
& = \frac{2}{\epsilon} \sum_{i=1}^N \int_{I_i} \sum_{j=l+1}^d \sum_{k=l+1}^d (y_\delta|_{I_i}, \psi_j)_{\mathbf{M}} (y_\delta|_{I_i}, \psi_k)_{\mathbf{M}} dt a^d(\psi_j, \psi_k) \\
& = \frac{2}{\epsilon} \sum_{j=l+1}^d \sum_{k=l+1}^d \left(\sum_{i=1}^N \int_{I_i} (y_\delta|_{I_i}, \psi_j)_{\mathbf{M}} y_\delta|_{I_i}, \psi_k dt \right)_{\mathbf{M}} a^d(\psi_j, \psi_k) \\
& = \frac{2}{\epsilon} \sum_{j=l+1}^d \sum_{k=l+1}^d (YY^T \psi_j, \psi_k)_{\mathbf{M}} a^d(\psi_j, \psi_k) \stackrel{(5.33)}{=} \frac{2}{\epsilon} \sum_{j=l+1}^d \sum_{k=l+1}^d (\lambda_j \psi_j, \psi_k)_{\mathbf{M}} a^d(\psi_j, \psi_k) \\
& = \frac{2}{\epsilon} \sum_{j=l+1}^d \sum_{k=l+1}^d \lambda_j \delta_{jk} a^d(\psi_j, \psi_k) \stackrel{(4.26)}{\leq} 2C_A \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2, \tag{5.34}
\end{aligned}$$

which proves (5.32b). \square

We define the dG-elliptic projection $P^h : V \rightarrow V_{h,p}$ (if P^h is restricted to the dG-elliptic projection from $V_{h,p}$ to $V_{h,l}$, it is written as P^l) such that $P^h|_{V_{h,p}} = P^l : V_{h,p} \rightarrow V_{h,l}$ and $P^h : V \setminus V_{h,p} \rightarrow V_{h,p} \setminus V_{h,l}$ denoted by

$$a^d(y - P^h y, v_h) = 0, \quad \forall v_h \in V_{h,p}, \forall y \in V, \forall t \geq 0. \tag{5.35}$$

The projection operator P^l satisfies the following bounds.

Lemma 5.8. *For every l ($1 \leq l \leq d$), the projection operator P^l satisfies*

$$\sum_{i=1}^N \int_{I_i} \|y - P^l y\|_{DG}^2 dt \leq C \left(h^{2p} + \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right), \tag{5.36a}$$

$$\sum_{i=1}^N \int_{I_i} \|y - P^l y\|^2 dt \leq C \left(h^{2(p+1)} + h^2 \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right), \tag{5.36b}$$

$$\sum_{i=1}^N \|y^i - P^l y^i\|^2 \leq C \left(h^{2p+1} + h \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right). \tag{5.36c}$$

Proof. We start with (5.36a). For any $v \in V$, we deduce from (5.35),(4.27) that

$$\begin{aligned}
\|v - P^h v\|_{DG}^2 & \leq \frac{2}{\epsilon} a^d(v - P^h v, v - P^h v) = \frac{2}{\epsilon} a^d(v - P^h v, v - v_h) \\
& \leq 2C_A \|v - P^h v\|_{DG} \|v - v_h\|_{DG}, \quad \forall v_h \in V_{h,p}. \tag{5.37}
\end{aligned}$$

Then, we obtain that $\|v - P^h v\|_{DG} \leq 2C_A \|v - v_h\|_{DG}$, $\forall v_h \in V_{h,p}$. If P^h is restricted to dG-elliptic projection from $V_{h,p}$ to $V_{h,l}$, i.e., $P^h y = P^l y \in V_{h,l}$, then we obtain

$$\|v - P^l v\|_{DG} \leq 2C_A \|v - v_h\|_{DG}, \quad \forall v_h \in V_{h,l}. \quad (5.38)$$

We choose $v = y$ with $y = y(\cdot, t)$ and decompose the right-hand side of (5.38) as $y - v_h = (y - y_h) + (y_h - v_h)$, where y_h is the solution of the semi-discrete problem.

We choose $v_h = \sum_{j=1}^l (y_h, \psi_j) \mathbf{M} \psi_j \in V_{h,l} \subset V_{h,p}$. Using the estimates [5, Lemma 5] and 5.32), we prove (5.36a).

In order to prove (5.36b), we follow the approach given in the proof of [42, Lemma 4.2]. Let us mention the techniques in that study. Firstly, the following problem is considered:

$$-\Delta w = v - P^h v, \quad \text{in } \Omega, \quad w = 0, \quad \text{on } \partial\Omega. \quad (5.39)$$

With $\bar{V} = \{v \in C^\infty(\Omega) : \text{supp } v \subset \Omega\}$, the weak formulation of (5.39) is written as follows: Find $w \in H^1(\Omega)$ such that $w|_{\partial\Omega} = 0$ and

$$(\nabla w, \nabla q) = (v - P^h v, q), \quad \forall q \in \bar{V}. \quad (5.40)$$

Suppose that $w \in H^2(\Omega)$ and there is a constant $C > 0$, independent of $v - P^h v$, such that

$$\|w\|_{H^2(\Omega)} \leq C \|v - P^h v\|. \quad (5.41)$$

As the domain Ω is convex, this is true. In addition, $H^2(\Omega) \subset C(\bar{\Omega})$. Secondly, let $w \in H^2(\Omega)$ be the solution of the problem (5.39) satisfying (5.41), w_h be the piecewise Lagrange interpolant of w satisfying $w_h \in C(\bar{\Omega}) \cap V_{h,p}$ and $w|_{\partial\Omega} = 0$. Then,

$$\|w - w_h\|_{DG}^2 = |w - w_h|_{H^1(\Omega, \mathcal{T}_h)}^2 \leq Ch^2 |w|_{H^2(\Omega)}^2, \quad (5.42)$$

(see, for example, [32, Thm. 3.1.6]). Due to $w \in H^2(\Omega)$ [42], we have $[[v]] = 0 = [[\nabla v]]$, $\forall E \in \mathcal{E}_h$. Now, we continue as in [98, Lem. 3.3] by taking $q = v - P^h v$ in (5.40)

$$\begin{aligned} \|v - P^h v\|^2 &= (\nabla w, \nabla(v - P^h v)) = (\nabla(w - w_h), \nabla(v - P^h v)) \quad \forall w_h \in V_{h,p} \\ &\leq |w - w_h|_{H^1(\Omega, \mathcal{T}_h)} \|v - P^h v\|_{DG} \\ &\stackrel{(5.42)}{\leq} Ch |w|_{H^2(\Omega)} \|v - P^h v\|_{DG} \stackrel{(5.41)}{\leq} Ch \|v - P^h v\| \|v - P^h v\|_{DG}. \end{aligned} \quad (5.43)$$

Then, we obtain

$$\|v - P^h v\| \leq Ch \|v - P^h v\|_{DG}.$$

If P^h is restricted to dG-elliptic projection from $V_{h,p}$ to $V_{h,l}$, i.e., $P^h y = P^l y \in V_{h,l}$, then we obtain

$$\|v - P^l v\| \leq Ch \|v - P^l v\|_{DG}.$$

The rest of the proof follows from (5.38). For (5.36a), we have

$$\begin{aligned}
\sum_{i=1}^N \|y^i - P^l y^i\|^2 &= C \sum_{i=1}^N h^2 \|y^i - P^l y^i\|_{DG}^2 \\
&\leq CNh^2 \left(C_P h^{2p} + C_A \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right) \\
&\leq C \left(C_P h^{2p+1} + C_A h \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right).
\end{aligned}$$

□

Since we use space-time discontinuous Galerkin method as a discretization technique, we adapt the interpolant given in [49, Sec.3] as follows: $\pi^l y \in V_{h,l}^{k,q}$ of y as $\pi^l y = \Pi_{k,n}(P^l y)$, $n = 1, \dots, N$. For a function $v(t)$, $\Pi_{k,n}v$ is the unique polynomial of degree q determined by $(\Pi_{k,n}v)^n = v^n$. For $q \geq 1$,

$$\int_{I_n} \Pi_{k,n}v t^j dt = \int_{I_n} v t^j dt, \quad j = 0, \dots, q-1, \quad n = 1, \dots, N. \quad (5.44)$$

Lemma 5.9. *For every l ($1 \leq l \leq d$), the projection operator π^l satisfies*

$$\sum_{i=1}^N \int_{I_i} \|y - \pi^l y\|_{DG}^2 dt \leq C_\pi (k^{2(q+1)} + h^{2p}) \|y\|_{*,I}^2 + C_A \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2, \quad (5.45a)$$

$$\sum_{i=1}^N \int_{I_i} \|y - \pi^l y\|^2 dt \leq C_\pi (k^{2(q+1)} + h^{2(p+1)}) \|y\|_{*,I}^2 + C_A h^2 \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2, \quad (5.45b)$$

$$\sum_{i=1}^N \|y^{i-1} - \pi^l y^{i-1}\|^2 \leq C_P h^{2p+1} \|y\|_{L^\infty(I; H^{p+1}(\Omega))}^2 + C_A h \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2. \quad (5.45c)$$

Proof. We apply triangle inequality as follows

$$\|y - \pi^l y\| \leq \|y - \pi^h y\| + \|\Pi_{k,n}(P^h - I)y\| + \|\Pi_{k,n}(I - P^l)y\|.$$

Then, we use the properties of the operators P^h , $\Pi_{k,n}$, π^h and the estimate (5.36). For the last term, we use the properties of $\pi^l y$ which leads to the following relation

$$\|y^{n-1} - \pi^l y^{n-1}\|^2 \leq \|y^{n-1} - P^l y^{n-1}\|^2.$$

Then, the rest follows from (5.36c). □

Corollary 5.10. For every l ($1 \leq l \leq d$), the projection operator π^l satisfies

$$\sum_{i=1}^N \int_{I_i} \|y_\delta - \pi^l y_\delta\|_{DG}^2 dt \leq C_A \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2, \quad (5.46a)$$

$$\sum_{i=1}^N \int_{I_i} \|y_\delta - \pi^l y_\delta\|^2 dt \leq C_A h^2 \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2, \quad (5.46b)$$

$$\sum_{i=1}^N \|y_\delta^{i-1} - \pi^l y_\delta^{i-1}\|^2 \leq C_A h \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2. \quad (5.46c)$$

5.3.2 Main Result

We define auxiliary state and adjoint equation which are needed for a priori error analysis

$$\begin{aligned} \sum_{m=1}^{N_T} \int_{I_m} (\partial_t y_\delta^{l,u}, v_\delta) dt + \int_0^T a_h^s(y_\delta^{l,u}, v_\delta) dt + \sum_{m=1}^{N_T} ([y_\delta^{l,u}]_{m-1}, v_{\delta,+}^{m-1}) \\ = \int_0^T (f_\delta^l + u_\delta, v_\delta) dt, \quad y_{\delta,-}^{u,l,0} = (y_0)_\delta^l, \end{aligned} \quad (5.47a)$$

$$\begin{aligned} \sum_{m=1}^{N_T} \int_{I_m} (-\partial_t \lambda_\delta^{l,u}, \phi_\delta) dt + \int_0^T a_h(\lambda_\delta^{l,u}, \phi_\delta) dt - \sum_{m=1}^{N_T} ([\lambda_\delta^{l,u}]_m, \phi_{\delta,-}^m) \\ = - \int_0^T (y_\delta^{l,u} - y_\delta^d, \phi_\delta) dt, \quad \lambda_{\delta,+}^{u,l,N} = 0. \end{aligned} \quad (5.47b)$$

We shall derive some useful lemmas before stating the main estimate of this study.

Lemma 5.11. Let $(y_\delta^l, \lambda_\delta^l)$ be the reduced solution in (5.24) and $(y_\delta^{l,u}, \lambda_\delta^{l,u})$ be the auxiliary solution in (5.47). Then, there exists a constant C independent of h, k, l such that

$$\|y_\delta^l - y_\delta^{l,u}\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda_\delta^l - \lambda_\delta^{l,u}\|_{L^\infty(0,T;L^2(\Omega))} \leq C \|u_\delta - u_\delta^l\|_{L^2(0,T;L^2(\Omega))}. \quad (5.48)$$

Proof. Firstly, we consider the state equation. We subtract (5.24a) from (5.47a) to obtain the local error equation with $\theta = y_\delta^{l,u} - y_\delta^l$:

$$\int_{I_m} ((\partial_t \theta, v) + a_h^s(\theta, v)) dt + ([\theta]_{m-1}, v_+^{m-1}) = \int_{I_m} (u_\delta - u_\delta^l, v) dt, \quad \forall v \in V_{h,l}^{k,q}. \quad (5.49)$$

We substitute $v_\delta = 2\theta$ in (5.49). We know that

$$\int_{I_m} 2(\partial_t \theta, \theta) dt + 2([\theta]_{m-1}, \theta_+^{m-1}) = \|\theta_-^m\|^2 - \|\theta_-^{m-1}\|^2 + \|[\theta]_{m-1}\|^2, \quad (5.50)$$

is valid. For the right-hand side, we employ Cauchy-Schwarz, Young inequalities, Poincaré inequality (2.4) and the definition of dG norm (2.1). For the left-hand side, we use (4.27) for diffusion term and follow the technique in (see [54, Theorem 5.1]) for convection and reaction terms. Then, we derive the following estimate

$$\begin{aligned}
\|\theta_-^m\|^2 &= \|\theta_-^{m-1}\|^2 + \frac{\epsilon}{2} \int_{I_m} \|\theta\|_{DG}^2 dt + 2C_0 \int_{I_m} \|\theta\|^2 dt \\
&+ \int_{I_m} \left(\sum_{K \in \mathcal{T}_h} \left(\|\theta\|_{\partial K^- \cap \Gamma^-}^2 + \|[\![\theta]\!] \|_{\partial K^- \setminus \Gamma^-}^2 + \|\theta\|_{\partial K^+ \cap \Gamma^+}^2 \right) \right) dt \\
&\leq C \int_{I_m} \|u_\delta - u_\delta^l\|^2 dt.
\end{aligned} \tag{5.51}$$

Then, we sum the resulting inequality from $m = 1$ to n to arrive at

$$\begin{aligned}
\|\theta_-^n\|^2 + \frac{\epsilon}{2} \int_0^{t_n} \|\xi\|_{DG}^2 dt \\
+ \int_0^{t_n} \sum_{K \in \mathcal{T}_h} \left(\|\xi\|_{\partial K^- \cap \Gamma^-}^2 + \|[\![\xi]\!] \|_{\partial K^- \setminus \Gamma^-}^2 + \|\xi\|_{\partial K^+ \cap \Gamma^+}^2 \right) dt \\
\leq \|\theta_-^0\|^2 + C \int_0^{t_n} \|u_\delta - u_\delta^l\|^2 dt.
\end{aligned} \tag{5.52}$$

Then, (5.52) yields that

$$\|\theta\|_{L^\infty(0,T;L^2(\Omega))}^2 \leq C \|u_\delta - u_\delta^l\|_{L^2(0,T;L^2(\Omega))}^2. \tag{5.53}$$

Secondly, we consider the adjoint equation. We subtract (5.24b) from (5.47b) to obtain the local error equation with $\varphi = p_\delta^{l,u} - p_\delta^l$:

$$\int_{I_m} \left(-(\partial_t \varphi, v) + a_h(\varphi, v) \right) dt - ([\varphi]_m, v_-^m) = - \int_{I_m} (y_\delta^{l,u} - y_\delta^l, v) dt, \quad \forall v \in V_{h,l}^{k,q}. \tag{5.54}$$

Apply the same idea above to find

$$\|\varphi\|_{L^\infty(0,T;L^2(\Omega))}^2 \leq C \|y_\delta^l - y_\delta^{l,u}\|_{L^2(0,T;L^2(\Omega))}^2 \leq C \|u_\delta - u_\delta^l\|_{L^2(0,T;L^2(\Omega))}^2. \tag{5.55}$$

The final result (5.48) is obtained through the inequalities (5.53,5.55). \square

We shall proceed with the estimate associated to the control.

Lemma 5.12. *Let (y, λ, u) and $(y_\delta^l, \lambda_\delta^l, u_\delta^l)$ be the solutions of (2.6) and (5.24), respectively. Then, we have*

$$\|u_\delta - u_\delta^l\|_{L^2(0,T;L^2(\Omega))} \leq \frac{1}{\alpha} \|\lambda - \lambda_\delta^{l,u}\|_{L^2(0,T;L^2(\Omega))}. \tag{5.56}$$

Proof. We apply the technique used for the steady-state optimal control problem in [91, Section 4.2]. We start using the continuous and reduced optimality conditions (2.6c)-(5.24c) to obtain the following equation

$$\begin{aligned}
\alpha \|u_\delta - u_\delta^l\|_{L^2(0,T;L^2(\Omega))}^2 &= \alpha \int_0^T (u_\delta - u_\delta^l, u_\delta - u_\delta^l) dt \\
&= \int_0^T (\alpha u_\delta - \lambda, u_\delta - u_\delta^l) dt - \int_0^T (\alpha u_\delta^l - \lambda_\delta^l, u_\delta - u_\delta^l) dt + \int_0^T (\lambda - \lambda_\delta^l, u_\delta - u_\delta^l) dt \\
&= \int_0^T (\lambda - \lambda_\delta^{l,u}, u_\delta - u_\delta^l) dt + \int_0^T (\lambda_\delta^{l,u} - \lambda_\delta^l, u_\delta - u_\delta^l) dt = J_1 + J_2. \tag{5.57}
\end{aligned}$$

We use Cauchy-Schwarz and Young inequalities to show that

$$0 \leq J_1 \leq \frac{1}{2\alpha} \|\lambda - \lambda_\delta^{l,u}\|_{L^2(0,T;L^2(\Omega))}^2 + \frac{\alpha}{2} \|u_\delta - u_\delta^l\|_{L^2(0,T;L^2(\Omega))}^2. \tag{5.58}$$

We proceed with J_2 and use the auxiliary state equation (5.47a) to obtain

$$\begin{aligned}
J_2 &= \int_0^T (\lambda_\delta^{l,u} - \lambda_\delta^l, u - u_\delta^l) dt \\
&= \sum_{m=1}^{N_T} \int_{I_m} (\partial_t (y_\delta^{l,u} - y_\delta^l), \lambda_\delta^{l,u} - \lambda_\delta^l) dt + \int_0^T a_h^s (y_\delta^{l,u} - y_\delta^l, \lambda_\delta^{l,u} - \lambda_\delta^l) dt \\
&\quad + \sum_{m=1}^N ([y_\delta^{l,u} - y_\delta^l]_{m-1}, (\lambda_\delta^{l,u} - \lambda_\delta^l)_+^{m-1}).
\end{aligned}$$

We proceed applying integration by parts in time and use the auxiliary adjoint equation (5.47b) to arrive at

$$\begin{aligned}
J_2 &= - \sum_{m=1}^{N_T} \int_{I_m} (\lambda_\delta^{l,u} - \lambda_\delta^l, \partial_t (y_\delta^{l,u} - y_\delta^l)) dt + \sum_{m=1}^N (y_\delta^{l,u} - y_\delta^l, \lambda_\delta^{l,u} - \lambda_\delta^l)|_{t_{m-1}}^{t_m} \\
&\quad + \int_0^T a_h^s (y_\delta^{l,u} - y_\delta^l, \lambda_\delta^{l,u} - \lambda_\delta^l) dt + \sum_{m=1}^N ([y_\delta^{l,u} - y_\delta^l]_{m-1}, (\lambda_\delta^{l,u} - \lambda_\delta^l)_+^{m-1}) \\
&= - \sum_{m=1}^{N_T} \int_{I_m} (\lambda_\delta^{l,u} - \lambda_\delta^l, \partial_t (y_\delta^{l,u} - y_\delta^l)) dt + \int_0^T a_h^s (y_\delta^{l,u} - y_\delta^l, \lambda_\delta^{l,u} - \lambda_\delta^l) dt \\
&\quad - \sum_{m=1}^N ((y_\delta^u - y_\delta)_-^m, [\lambda_\delta^u - \lambda_\delta]_m) = - \int_0^T (y_\delta^u - y_\delta, y_\delta^u - y_\delta) dt \leq 0. \tag{5.59}
\end{aligned}$$

Then, using (5.57)-(5.59), we derive the final result (5.12). \square

Lemma 5.13. *For every l ($1 \leq l \leq d$), Then, the error between dG solution $(y_\delta, \lambda_\delta)$ in (4.12) and the auxiliary solution $(y_\delta^{l,u}, \lambda_\delta^{l,u})$ in (5.47) satisfies*

$$\|y_\delta - y_\delta^{l,u}\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda_\delta - \lambda_\delta^{l,u}\|_{L^\infty(0,T;L^2(\Omega))} \leq C \left(\sqrt{\sum_{j=l+1}^d \lambda_j} \|\psi_j\|_{DG}^2 \right). \quad (5.60)$$

Proof. We consider the following local error equation of the state:

$$(e^n, v^n) + \int_{I_m} (-e, \partial_t v) + a_h(e, v) dt - (e^{n-1}, v_+^{n-1}) = 0, \quad \forall v \in V_{h,l}^{k,q}. \quad (5.61)$$

Step I Now, we decompose the error as $e = y_\delta - y_\delta^{l,u} = (y_\delta - \pi^l y_\delta) + (\pi^l y_\delta - y_\delta^{l,u}) = \eta + \theta$. The aim is to find an upper bound for θ , since estimates associated to η are given in (5.46). We choose $v = \theta \in V_{h,l}^{k,q}$ in (5.61). Then, we use the estimates (4.26) for diffusion term and (5.29) for convection and reaction terms. Then, we apply Young's inequality to obtain

$$\begin{aligned} & \frac{1}{2} \|\theta^n\|^2 + \frac{1}{2} \|\theta^{n-1} - \theta_+^{n-1}\|^2 - \frac{1}{2} \|\theta^{n-1}\|^2 + \int_{I_m} \frac{\epsilon}{2} \|\theta\|_{DG}^2 dt \\ & + \int_{I_m} \left(r \|\theta\|^2 + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \left(\|\theta\|_{\partial K - \cap \Gamma^-}^2 + \|[\theta]\|_{\partial K - \setminus \Gamma^-}^2 + \|\theta\|_{\partial K + \cap \Gamma^+}^2 \right) \right) dt \\ & \leq \left| - \int_{I_m} a_h(\eta, \theta) dt \right| + \|\eta_-^{n-1}\|^2 + \frac{1}{4} \|\theta^{n-1} - \theta_+^{n-1}\|^2 \\ & \leq \int_{I_m} (C_A \epsilon \|\eta\|_{DG} \|\theta\|_{DG} + r \|\eta\| \|\theta\|) dt \\ & + \int_{I_m} \frac{1}{4} \sum_{K \in \mathcal{T}_h} \left(\|\theta\|_{\partial K + \cap \partial \Omega^+}^2 + \|[\theta]\|_{\partial K - \setminus \partial \Omega^-}^2 \right) dt \\ & + \int_{I_m} \left(C_\beta \|\eta\|^2 + C_\beta \sqrt{\|\mathcal{S}^l\|_2} \|\eta\| \|\theta\| \right) dt + \|\eta^{n-1}\|^2 + \frac{1}{4} \|\theta^{n-1} - \theta_+^{n-1}\|^2. \quad (5.62) \end{aligned}$$

We put θ terms onto the right-hand side after applying Young's inequality and sum the

resulting inequality from $n = 1$ to $n = s$. Then, use the estimates (5.46) to obtain

$$\begin{aligned}
& \frac{1}{2} \|\theta^s\|^2 - \frac{1}{2} \|\theta^0\|^2 + \frac{1}{4} \sum_{n=1}^s \|\theta^{n-1} - \theta_+^{n-1}\|^2 + \int_0^{t_s} \left(\frac{r}{2} \|\theta\|^2 + \frac{\epsilon}{4} \|\|\theta\|\|_{DG}^2 \right) dt \\
& + \frac{1}{4} \int_0^{t_s} \left(\sum_{K \in \mathcal{T}_h} \left(2\|\theta\|_{\partial K^- \cap \partial \Omega^-}^2 + \|\llbracket \theta \rrbracket\|_{\partial K^- \setminus \partial \Omega^-}^2 + \|\theta\|_{\partial K^+ \cap \partial \Omega^+}^2 \right) \right) dt \\
& \leq C_A^3 \epsilon \sum_{j=l+1}^d \lambda_j \|\|\psi_j\|\|_{DG}^2 + \left(C_\beta + \frac{C_\beta^2 \|\mathcal{S}^l\|_2}{r} \right) C_A h^2 \sum_{j=l+1}^d \lambda_j \|\|\psi_j\|\|_{DG}^2 \\
& + C_A h \sum_{j=l+1}^d \lambda_j \|\|\psi_j\|\|_{DG}^2. \tag{5.63}
\end{aligned}$$

Step II We choose $v = \tilde{\theta} \in V_{h,l}^{k,q}$ in (5.61). We use the properties of discrete characteristic function defined in Sec. 4.3.2 which leads to

$$\|\theta(s)\| = \sup_{s \in I_n} \|\theta(s)\| \quad \text{and} \quad - \sup_{s \in I_{n-1}} \|\theta(s)\| \leq -\|\theta^{n-1}\|.$$

Then, we sum the resulting inequality from $n = 1$ to $n = s$ and use the estimates (5.46) to obtain

$$\begin{aligned}
& \frac{1}{2} \sup_{s \in I_r} \|\theta(s)\|^2 - \frac{1}{2} \|\theta^0\|^2 + \frac{1}{2} \sum_{n=1}^s \|\theta^{n-1} - \theta_+^{n-1}\|^2 \\
& \leq \int_0^{t_s} \frac{3}{2} C_A C_D \epsilon \|\|\theta\|\|_{DG}^2 dt + \int_0^{t_s} C_D \left(C_\beta + \frac{3r + 3C_D C_\beta \sqrt{\|\mathcal{S}^l\|_2}}{2} \right) \|\theta\|^2 dt \\
& + \int_0^{t_s} \frac{C_D}{2} \sum_{K \in \mathcal{T}_h} \left(\|\theta\|_{\partial K^+ \cap \partial \Omega^+}^2 + \|\llbracket \theta \rrbracket\|_{\partial K^- \setminus \partial \Omega^-}^2 \right) dt \\
& + \int_0^{t_s} \frac{C_A C_D \epsilon}{2} \|\|\eta\|\|_{DG}^2 dt + \int_0^{t_s} \left(C_\beta + \frac{C_D(r + C_\beta \sqrt{\|\mathcal{S}^l\|_2})}{2} \right) \|\eta\|^2 dt \\
& + \sum_{n=1}^s \|\eta^{n-1}\|^2 + \frac{1}{2} \sum_{n=1}^s \|\theta^{n-1} - \theta_+^{n-1}\|^2. \tag{5.64}
\end{aligned}$$

Then, by bounding $\|\theta\|$ terms on the right-hand side of (5.64) using (5.63), we find the following estimate for state:

$$\|\theta\|_{L^\infty(0,T;L^2(\Omega))}^2 \leq C \left(\sqrt{\sum_{j=l+1}^d \lambda_j \|\|\psi_j\|\|_{DG}^2} \right). \tag{5.65}$$

We continue with the adjoint equation and write the local error equation as follows

$$-(r_+^n, v^n) + \int_{I_m} ((r, \partial_t v) + a_h(v, r)) dt + (r_+^{n-1}, v_+^{n-1}) = -(\xi, v), \quad \forall v \in V_{h,l}^{k,q}.$$

We write the error as

$$r = (\lambda_\delta - \pi^l \lambda_\delta) + (\pi^l \lambda_\delta - \lambda_\delta^{l,u}) = \psi + \mu$$

with the projection π^l defined above. We proceed as we do for the state equation to obtain

$$\|\mu\|_{L^\infty(0,T;L^2(\Omega))}^2 \leq C \|\xi\|_{L^\infty(0,T;L^2(\Omega))}^2. \quad (5.66)$$

Then, an upper bound for the adjoint follows from (5.65). The final result (5.60) is obtained by combining (5.65) and (5.66). \square

We estimate the error between dG and POD solution of the OCP by combining Lemmas 5.11-5.13.

Theorem 5.14. *Suppose that $(y_\delta, \lambda_\delta, u_\delta)$ and $(y_\delta^l, \lambda_\delta^l, u_\delta^l)$ are the solutions of (2.6) and (5.24), respectively. We assume that all conditions of Lemmas 5.11-5.13 are satisfied. Then, there exists a constant C independent of h, k and l such that*

$$\begin{aligned} & \|y_\delta - y_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda_\delta - \lambda_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} + \|u_\delta - u_\delta^l\|_{L^2(0,T;L^2(\Omega))} \\ & \leq C \left(\sqrt{\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2} \right). \end{aligned} \quad (5.67)$$

We estimate the error between the exact and the POD solution of the OCP by combining Theorems 4.8-5.14.

Theorem 5.15. *Suppose that (y, λ, u) and $(y_\delta^l, \lambda_\delta^l, u_\delta^l)$ are the solutions of (2.6) and (5.24), respectively. We assume that all conditions of Theorems 5.14-4.8 are satisfied. Then, there exists a constant C independent of h, k and l such that*

$$\begin{aligned} & \|y - y_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda - \lambda_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} + \|u - u_\delta^l\|_{L^2(0,T;L^2(\Omega))} \\ & \leq C \left(k^{q+1} + h^p + \sqrt{\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2} \right). \end{aligned} \quad (5.68)$$

The estimate (5.68) reveals that the error between the exact and the numerical results is of order q in time, of order p in space and of order one with respect to the POD

truncation error $\sqrt{\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2}$.

5.3.3 Numerical Results

In this section, we present some numerical results to investigate the experimental order of convergence associated to the reduced-order model. Full-order problem is solved

using piecewise linear discontinuous finite elements on a uniform mesh with $h = k = 1/80$ leading to linear systems of size 38400. Three different snapshot sets for W are used to generate the POD basis functions, namely the state Y , the adjoint P and the combination of them $Y \cup P$, as in [72]. POD basis is computed using the singular value decomposition (SVD), because it is more stable than the eigenvalue decomposition, i.e. the singular values decay to machine precision, whereas the eigenvalues stagnate above [118]. We measure the error in the state, in the adjoint and control approximation in terms of $L^\infty(0, 1; L^2(\Omega))$ and $L^2(0, 1; L^2(\Omega))$ norm, respectively.

We consider the optimal control problem with

$$Q = (0, 1] \times \Omega, \quad \Omega = (0, 1)^2,$$

$$\epsilon = 10^{-2}, \quad \beta = (y - 1/2, -x + 1/2)^T, \quad r = 1, \quad \alpha = 1.$$

We take the source function f , the desired state y_d and the initial condition y_0 as

$$f(x, t) = y_d(x, t) = 1, \quad y_0(x, t) = 0.$$

The exact solution of this problem is not known and the convection field is not a constant vector. In Figures 6.2-6.1, we plot the convection field of the state and the adjoint equation for $p = 0.1, 0.5, 0.9$.

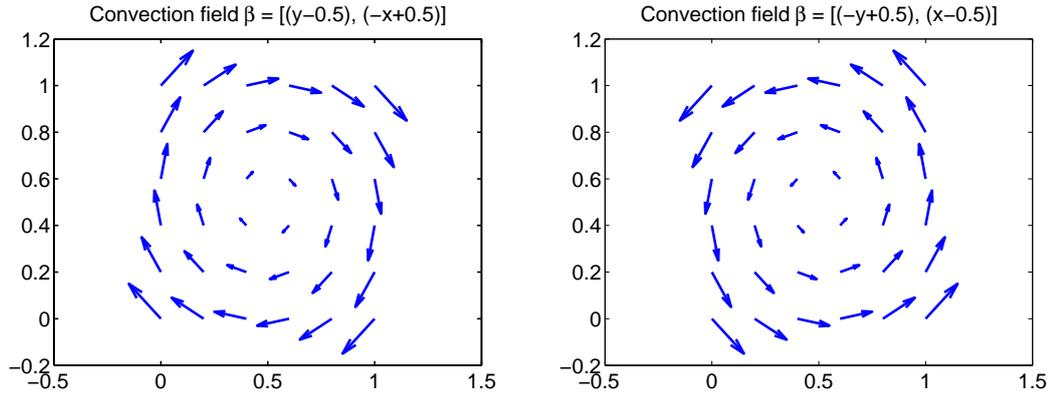


Figure 5.1: Convection field of the state and the adjoint, respectively

In Figure 5.2, we present the numerical solutions of the state at the time instances $t = 0.2, 0.6, 1$ from left to right. In Figure 5.3, numerical solution of the adjoint are given at the time instances $t = 0.8, 0.4, 0$ from left to right. Due to the convection field β , state rotates clockwise as $t \rightarrow T = 1$, while adjoint follows the counter-clockwise direction as $t \rightarrow t_0 = 0$.

In Figure 5.4, we present the decay of the eigenvalues for three different snapshot sets. We observe that the eigenvalues decrease rapidly showing that POD can be successfully applied.

In Figure 5.5, we plot the error in the state, the adjoint and the control with respect to the number of POD basis functions. Firstly, we comment on the state solution. Snap-

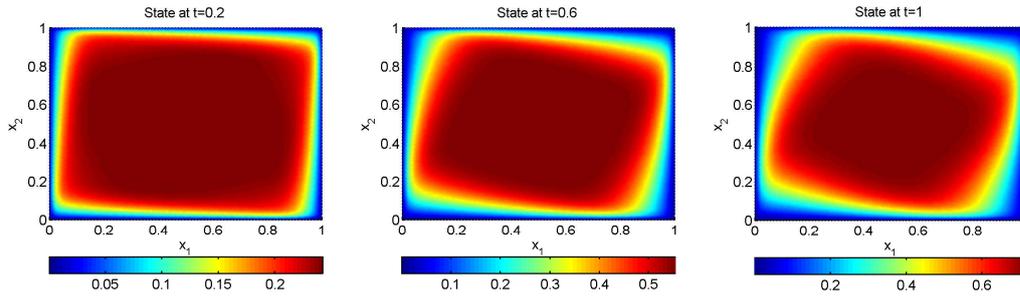


Figure 5.2: DCR Eqn: Numerical solution of the state at $t = 0.2, 0.6, 1$, respectively

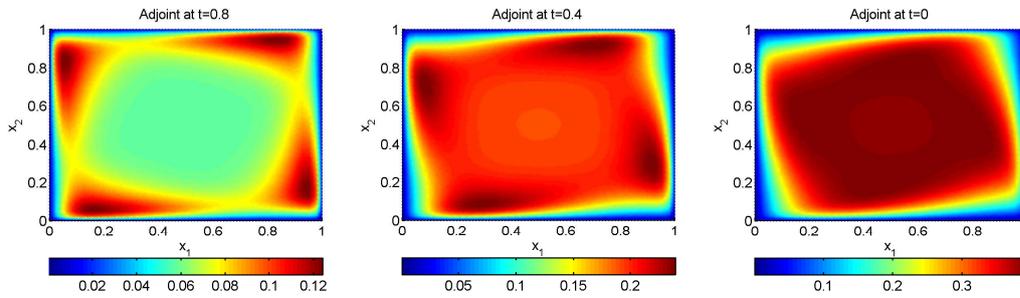


Figure 5.3: DCR Eqn: Numerical solution of the adjoint at $t = 0.8, 0.4, 0$, respectively

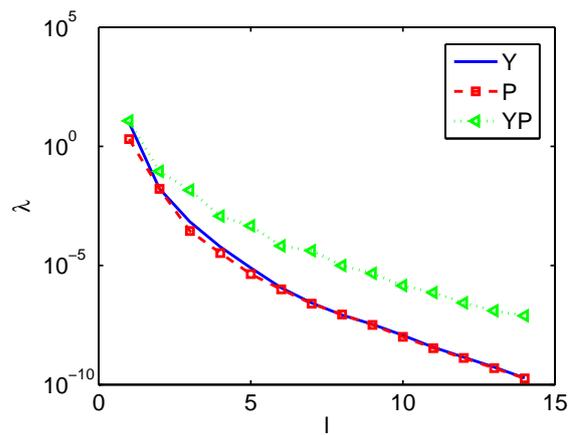


Figure 5.4: DCR Eqn: Eigenvalues for snapshot ensemble Y, P and $Y \cup P$

shot set P fails to predict the state, because adjoint and state have different characteristics as shown in Figures 5.2-5.3. For small values of POD basis functions, snapshot set Y gives the best results. However, as we increase the number of basis functions, the results obtained through the snapshot set Y and $Y \cup P$ become almost the same. Although the snapshot set $Y \cup P$ contains information about the adjoint, the best results are achieved through this basis. It is because the control, which acts on the right-hand side of the state equation as a forcing term, is approximated well using this snapshot set.

We proceed with the solutions of the adjoint and the control. Their results are almost the same because of the Tikhonov regularization parameter $\alpha = 1$ and the optimality condition (2.6c), as expected. Snapshot set Y cannot predict the adjoint and the control well, because their numerical solutions are totally different from each other. Although snapshot set P reveals the properties of the adjoint and the control up to four POD basis functions, snapshot set $Y \cup P$ leads to the best results as we increase the number of POD basis functions. Since the state equation is solved accurately using the snapshot set $Y \cup P$, this accuracy reflected not only to the adjoint but also to the control.

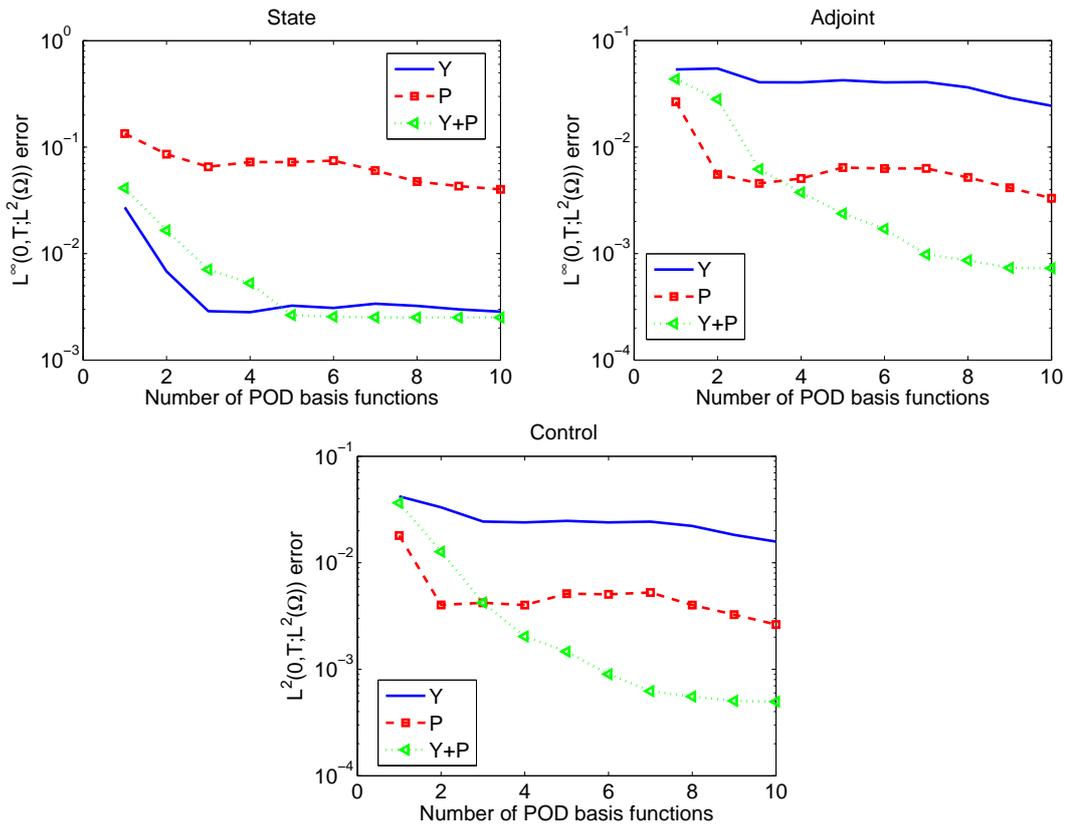


Figure 5.5: DCR Eqn: Error versus number of POD basis functions for state (*top-left*), adjoint (*top-right*) and control (*bottom*)

We continue to discuss the rate of convergence associated to the reduced-order model. We make a comparison between the numerical results and the theoretical convergence

rate with respect to $\Lambda_\epsilon = \sqrt{\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2}$ given in Theorem 5.14.

In Figure 5.6, we present the error in the state, the adjoint and the control with respect to the POD truncation error Λ_ϵ . For the state and the adjoint ensembles, the errors associated to 1, the first 2, 3 and 4 POD basis functions are plotted using the linear regression. For the set $Y \cup P$, the first 2, 4, 6, and 7 POD basis functions are chosen in order to equate Λ_ϵ in the x-axis as much as possible. This difference in the adjoint and in the control is because of the use of different norms, namely $L^\infty(0, 1; L^2(\Omega))$ -norm for the adjoint and $L^2(0, 1; L^2(\Omega))$ -norm for the control.

The highest convergence rate for the state is achieved through the snapshot set Y , namely around 0.63, which is also the case in the Figure 5.5. Snapshot ensemble $Y \cup P$ leads to the order 0.58. Dynamics of the state solution cannot be predicted using the adjoint information at all. The best rate associated to the snapshot set Y is less than the theoretical estimate, which means that the temporal and spatial terms in the estimate (5.67) are more dominant than Λ_ϵ and the error decays slower than the predicted rate.

For the adjoint, theoretical convergence rate is achieved using the snapshot set $Y \cup P$. For the control, the best order, namely 0.89, is obtained using the same snapshot set, too. This results are compatible with the Figures 5.5. We note that the snapshot set Y leads to poor results, because the error decays very slowly.

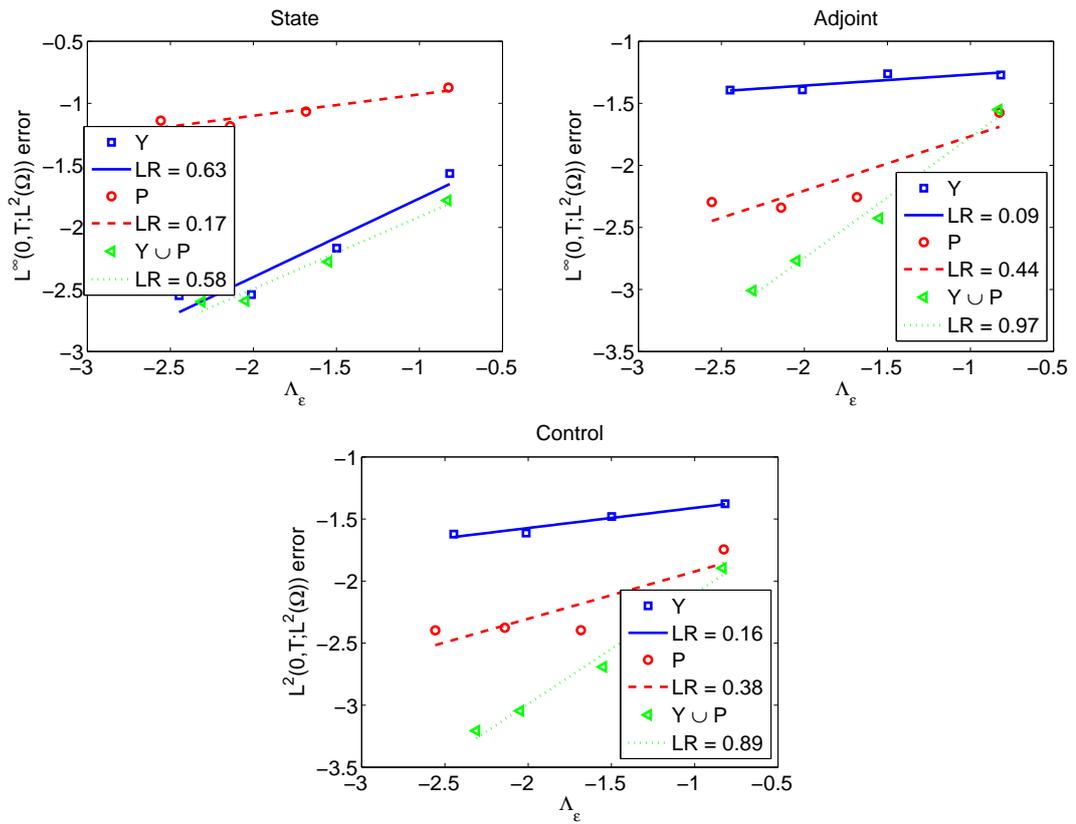


Figure 5.6: DCR Eqn: Order convergence for state (*top-left*), adjoint (*top-right*) and control (*bottom*) $h = k = 1/80$

5.4 Discrete Empirical Interpolation Method

In this section, we discuss the reduced-order model for optimal control of Burgers equation. Then, discrete empirical interpolation method (DEIM), which is used to increase the efficiency of POD when applied to nonlinear problems, will be explained [27].

We proceed by projecting the systems given in Sec.4.6 onto the low-dimensional space for Burgers equation. We note that the initial conditions for state equation on each time interval I_m must be projected on to the low-dimensional space as follows:

$$Y_m^{l,0} = \Psi^T (y_\delta)_{m-1}^- \quad \text{if } m \geq 2 \quad \text{or} \quad (y_\delta)_{m-1}^{l,-} = \Psi^T y_0^- \quad \text{if } m = 0.$$

Then, for dG(1)-method, the reduced-order system associated to state equation, which is the linearised using Newton's method, is written as

$$\begin{pmatrix} \frac{3}{4}\mathbf{M}^l + \frac{k_m}{2}(\mathbf{A}^l + k\mathbf{N}_y^l(Y_m^{1,\alpha})) & \frac{1}{4}\mathbf{M}^l \\ -\frac{9}{4}\mathbf{M}^l & \frac{5}{4}\mathbf{M}^l + \frac{k_m}{2}(\mathbf{A}^l + k\mathbf{N}_y^l(Y_m^{2,\alpha})) \end{pmatrix} \begin{pmatrix} \delta Y_m^{l,1} \\ \delta Y_m^{l,2} \end{pmatrix} \\ = \begin{pmatrix} \left(\frac{3}{4}\mathbf{M}^l + \frac{k}{2}\mathbf{A}^l \right) Y_m^{l,1} + \frac{k}{2}\mathbf{N}^l((Y_m^1)^2) + \frac{1}{4}\mathbf{M}^l Y_m^{l,2} - \mathbf{M}^l Y_m^{l,0} - \frac{k}{2}\mathbf{F}_h^l(t_{m,1}) - \frac{k}{4}\mathbf{M}^l (U_m^{l,0} + U_m^{l,1}) \\ -\frac{9}{4}\mathbf{M}^l Y_m^{l,1} + \left(\frac{5}{4}\mathbf{M}^l + \frac{k}{2}\mathbf{A}^l \right) Y_m^{l,2} + \frac{k}{2}\mathbf{N}^l((Y_m^2)^2) + \mathbf{M}^l Y_m^{l,0} - \frac{k}{2}\mathbf{F}_h^l(t_{m,2}) - \frac{k}{4}\mathbf{M}^l (3U_m^{l,1} - U_m^{l,0}) \end{pmatrix}. \quad (5.69)$$

This system is solved up to a given tolerance for $\delta Y_m^{l,1}, \delta Y_m^{l,2}$ on each time interval $I_m = (t_{m-1}, t_m]$. Then, the solution is updated as

$$Y_m^{l,1,\alpha+1} = Y_m^{l,1,\alpha} - \delta Y_m^{l,1}$$

and

$$Y_m^{l,2,\alpha+1} = Y_m^{l,2,\alpha} - \delta Y_m^{l,2}.$$

Then, we denote $Y_m^{l,1} := Y_m^{l,1,\alpha+1}$ and $Y_m^{l,2} := Y_m^{l,2,\alpha+1}$.

To derive the reduced-order system associated to adjoint equation, the initial conditions for the adjoint equation on each time interval I_m must be projected onto the low-dimensional space as:

$$\lambda_m^{l,2} = \Psi^T (\lambda_\delta)_m^+ \quad \text{if } m \leq N_T \quad \text{or} \quad \lambda_m^{l,2} = 0 \quad \text{if } m = N_T.$$

On each time interval $I_m = (t_{m-1}, t_m]$, we solve the following linear system for $\lambda_m^{l,1}, \lambda_m^{l,0}$

$$\begin{pmatrix} \frac{3}{4}\mathbf{M}^l + \frac{k_m}{2}(\mathbf{A}^l - \mathbf{N}^{\lambda,l}(\frac{Y_m^1 + Y_m^2}{2})) & \frac{1}{4}\mathbf{M}^l \\ -\frac{9}{4}\mathbf{M}^l & \frac{5}{4}\mathbf{M}^l + \frac{k_m}{2}(\mathbf{A}^l - \mathbf{N}^{\lambda,l}(3\frac{Y_m^1 - Y_m^2}{2})) \end{pmatrix} \begin{pmatrix} \lambda_m^{l,1} \\ \lambda_m^{l,0} \end{pmatrix} \\ = \begin{pmatrix} \mathbf{M}^l \lambda_m^2 + \frac{k_m}{2}\mathbf{Y}_h^{l,d}(t_{m,1}) - \frac{k_m}{4}\mathbf{M}^l (Y_m^{l,1} + Y_m^{l,2}) \\ -\mathbf{M}^l \lambda_m^2 + \frac{k_m}{2}\mathbf{Y}_h^{l,d}(t_{m,0}) - \frac{k_m}{4}\mathbf{M}^l (3Y_m^{l,1} - Y_m^{l,2}) \end{pmatrix}. \quad (5.70)$$

Similar to the cGP(q+1) method, the discrete adjoint at t_{m-1} is given as $\lambda_{l,m-1} = \lambda_m^{l,0}$.

Dimension of the reduced-order model obtained from the POD method is several times smaller than the full-order problem. However, for nonlinear problems, computational complexity of the terms such as N^l and N_y^l in (5.69) still depend on the degrees of freedom. To overcome this drawback of POD and increase the efficiency of this method, DEIM proposed in [27] can be used.

The idea behind this method is to project the nonlinear term (5.71)

$$\mathbf{N}^l = \Psi^T \mathbf{N}, \quad \mathbf{N}_y^l = \Psi^T \mathbf{N}_y \Psi, \quad (5.71)$$

onto a subspace of dimension $l_D \approx l \ll \text{dof}$ that approximates the space generated by the nonlinear function. For nonlinear structural analysis, unassembled DEIM is proposed in order to decrease the number of function calls to calculate the contribution of each finite element associated to the DEIM points in [121]. Now, we explain the method briefly and discuss the details arising due to the discontinuous Galerkin discretization following [27].

Remark 5.3. Since we use the mass matrix \mathbf{M} as a weighting matrix in the computation of the POD basis (see, Defn 5.19), we follow the same idea in order to compute a new basis for the nonlinear term \mathbf{N} before applying the DEIM. Therefore, we proceed the following vector and the matrices:

$$\hat{\mathbf{N}}^l = \Psi^T \mathbf{M} \mathbf{N}, \quad \hat{\mathbf{N}}_y^l = \Psi^T \mathbf{M} \mathbf{N}_y \Psi.$$

Firstly, let us consider the nonlinear term and its Jacobian in the reduced-order system

$$\hat{\mathbf{N}}^l = \underbrace{\Psi^T \mathbf{M}}_{l \times \text{dof}} \underbrace{\hat{\mathbf{N}}^l(\Psi \mathbf{y}_{l,m})}_{\text{dof} \times 1}, \quad \hat{\mathbf{N}}_y^l = \underbrace{\Psi^T \mathbf{M}}_{l \times \text{dof}} \underbrace{\frac{\partial \mathbf{N}(\Psi \mathbf{y}_{l,m})}{\partial y}}_{\text{dof} \times \text{dof}} \underbrace{\Psi}_{\text{dof} \times l}. \quad (5.72)$$

For ease of notation, we denote the nonlinear term by f_N . The aim is to find an approximation to f_N by projecting it onto a subspace spanned by $\{\psi_1^D, \dots, \psi_{l_D}^D\} \subset \mathbb{R}^{\text{dof}}$ as follows

$$f_N \approx \Psi^D \mathbf{c},$$

where $\{\Psi_1^D, \dots, \Psi_{l_D}^D\} \in \mathbb{R}^{\text{dof}}$ is the space of the first l_D POD basis functions of the space spanned by the nonlinear snapshots $\{\mathbf{N}(\mathbf{y}_{\delta,1}), \dots, \mathbf{N}(\mathbf{y}_{\delta,N})\}$ associated to the largest singular values and \mathbf{c} is the corresponding coefficient vector. To find \mathbf{c} , we choose l_D distinct rows from the overdetermined system $f_N \approx \Psi^D \mathbf{c}$. Let us define a matrix $P = [e_{\varphi_1}, \dots, e_{\varphi_{l_D}}] \in \mathbb{R}^{\text{dof} \times l_D}$, where $e_{\varphi_i} = [0, \dots, 0, 1, 0, \dots, 1]^T \in \mathbb{R}^{\text{dof}}$ is the φ_i -th column of the identity matrix I_{dof} . If $P^T \Psi^D$ is singular, then \mathbf{c} can be determined uniquely through the relation

$$P^T f_N = (P^T \Psi^D) \mathbf{c} \implies f_N \approx \Psi^D \mathbf{c} = \Psi^D (P^T \Psi^D)^{-1} P^T f_N. \quad (5.73)$$

DEIM algorithm can be found in Algorithm 5.2.

Algorithm 5.2 Discrete Empirical Interpolation Method

```

1: procedure DEIM( $l_D, \Psi^D$ )
2:    $[\rho, \wp_1] = \max\{|\Psi_1^D|\}$ 
3:    $\Psi^D = [\Psi_1^D], P = [e_{\wp_1}], \vec{\wp} = [\wp_1]$ 
4:   for  $i = 2$  to  $l^D$  do
5:     Solve  $(P^T \Psi^D) \mathbf{c} = P^T \Psi_i^D$  for  $\mathbf{c}$ 
6:      $\mathbf{r} = \Psi_i^D - \Psi_i^D \mathbf{c}$ 
7:      $[\rho, \wp_i] = \max\{|\mathbf{r}|\}$ 
8:      $\Psi^D \leftarrow [\Psi^D, \Psi_i^D], P \leftarrow [P, e_{\wp_i}], \vec{\wp} \leftarrow \begin{bmatrix} \vec{\wp} \\ \wp_i \end{bmatrix}$ 
9:   end for
10:  return  $\vec{\wp} = [\wp_1, \dots, \wp_{l_D}]^T \in \mathbb{R}^{l_D}$ 
11: end procedure

```

DEIM approximates the projected terms (5.72) as

$$\hat{\mathbf{N}}^l \approx \underbrace{\Psi^T \mathbf{M} \Psi^D (P^T \Psi^D)^{-1} P^T \mathbf{N}(\Psi \mathbf{y}_{l,m})}_{:= \mathcal{Z}} = \underbrace{\mathcal{Z}}_{l \times l_D} \underbrace{P^T \mathbf{N}(\Psi \mathbf{y}_{l,m})}_{l_D \text{ evaluations}}, \quad (5.74a)$$

$$\hat{\mathbf{N}}_y^l(\mathbf{y}_{l,m}) \approx \underbrace{\mathcal{Z}}_{l \times l_D} \underbrace{\frac{\partial P^T \mathbf{N}(\Psi \mathbf{y}_{l,m})}{\partial \mathbf{y}}}_{l_D \times \text{dof, sparse}} \underbrace{\Psi}_{\text{dof} \times l}. \quad (5.74b)$$

However, the computation of the terms like $\Psi \mathbf{y}_{l,m}$ still depends on the degrees of freedom. Therefore, only the necessary terms of $\mathbf{y}_{l,m}$ in the matrix multiplication $P^T \mathbf{N}(\Psi \mathbf{y}_{l,m})$ are selected so that the dependency of the reduced-order model on dof is eliminated. A similar argument can also be found in [70, Sec. 3].

Remark 5.4. In discontinuous Galerkin framework, DEIM indices are connected with the right $y(x_n^+)$ or left approximations $y(x_n^-)$ instead of the value of the functions $y(x_n)$ at a node $x = x_n$ as in the continuous finite element method. In particular for linear finite element approximation, let the matrix of the indices and the nonlinear term be given as $P = [e_3, e_1, e_{15}, e_{16}, e_{\text{dof}/2}] \in \mathbb{R}^{\text{dof} \times 5}$ and

$$f_N = [f_1^+, f_2^-, f_2^+, \dots, f_{\text{dof}/2}^-, f_{\text{dof}/2}^+, f_{\text{dof}/2+1}^-]^T$$

, for an approximation of dof/2 subintervals. Then, $P^T f_N = [f_2^+, f_1^+, f_8^+, f_9^-, f_{\text{dof}/2}^-]^T$. Therefore, the values of $\mathbf{y}_{l,m}$ on the 2nd, 1st, 8th and (dof/2 - 1)th elements are required, instead of dof/2 elements which is the case without using DEIM.

5.5 A Priori Error Estimates for Optimal Control of Burgers Equation

In this section, we derive a priori error estimates for reduced-order solution of Burgers equation based on space-time DG and POD method.

5.5.1 Auxiliary Results

We make the following assumptions using the study [53, Sec.4.2]:

$$|n_h^s(y_\delta, v) - n_h^s(y_\delta^l, v)| \leq \frac{\epsilon}{8} \|v\|_{DG}^2 + C \left(\|y_\delta^l - \pi^l y_\delta\|^2 + \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right). \quad (5.75)$$

Therefore, we obtain the following estimate

$$\begin{aligned} & |n_h^a(\lambda_\delta, y_\delta, \frac{1}{\alpha}\phi) - n_h^a(\lambda_\delta^l, y_\delta^l, \frac{1}{\alpha}\phi)| \\ & \leq \frac{\epsilon}{4\alpha} \|\lambda_\delta^l - \pi^l \lambda_\delta\|_{DG}^2 + C(\|y_\delta - \pi^l y_\delta\|^2 + (1 + \frac{2}{\epsilon})\|\phi\|^2) \\ & \leq \frac{\epsilon}{4\alpha} \|\lambda_\delta^l - \pi^l \lambda_\delta\|_{DG}^2 + C \left((1 + \frac{2}{\epsilon})\|\phi\|^2 + h^2 \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right). \end{aligned} \quad (5.76)$$

5.5.2 Main Result

Theorem 5.16. *Let $(y_\delta, \lambda_\delta)$ and $y_\delta^l, \lambda_\delta^l$ be the solutions of (4.12) and the associated reduced-order solution, respectively. Then, there is a constant C independent of h, k and l such that the following inequalities hold:*

$$\begin{aligned} & \|y_\delta - y_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda_\delta - \lambda_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} + \|u_\delta - u_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} \\ & \leq C \left(\sqrt{\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2} \right). \end{aligned} \quad (5.77)$$

Proof. Firstly, we write the error in the state and the adjoint as

$$\begin{aligned} e &= y_\delta - y_\delta^l = (\pi^l y_\delta - y_\delta^l) + (y_\delta - \pi^l y_\delta) = \xi + \eta, \\ r &= \lambda_\delta - \lambda_\delta^l = (\pi^l \lambda_\delta - \lambda_\delta^l) + (\lambda_\delta - \pi^l \lambda_\delta) = \varphi + \phi, \end{aligned}$$

where π^l denotes the interpolation operator which is explained in Sec.5.3.1.

The local error equation for the state equation can be written as follows:

$$\begin{aligned} & \int_{I_m} ((\partial_t \xi, v) + a_h^s(\xi, v)) dt + ([\xi]_{m-1}, v_+^{m-1}) \\ & = - \left(\int_{I_m} (\partial_t \eta, v) dt + ([\eta]_{m-1}, v_+^{m-1}) \right) - \int_{I_m} a_h^s(\eta, v) dt \\ & + \left(\int_{I_n} n_h^s(y_\delta, v) dt - \int_{I_n} n_h^s(y_\delta^l, v) dt \right) - \frac{1}{\alpha} \int_{I_m} (\lambda_\delta - \lambda_\delta^l, v) dt, \quad \forall v \in V_{h,l}^{k,q}. \end{aligned} \quad (5.78)$$

Step I We choose $v = 2\xi$ in (5.78). Then, we use coercivity of DG-bilinear form (4.27), the equality (4.32) and (4.47). We proceed with the estimates (5.75) and simplify the common terms. Then, we sum from $m = 1$ to n to derive the following inequality

$$\begin{aligned} & \|\xi_-^m\|^2 + \frac{\epsilon}{2} \sum_{m=1}^n \int_{I_m} \|\xi\|_{DG}^2 dt & (5.79) \\ & \leq C \left(\|\xi_0^{l,-}\|^2 + \sum_{m=1}^n \frac{C}{\epsilon} k_m \|\xi_-^{m-1}\|^2 - \frac{2}{\alpha} \sum_{m=1}^n \int_{I_m} (\lambda_\delta - \lambda_\delta^l, \xi) dt \right) \\ & + \mathcal{O} \left(\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right). \end{aligned}$$

By applying the discrete Gronwall's lemma, we obtain the desired result (5.80a).

We proceed with the local error equation for the reduced-order adjoint equation. We proceed similarly and use the properties of the interpolation operator in Sec.5.46, we obtain the following result:

$$\begin{aligned} & \|e^n\|^2 + \frac{\epsilon}{2} \int_0^{t_n} \|\xi\|_{DG}^2 dt & (5.80a) \\ & \leq C \left(\|e_0^{l,-}\|^2 - \frac{2}{\alpha} \sum_{m=1}^n \int_{I_m} (\lambda_\delta - \lambda_\delta^l, \xi) dt \right) + \mathcal{O} \left(\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right), \end{aligned}$$

$$\begin{aligned} & \|r^n\|^2 + \frac{\epsilon}{2} \int_0^{t_n} \|\varphi\|_{DG}^2 dt & (5.80b) \\ & \leq C \left(\frac{2}{\alpha} \sum_{m=1}^n \int_{I_m} (y_\delta - y_\delta^l, \varphi) dt \right) + \mathcal{O} \left(\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right). \end{aligned}$$

Step II Now, we proceed with the estimates at arbitrary time steps using the previous estimate and the discrete characteristic function (4.3.2). The proof is similar to the case in Lemma 5.16. To do this, for any fixed $t \in [t^{n-1}, t^n)$, we set $v_\delta = \tilde{e}_h$ in (5.78).

$$\begin{aligned} & \sup_{t \in I_m} \|\xi(t)\|^2 - \|\xi_-^{m-1}\|^2 & (5.81) \\ & \leq \frac{\epsilon}{2} \int_{I_m} \|\tilde{\xi}\|_{DG}^2 dt + \frac{C}{\epsilon} \int_{I_m} \|y_\delta^l - \pi^l y_\delta\|^2 dt - \frac{2}{\alpha} \int_{I_m} (\lambda_\delta - \lambda_\delta^l, \tilde{\xi}) dt \\ & + \mathcal{O} \left(\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2 \right). \end{aligned}$$

For the adjoint equation, for any fixed $t \in (t^{N-1}, t^N]$, we set $v_\delta = \frac{1}{\alpha} \tilde{r}_h$ in the local error equation of the adjoint and proceed similarly.

$$\begin{aligned}
& \sup_{t \in I_m} \|\varphi(t)\|^2 - \|\varphi^m\|^2 & (5.82) \\
& \leq \frac{\epsilon}{2} \int_{I_m} \|\tilde{\varphi}\|_{DG}^2 dt + \frac{C}{\epsilon} \int_{I_m} \|y_\delta - \pi^l y_\delta\|^2 dt + \frac{2}{\alpha} \int_{I_m} (y_\delta - y_\delta^l, \tilde{\xi}) dt \\
& + \mathcal{O}\left(\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2\right).
\end{aligned}$$

In order to eliminate the terms $\lambda_\delta - \lambda_\delta^l$ and $y - y_\delta^l$ in the estimates 5.81 and 5.82, we adapt the approach in [29] and write the coupling terms as follows:

$$\begin{aligned}
& -\frac{1}{\alpha} \int_{I_m} (r, \tilde{\xi}) dt + \frac{1}{\alpha} \int_{I_m} (e, \tilde{\varphi}) dt & (5.83) \\
& \leq \mathcal{O}(h^2 \sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2) + C_D \int_{I_m} \left(\frac{\epsilon}{8} \|\xi\|_{DG}^2 + \frac{\epsilon}{8\alpha} \|\varphi\|_{DG}^2 \right) dt.
\end{aligned}$$

We add the resulting inequalities for the state and the adjoint equation and sum from $m = 1$ to n . Then, we apply discrete Gronwall's inequality to arrive at the estimate (5.77). \square

Using the triangle inequality and the estimate in Lemma 4.11-Theorem 5.16, we state the main estimate for the error between the exact and the reduced-order solution:

Theorem 5.17. *Suppose that (y, λ, u) and $(y_\delta, \lambda_\delta, u_\delta)$ are the solutions of (2.8) and (4.12), respectively. We assume that all conditions of Lemmas 5.16 and 4.11 are satisfied. Then, there exists a constant C independent of h and k such that*

$$\begin{aligned}
& \|y - y_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda - \lambda_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} + \|u - u_\delta^l\|_{L^\infty(0,T;L^2(\Omega))} \\
& \leq C \left(h^{p+1} \|y_0^l\| + \mathcal{O}(h^p, k^{q+1}, \sqrt{\sum_{j=l+1}^d \lambda_j \|\psi_j\|_{DG}^2}) \right). & (5.84)
\end{aligned}$$

5.5.3 Numerical Results

In this section, we present some numerical results to investigate the experimental order of convergence associated to the reduced-order model. Full-order problem is solved using piecewise linear discontinuous finite elements on a uniform mesh with $h = k = 1/200$ leading to linear systems of size 400. Three different snapshot sets for W are used to generate the POD basis functions, namely the state Y , the adjoint P and the combination of them $Y \cup P$, as in [72]. POD basis is computed using the singular value decomposition (SVD), because it is more stable than the eigenvalue decomposition, i.e.

the singular values decay to machine precision, whereas the eigenvalues stagnate above [118]. We measure the error in the state, in the adjoint and control approximation in terms of $L^\infty(0, 1; L^2(\Omega))$ -norm.

We consider the optimal control problem in [82], which has been discussed in Sec. 4.6.1, with

$$Q = (0, 1] \times \Omega, \quad \Omega = (0, 1), \quad \epsilon = 10^{-2}, \quad \text{and} \quad \alpha = 0.05.$$

We take the source function $f = 0$, the desired state y_d and the initial condition y_0 are defined as

$$y_d(x, t) = \begin{cases} 1 & \text{in } (0, 1/2], \\ 0 & \text{otherwise.} \end{cases}$$

In Figure 5.7, we present the decay of the eigenvalues for three different snapshot sets. We observe that the eigenvalues decrease rapidly showing that POD can be successfully applied.

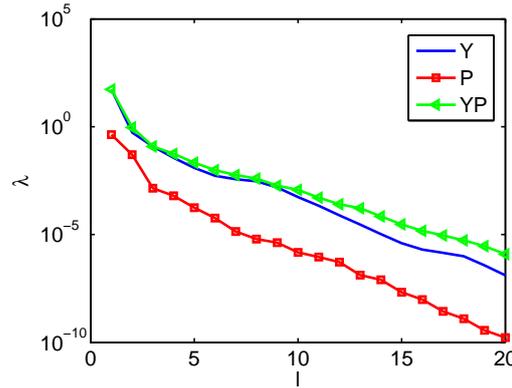


Figure 5.7: Burgers Eqn: Eigenvalues for snapshot ensemble Y , P and $Y \cup P$

In Figure 5.8, we present the error in the state, the adjoint and the control with respect to the number of POD basis functions.

First of all, the error in the state is decreasing for three different snapshot ensemble as we increase the number of POD basis functions. The use of adjoint information leads to poor results of the state solution. For small number of POD basis functions, the snapshot sets Y and $Y \cup P$ result in almost the same results. As we increase the number of POD basis functions, more accurate results are obtained using the snapshot set $Y \cup P$ and the error is decreased up to 10^{-4} with 15 POD basis functions.

The error in the adjoint equation is decreased up to 10^{-2} using the basis computed with the snapshot ensemble Y . On the other hand, the snapshot set P gives the best results until 10 POD basis functions are used. Then, the error stops around 10^{-3} . For the snapshot set $Y \cup P$, the smallest error is achieved.

We observe that the graphs of the control and the adjoint have the same pattern due to the optimality condition (2.8c). Since the regularization parameter is fixed as $\alpha = 0.05$, the error in the adjoint is α -times smaller than the error in the control.

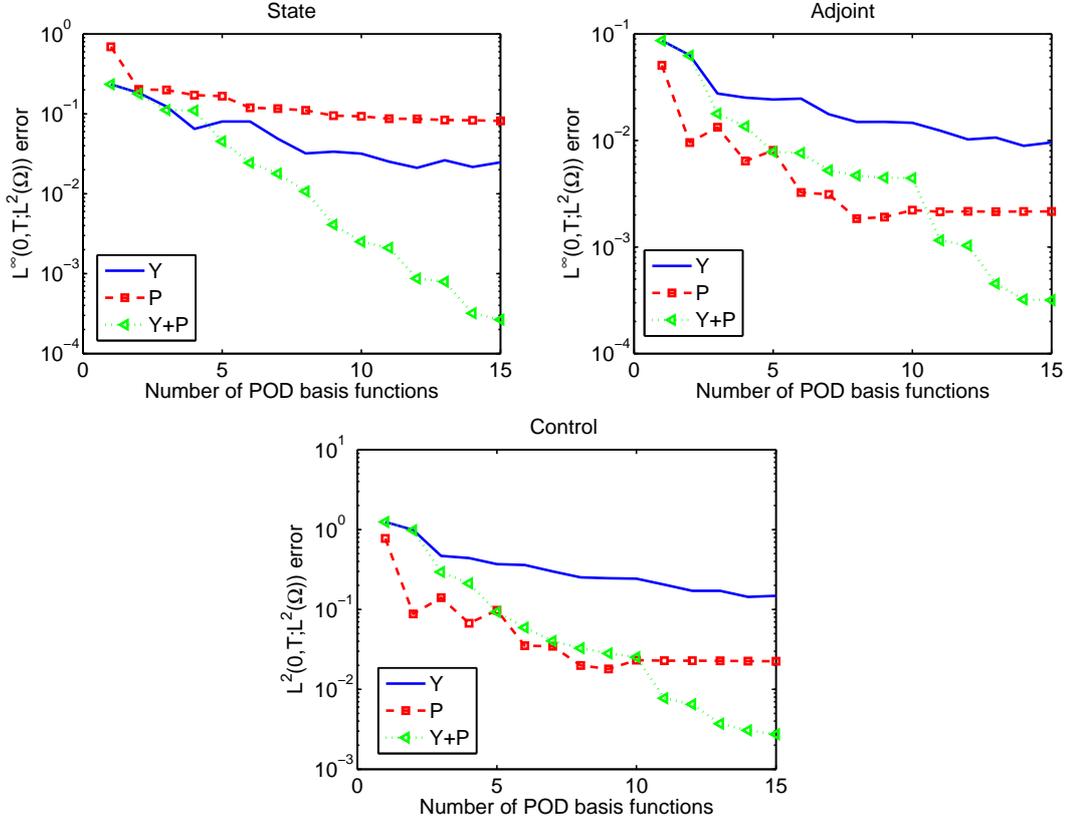


Figure 5.8: Burgers Eqn: Error versus number of POD basis functions for state (*top-left*), adjoint (*top-right*) and control (*bottom*)

In Figure 5.9, we plot the error in the state, the adjoint and the control with respect to the POD truncation error. For the state ensemble Y , the first 7, 9, 10 and 11 POD basis functions are used; while for the adjoint ensemble P , the first 2, 4, 6 and 7 POD basis functions are decided. For the snapshot set $Y \cup P$, the first 8, 11, 13 and 14 POD basis functions are chosen for the state solution; the first 9, 10, 13 and 14 POD basis functions are taken for the adjoint solution in order to equate the POD truncation error in the x-axis. Since the error in the state obtained using the snapshot ensemble Y and P are parallel in Figure 5.8, the rate of convergence associated to this sets are the same, namely 0.18. Using the snapshot ensemble $Y \cup P$ approximates the state well and the theoretically estimated convergence rate in Thm 5.16 is achieved.

For the state and the control, the state snapshot Y is unsuccessful to predict both of the adjoint and the control. Using the snapshot set P , the order of convergence is increased up to 0.36. On the other hand, the snapshot set $Y \cup P$ gives the best results and the theoretical convergence rate is reached for the adjoint and the control.

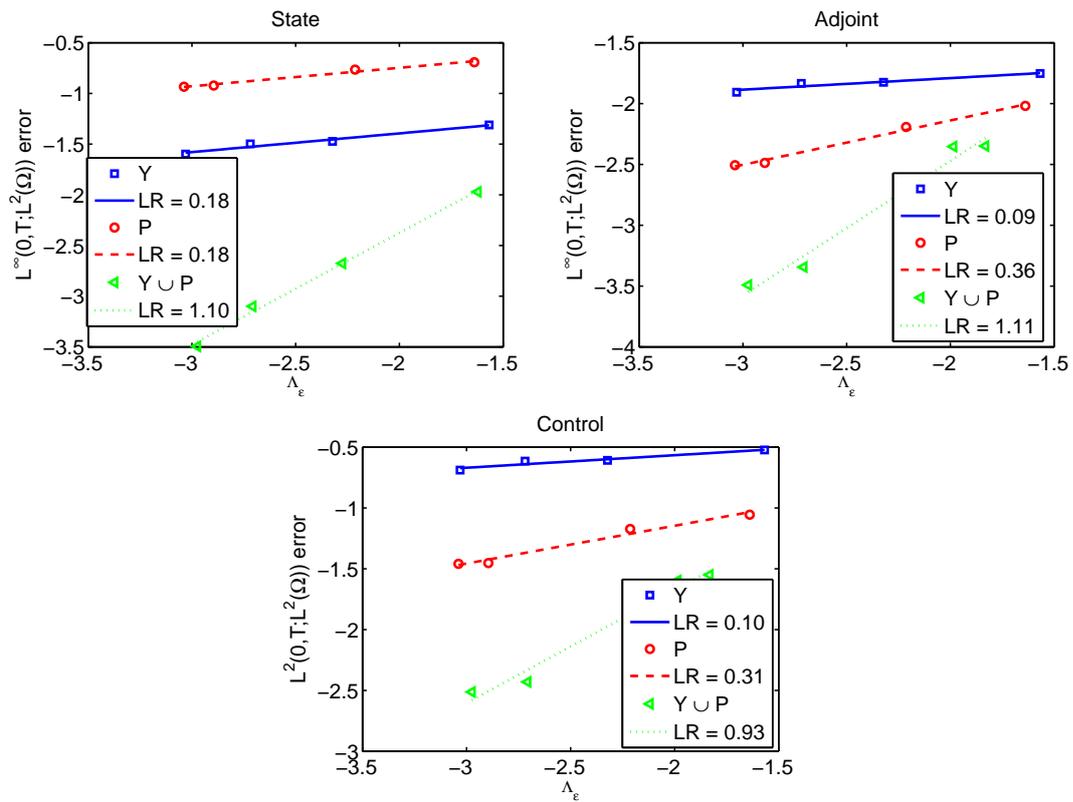


Figure 5.9: Burgers Eqn: Order convergence for state (*top-left*), adjoint (*top-right*) and control (*bottom*)

CHAPTER 6

LOCAL IMPROVEMENTS TO REDUCED-ORDER SOLUTION

POD basis is computed using the snapshots of a particular problem which is interpreted by a mathematical model and data. Because there is a link between the data and the snapshots, some perturbation in the data may lead to larger changes in the snapshots depending on the problem at hand. This leads the nominal/baseline POD basis, which depends on the nominal/baseline parameters, not to approximate the perturbed problem accurately. In such cases, one has to solve the full problem for each parameter in the data set again and regenerate the POD basis. This approach is expensive especially for nonlinear problems or optimal control problems which requires the solution of a set of differential equations.

In the literature, some alternatives are derived for a single PDE. For example, subspace angle interpolation method motivated from [20] is applied to linearised Euler equations in [128]. For a robust method, the interpolation method based on Grassman manifold and its tangent space at a point is employed in [7, 8].

Another choice is to use the sensitivities of the trajectories which require the solution of the sensitivity equations. They can be obtained by applying continuous sensitivity equation (CSE) method or finite difference (FD) approximation. Because sensitivity equations are always linear, the former method is especially preferable for nonlinear problems. The latter one requires the computation of the full problem at least one more time, so it is expensive for nonlinear case. This approach, for example, is used in [64] where the sensitivity of Navier-Stokes equation with respect to Reynolds number is computed. Sensitivity of the solution is used to compute the POD basis sensitivities and then the nominal basis is extrapolated or expanded. In [62], the sensitivity with respect to the angle of inclination of the incoming free-stream flow is utilized. In [63], Burgers equation is solved on a domain whose shape depends on the parameter α and the sensitivity with respect to α is derived and used to enrich the low-dimensional space. The sensitivity of the shallow water equations [143] based on Manning roughness coefficient is considered and utilized using POD. However, to the best of our knowledge, our work [4] is the first study combining the POD sensitivities and the PDE-constrained optimization. For a comparison of other types of reduced-order bases, we refer the reader to the study [2].

In this study, motivated by the papers [62, 63, 64], POD sensitivities are used to enrich the low-dimensional subspace for a wider range of parameters and the quantity

of interest is the diffusion term ϵ , the convective field β and the reaction term r for DCR equation. For Burgers equation, sensitivity with respect to the diffusion term ϵ is computed. We generate two new bases, i.e. extrapolated POD (ExtPOD) and expanded POD (ExpPOD) and compare these bases in terms of advantages and discuss the main drawbacks of them.

In this chapter, we explain how to compute the POD sensitivities in Sec. 6.1 and derive two new bases. Then, we derive sensitivity equations of DCR equation in Sec. 6.2. We present the numerical results in Sec. 6.2.4. Then, we proceed with the sensitivity equations associated to Burgers equation in Sec. 6.3 and present the numerical results in Sec. 6.3.1 and compare two new bases and three different snapshot ensembles.

6.1 Proper Orthogonal Decomposition Sensitivities

Sensitivity of a term is defined as the derivative of that term with respect to a quantity of interest μ . We assume that the state, the adjoint and the control are functions depending on space, time and the quantity of interest μ ,

$$y = y(x, t, \mu), \quad \lambda = \lambda(x, t, \mu), \quad u = u(x, t, \mu).$$

We define the sensitivities as

$$s_y = \frac{\partial y}{\partial \mu}, \quad s_\lambda = \frac{\partial \lambda}{\partial \mu}, \quad s_u = \frac{\partial u}{\partial \mu}.$$

After finding the sensitivities of the state s_y and the adjoint s_λ , POD sensitivities are obtained. To do so, we treat each POD mode as a function of both space and the parameter, i.e. $\psi = \psi(x, \mu)$. Then, we differentiate the relation (5.20) with respect to μ and solve the resulting equation for Ψ_μ . We proceed with the relation (5.12) to derive the POD basis sensitivities ψ_μ

$$(\psi_j)_\mu = \sum_{i=1}^m (\Psi_{ij})_\mu \varphi_i(x), \quad j = 1, \dots, l.$$

We summarize the computation of POD basis sensitivities following [64, Sec. 3.2]. Computation of U_μ^l , which appears after differentiating (5.20) with respect to μ , is realised through the relation

$$U_\mu^l = (\tilde{W}V^l\Sigma^\dagger)_\mu = \tilde{W}_\mu V^l\Sigma^\dagger + \tilde{W}V_\mu^l\Sigma^\dagger + \tilde{W}V^l\Sigma_\mu^\dagger.$$

The term \tilde{W}_μ denotes the sensitivity of the snapshot matrix which is obtained by CSE or FD approximation. Computational details regarding the sensitivities will be explained in the next section. For the computation of V_μ^l and Σ_μ^\dagger , we consider the eigenvalue problem $BV^l = V^l\lambda^l$ with the l th column of V . Assume that B, V and $\Lambda = \text{diag}(\lambda^1, \dots, \lambda^k)$ are smooth with respect to μ . After differentiation, we arrive at

$$B_\mu V^k + BV_\mu^k = \lambda_\mu^k V^k + \lambda^k V_\mu^k \rightarrow (B - \lambda^k I)V_\mu^k = -(B_\mu - \lambda_\mu^k I)V^k. \quad (6.1)$$

V_μ^k is the solution of (6.1) only if $(B_\mu - \lambda_\mu^k I)V^k \in \text{Range}(B - \lambda^k I)$. Equivalently,

$$(V^l)^T (B_\mu - \lambda_\mu^l I) V^l = 0. \quad (6.2)$$

The orthonormal matrix V^l has already been computed via SVD. Then, the eigenvalue sensitivities are given by

$$\lambda_\mu^l = (V^l)^T B_\mu V^l.$$

Each term of Σ_μ^\dagger is computed due to the relation between the singular values σ_μ and the eigenvalues λ_μ , i.e. $\sigma_\mu^2 = \lambda_\mu$.

The equation (6.2) is solved in the least-squares sense and we denote one particular solution by s^l . The general solution to (6.2) is expressed as $s^l + \gamma V^l$ for $\gamma \in \mathbb{R}$ with a simple λ^l . In addition, we differentiate the normalization condition $V^l (V^l)^T = 1$ leading to $V_\mu^l (V^l)^T = 0$. Then, the sensitivity of V^l and γ are determined by

$$V_\mu^l = s^l - ((s^l)^T V^l) V^l, \quad \gamma = -(s^l)^T V^l.$$

In ExtPOD, POD basis depending on μ is written using the first-order Taylor expansion as follows

$$\psi(x, \mu) = \psi(x, \mu_0) + \Delta\mu \frac{\partial \psi}{\partial \mu}(x, \mu_0) + \mathcal{O}((\Delta\mu)^2).$$

The reduced-order solution is expressed as

$$w_{\delta,i} = \sum_{j=1}^l \mathbf{w}_{\delta,i}^j (\psi_j(x, \mu_0) + \Delta\mu (\psi_j(x, \mu_0))_\mu).$$

In ExpPOD, POD basis sensitivities are also added to the original POD basis as follows

$$[\psi_1, \dots, \psi_l, (\psi_1)_\mu, \dots, (\psi_l)_\mu]$$

and the reduced-order solution is written as

$$w_{\delta,i} = \sum_{j=1}^l \mathbf{w}_{\delta,i}^j \psi_j(x, \mu_0) + \sum_{j=l+1}^{2l} \mathbf{w}_{\delta,i}^j (\psi_{j-l}(x, \mu_0))_\mu,$$

where the dimension of the reduced basis is doubled.

In Alg. 6.1, we summarize the derivation of the reduced-order solution of the OCP using POD sensitivities step by step.

Algorithm 6.1 Reduced-order approximation for the OCP associated to the perturbed parameter $\mu = \mu_0 \pm \Delta\mu$ using POD basis sensitivities

- 1: **if** Snapshots of the state and the adjoint associated to the nominal/reference parameter $\mu = \mu_0$ are not given **then**
 - 2: Solve the full-order problem and keep the solutions of the state and the adjoint equation .
 - 3: **end if**
 - 4: Construct the snapshot matrix W (using the state $W := Y$; the adjoint $W := P$ or a combination of them $W := Y \cup P$.)
 - 5: Compute the POD basis $\psi(x, \mu)$ associated to the nominal parameter $\mu = \mu_0$ using the snapshot matrix W and denote it by BPOD.
 - 6: **if** The state equation is nonlinear. **then**
 - 7: Compute a new basis for the nonlinear term using DEIM.
 - 8: **end if**
 - 9: Compute the sensitivity of the snapshot matrix W_μ (using the state $W_\mu := Y_\mu$; the adjoint $W_\mu := P_\mu$ or a combination of them $W_\mu := Y_\mu \cup P_\mu$).
 - 10: Compute the POD basis sensitivities $\psi_\mu(x, \mu)$ using the matrix W_μ .
 - 11: Construct new POD bases, namely ExtPOD and ExpPOD.
 - 12: Solve the reduced-order model of the OCP associated to $\mu = \mu_0 \pm \Delta\mu$.
-

6.2 Sensitivity Equations for Optimal Control of Diffusion-Convection-Reaction Equation

6.2.1 Sensitivity with respect to the diffusion term ϵ

In this section, the parameter is fixed as $\mu = \epsilon$. Sensitivity equations are obtained by differentiating the continuous state (1.1b) and the adjoint equation associated to (2.7a) and the optimality condition (2.6c) with respect to ϵ . The subscript ϵ denotes the derivative with respect to ϵ . The corresponding optimality system with s_y , s_λ and s_u is written as follows,

$$\begin{aligned} (\partial_t s_y, v) + a(s_y, v) + (\nabla y, \nabla v) &= (f_\epsilon + s_u, v), \quad \forall v \in V, \\ s_y(x, 0) &= (y_0)_\epsilon, \end{aligned} \tag{6.3a}$$

$$\begin{aligned} -(\partial_t s_\lambda, \psi) + a(\psi, s_\lambda) + (\nabla \psi, \nabla \lambda) &= -(s_y - y_\epsilon^d, \psi), \quad \forall \psi \in V, \\ s_\lambda(x, T) &= 0, \end{aligned} \tag{6.3b}$$

$$\alpha s_u = s_\lambda. \tag{6.3c}$$

We note that the homogeneous Dirichlet boundary conditions are imposed to (6.3) after differentiating in the same way. The optimality system (6.3) is discretized using the same numerical method, i.e., space-time discontinuous Galerkin discretization, as for (2.6).

Sensitivity equations are always linear, so CSE would be especially promising for non-linear problems. On the other hand, FD approximation can also be used to find the sensitivities. It requires the evaluation of the OCP depending on the perturbed parameters. In particular, sensitivity of the state can be computed via the centred difference as follows:

$$s_y(\epsilon_0) \approx \frac{y(\epsilon_0 + \Delta\epsilon) - y(\epsilon_0 - \Delta\epsilon)}{2\Delta\epsilon}. \quad (6.4)$$

The increment $\Delta\epsilon$ is chosen sufficiently small for an accurate FD approximation and it is chosen sufficiently large for the difference between two nearby POD vectors to be larger than the discretization error by one order of magnitude [64].

6.2.2 Sensitivity with respect to the convective term β

The quantity of interest is the parameter $\mu = p$ in the convection field $\beta = [y - 0.5, -x + p]$. In Figures 6.1-6.2, we plot the convection field of the state and the adjoint equation for $p = 0.1, 0.5, 0.9$. Such changes lead to different convective field and they results in the changes in the perturbed problem.

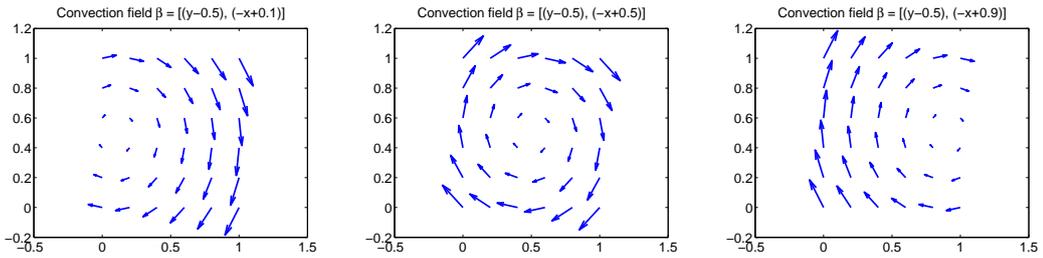


Figure 6.1: Convection field of the state $\beta = [y - 0.5, -x + p]$ with $p = 0.1, 0.5, 0.9$, respectively

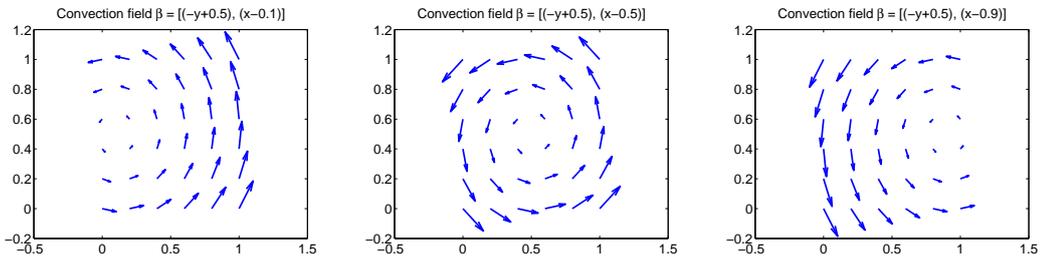


Figure 6.2: Convection field of the adjoint $\beta = [-y + 0.5, x - p]$ with $p = 0.1, 0.5, 0.9$, respectively

Sensitivity equations are obtained by differentiating the continuous state (1.1b) and the adjoint equation associated to (2.7a) and the optimality condition (2.6c) with respect to p . The subscript p denotes the derivative with respect to p . The corresponding

optimality system with s_y, s_λ and s_u is written as follows,

$$\begin{aligned} (\partial_t s_y, v) + a(s_y, v) + \left(\left[\frac{\partial \beta_1}{\partial p}, \frac{\partial \beta_2}{\partial p} \right] \cdot \nabla y, v \right) &= (f_p + s_u, v), \quad \forall v \in V, \\ s_y(x, 0) &= (y_0)_p, \end{aligned} \quad (6.5a)$$

$$\begin{aligned} -(\partial_t s_\lambda, \psi) + a(\psi, s_\lambda) - \left(\left[\frac{\partial \beta_1}{\partial p}, \frac{\partial \beta_2}{\partial p} \right] \cdot \nabla p, \psi \right) &= -(s_y - y_p^d, \psi), \quad \forall \psi \in V, \\ s_\lambda(x, T) &= 0, \end{aligned} \quad (6.5b)$$

$$\alpha s_u = s_\lambda. \quad (6.5c)$$

We note that the homogeneous Dirichlet boundary conditions are imposed to (6.5) after differentiating in the same way. The optimality system (6.5) is discretized using the same numerical method, i.e., space-time discontinuous Galerkin discretization, as for (2.6).

6.2.3 Sensitivity with respect to the reaction term r

In this section, the parameter is fixed as $\mu = r$. Sensitivity equations are obtained by differentiating the continuous state (1.1b) and the adjoint equation associated to (2.7a) and the optimality condition (2.6c) with respect to r . The subscript r denotes the derivative with respect to r . The corresponding optimality system with s_y, s_λ and s_u is written as follows,

$$\begin{aligned} (\partial_t s_y, v) + a(s_y, v) + (y, v) &= (f_r + s_u, v), \quad \forall v \in V, \\ s_y(x, 0) &= (y_0)_r, \end{aligned} \quad (6.6a)$$

$$\begin{aligned} -(\partial_t s_\lambda, \psi) + a(\psi, s_\lambda) + (\psi, \lambda) &= -(s_y - y_r^d, \psi), \quad \forall \psi \in V, \\ s_\lambda(x, T) &= 0, \end{aligned} \quad (6.6b)$$

$$\alpha s_u = s_\lambda. \quad (6.6c)$$

We note that the homogeneous Dirichlet boundary conditions are imposed to (6.6) after differentiating in the same way. The optimality system (6.6) is discretized using the same numerical method, i.e. space-time discontinuous Galerkin discretization, as for (2.6).

6.2.4 Numerical Results

In this section, we present some numerical results to investigate the performance of different bases. The full-order problem is solved using piecewise linear discontinuous

finite elements on a uniform mesh with $h = k = 1/80$ leading to linear systems of size 9600. Three different snapshot sets for W are used to generate the POD basis functions, namely the state Y , the adjoint P and the combination of them $Y \cup P$, as in [72]. Sensitivities derived from CSE method are calculated at the same time steps with FD approximation and we use the latter in the bases generation step. Computation times are obtained on a 2.13 GHz PC. The error between the DG and the reduced solution is measured with respect to $L^2(0, T; L^2(\Omega))$ norm.

We consider the optimal control problem in Sec. 5.3.3 with

$$Q = (0, 1] \times \Omega, \quad \Omega = (0, 1)^2, \\ \epsilon = 10^{-2}, \quad \beta = (y - 1/2, -x + 1/2)^T, \quad r = 1, \quad \alpha = 1.$$

We take the source function f , the desired state y_d and the initial condition y_0 as

$$f(x, t) = y_d(x, t) = 1, \quad y_0(x, t) = 0.$$

The total energy $\mathcal{E}(l)$ in the formula (5.22) is fixed up to $100(1 - \gamma)\%$ by keeping the most energetic POD modes. In this study, we present the results for 7 and 14 POD basis functions setting $\gamma = 10^{-2}$ for comparison purposes. We proceed with the sensitivities with respect to the diffusion term ϵ , the convection term β and the reaction term r .

Case I: Sensitivity with respect to the diffusion term ϵ

We note that the nominal value for the diffusion term is $\epsilon = \epsilon_0 = 10^{-2}$. We generate POD basis once using the snapshots associated to this nominal/baseline value and denote the corresponding results by BPOD in the figures. We choose the parameter range as $1/\epsilon = 80 : 5 : 120$.

Sensitivities of the state and the adjoint are depicted in Figures 6.3 and 6.4, respectively. For comparison purposes, we present the results obtained using CSE method and FD approximation and the results are almost the same. Since the convection field is fixed, the associated sensitivities rotate in the same direction. The state and the adjoint are highly sensitive where their solutions change mostly.

In Figure 6.5, we present the decay of the eigenvalue sensitivities. They decrease rapidly showing that POD can be successfully applied. Sensitivities are computed by CSE method and FD approximation for comparison purposes. We observe that both approaches give almost the same results. In addition, the eigenvalues (see, Figure 5.4) decay following the same pattern as the sensitivities do which means that the ordering will remain in case of parameter perturbations [64].

We present the computational time for the full-order model, CSE method and FD approximation in Table 6.1. In total, the full-problem and the sensitivity equations are solved in 63 and 70 seconds, respectively. The sensitivities are computed using FD approximation in 118 seconds.

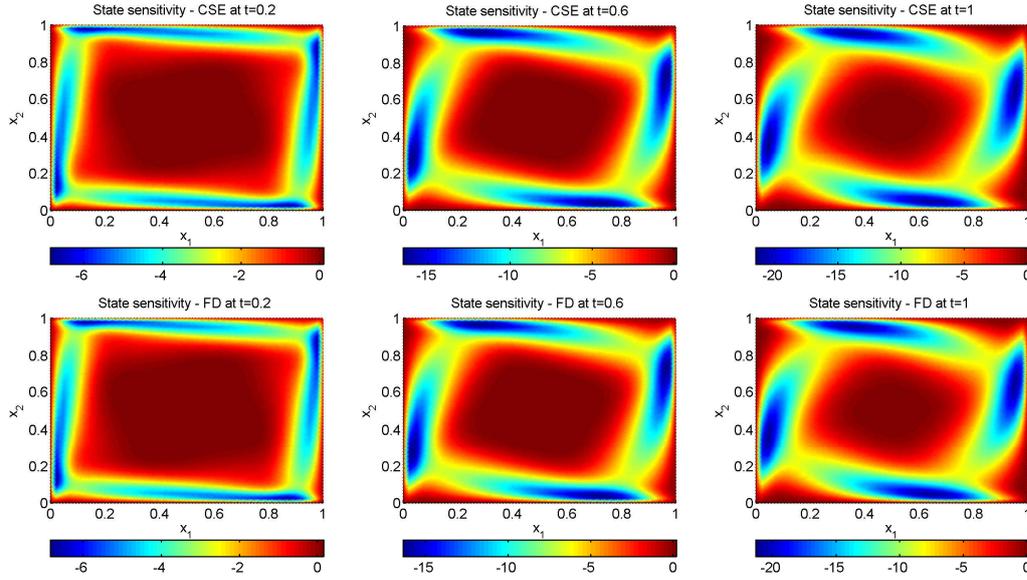


Figure 6.3: DCR Eqn - Case I: State sensitivities computed with CSE method (*first row*) and FD approximation (*second row*) at $t = 0.2, 0.6, 1$, respectively

Table 6.1: DCR Eqn - Case I: Computation times in seconds for the full-order problem

Computing the FE mesh and matrices	\approx	2 s
FE element solution	\approx	15 s
Sensitivity solution using CSE method	\approx	18 s
Sensitivity solution using FD approximation	\approx	2×15 s

In Table 6.2, we compare the computational cost of the reduced problem in terms of different bases and the snapshot ensemble. For each case, the reduced problem is solved less than 11 seconds, which is faster than for the full-problem. Using the snapshot set P , the problem is solved faster than with the snapshot set Y . It is because a better approximation of the control is achieved and led to fast convergence in the optimization step. On the other hand, the size of the set $Y \cup P$ is twice as large as Y or P . Therefore, it takes longer to compute the POD basis and the reduced solution. In terms of POD sensitivities, ExpPOD is slower than ExtPOD; because, its dimension is doubled. We note that the speed of POD gains importance when we have to solve the full-problem several times in case of parameter perturbations.

Table 6.2: DCR Eqn - Case I: Computation times in seconds for the computation of POD basis functions

	Y	P	$Y \cup P$
BPOD	\approx 1.90	1.87	2.80
POD basis sensitivities	\approx 2.00	1.96	2.84
ExtPOD	\approx 3.90	3.83	5.64
ExpPOD	\approx 3.90	3.83	5.64

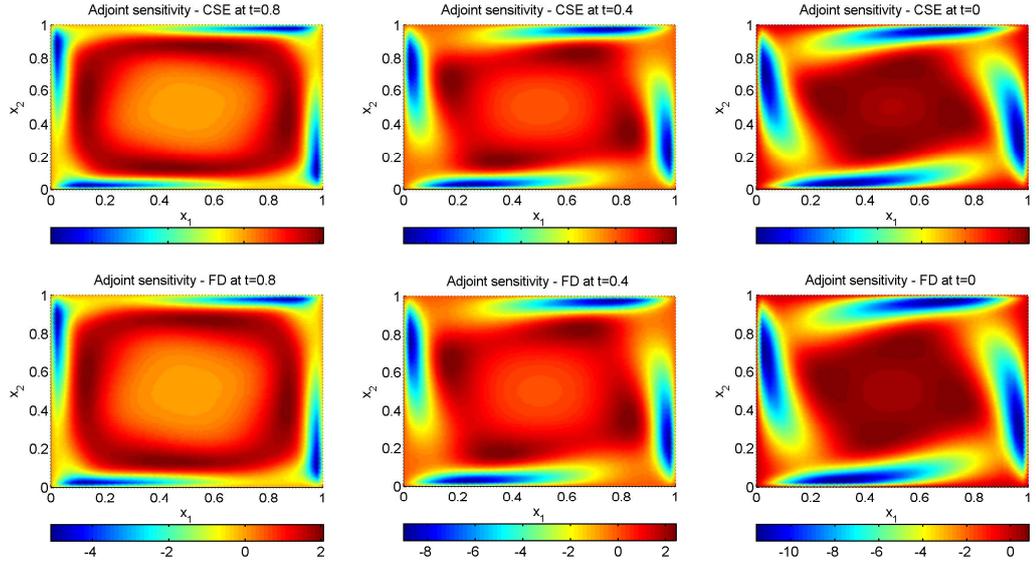


Figure 6.4: DCR Eqn - Case I: Adjoint sensitivities computed with CSE method (*first row*) and FD approximation (*second row*) at $t = 0.2, 0.6, 1$, respectively

In Table 6.3, speed-up of the reduced problem in terms of different bases is shown. We note that the speed-up is defined as the ratio of the computational time of the full-order model to the computational time of the reduced-order model. It is observed that the reduced-order model is solved faster than the full-order model, as expected. There is not a significant difference between the snapshot ensembles, whereas the dimension of the reduced-basis doubles the computation time, namely for ExpPOD basis.

Table 6.3: DCR Eqn - Case I: Computational speedup for the reduced-order model

	Y	P	$Y \cup P$
BPOD	≈ 6.00	5.90	4.42
ExtPOD	≈ 3.28	3.38	2.41
ExpPOD	≈ 2.75	2.79	2.09

Now, we compare the reduced-order solutions obtained by 3 different snapshot sets. In Figure 6.6, we present the error for the control with respect to the diffusion term using 7 and 14 POD basis functions. The control approximated with the POD bases generated from the state solution is poor because the characteristics of the control are totally different from the state solution. It is observed that the choice of the snapshot ensemble affects the approximation depending on whether it contains information about the term which will be approximated or not. The best result for the control is derived with the snapshot set P and $Y \cup P$ with 14 POD basis functions. Inclusion of the adjoint information in the POD basis generation step improves the performance of the method, because the relation between the adjoint and the control is determined through the optimality condition (4.12c). Therefore, the snapshot ensemble $Y \cup P$ leads to the

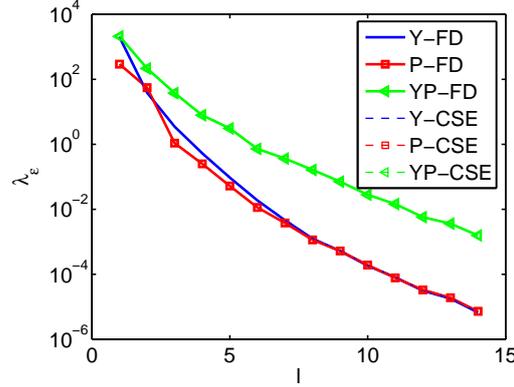


Figure 6.5: DCR Eqn - Case I: Eigenvalue sensitivities for snapshot ensemble Y , P and $Y \cup P$

smallest error. We observe that the baseline and ExtPOD obtained from the snapshot set Y fails to predict the control in case of parameter perturbations since the error remains the same and they gives almost the same results. For each cases, ExpPOD improves the results. However, as we increase the size of the bases, ExpPOD, whose size is doubled, and ExtPOD with snapshot set P and $Y \cup P$ give almost the same result, while a poor approximation to the control can be improved by using ExpPOD.

In Figure 6.7, we present the error for the state with respect to the diffusion term using 7 and 14 POD basis functions. Similarly, the error in the state approximated by the snapshots of the adjoint is higher due to the different properties of the adjoint solution than the state solution. In addition, a good approximation to the control influences the state solution directly due to acting on the right-hand side of (4.12a). For the state solution, the snapshot set Y and $Y \cup P$ outperforms the snapshot set P . It is because the former set contains information about state and the latter offers a good approximation to the control which influences the state solution directly due to the optimality condition. For this example, the best results for both of the state and the control are achieved through the snapshot set $Y \cup P$ with ExtPOD. On the other hand, ExpPOD of size 14 and the other bases of the same size give almost the same poor results with the snapshot set P . Doubling the size of ExpPOD does not lead to better results than the other bases using the snapshot set Y and $Y \cup P$. Thus, ExtPOD is the best choice.

In Figure 6.8, we present the results of the control for negative and positive changes in ϵ , particularly for $\epsilon = 1/80$ and $\epsilon = 1/120$. The results obtained from the snapshot set Y , the bases BPOD and ExtPOD give almost the same results since state solution cannot approximate the control well. For the snapshot ensemble P and $Y \cup P$, up to 5 POD basis functions, ExtPOD leads to worse results than BPOD, while ExtPOD surpass the BPOD as we increase the number of POD basis functions. For control approximation, with the snapshot set P and $Y \cup P$, ExtPOD approximates the state worse than ExpPOD. However, it catches the ExpPOD computed from the snapshot set $Y \cup P$ as we increase the number of POD bases. We note that ExtPOD cannot beat

ExpPOD. For each case, ExpPOD gives the best results.

In Figure 6.9, we present the results of the control error for negative and positive changes in ϵ , particularly for $\epsilon = 1/80$ and $\epsilon = 1/120$. The results obtained from the snapshot set Y , the bases BPOD and ExtPOD give almost the same results since state cannot approximate the control well. For the snapshot ensemble P and $Y \cup P$, up to 5 POD basis functions, ExtPOD leads to worse results than BPOD, while ExtPOD and ExpPOD give the same results for a larger set of POD basis functions.

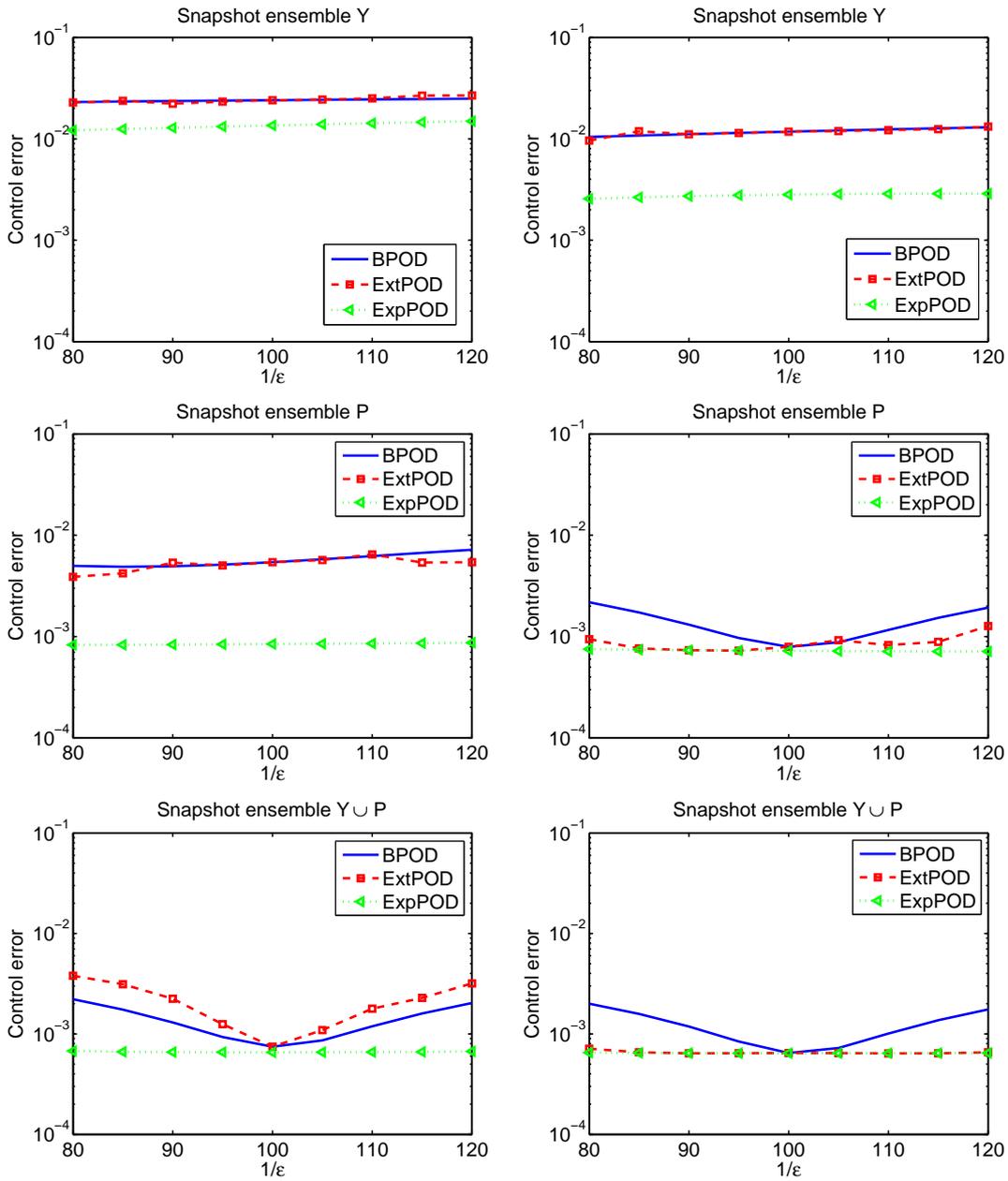


Figure 6.6: DCR Eqn - Case I: Local improvements in the error of the control with 7 (first column) and 14 (second column) POD basis functions

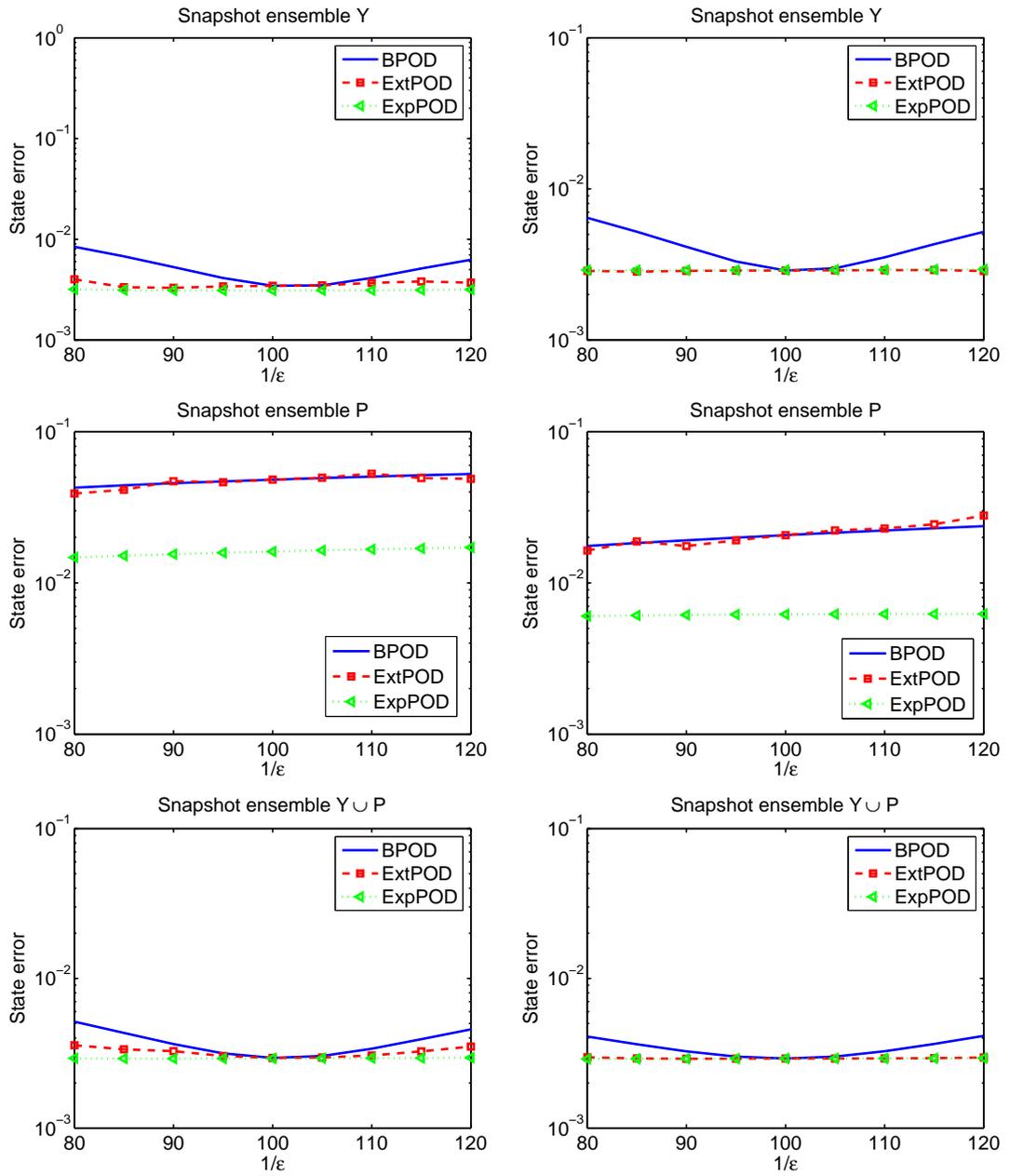


Figure 6.7: DCR Eqn - Case I: Local improvements in the error of the state with 7 (*first column*) and 14 (*second column*) POD basis functions

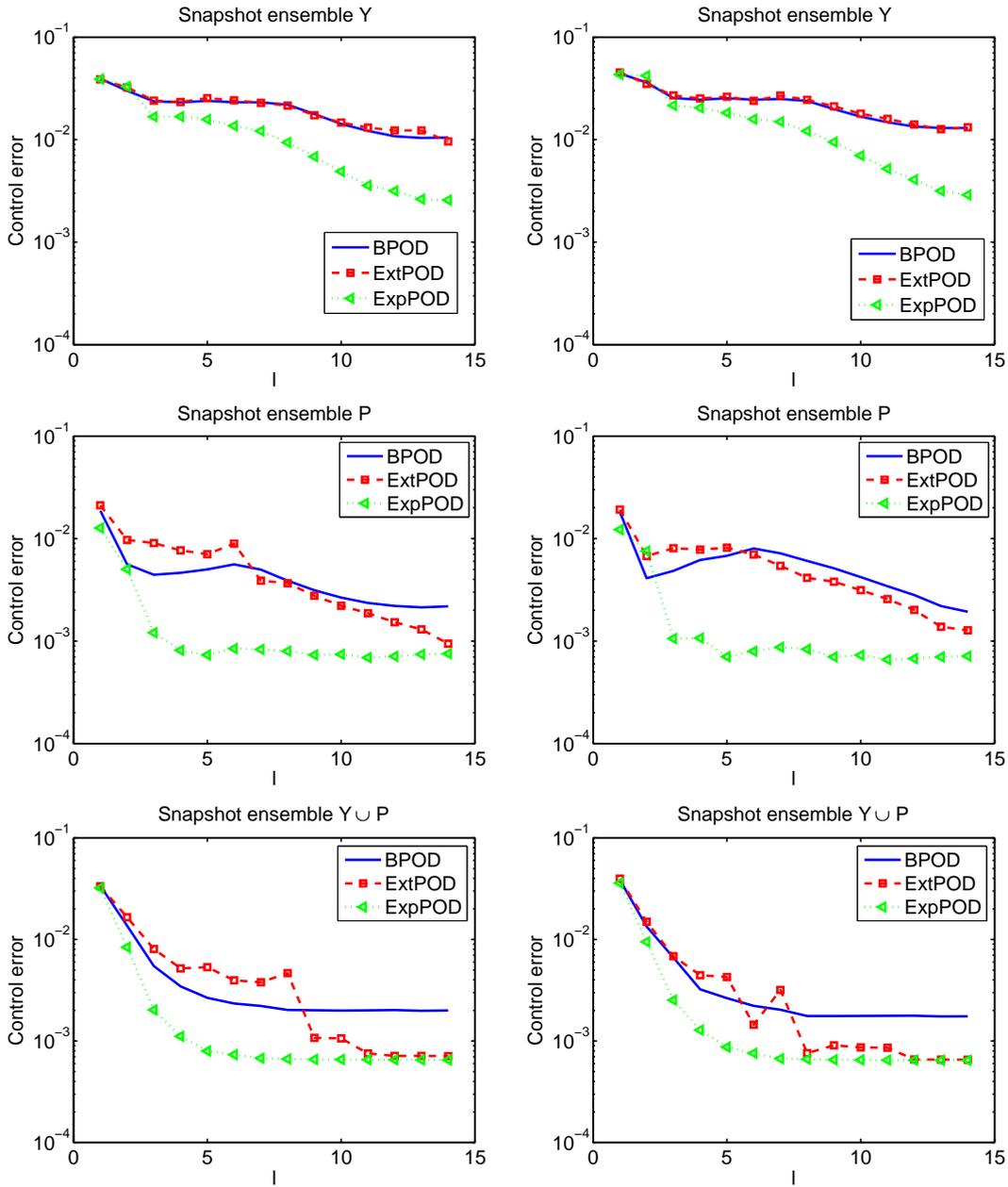


Figure 6.8: DCR Eqn - Case I: Error versus number of POD basis functions for control with $\epsilon = 1/80$ (first column) and $\epsilon = 1/120$ (second column)

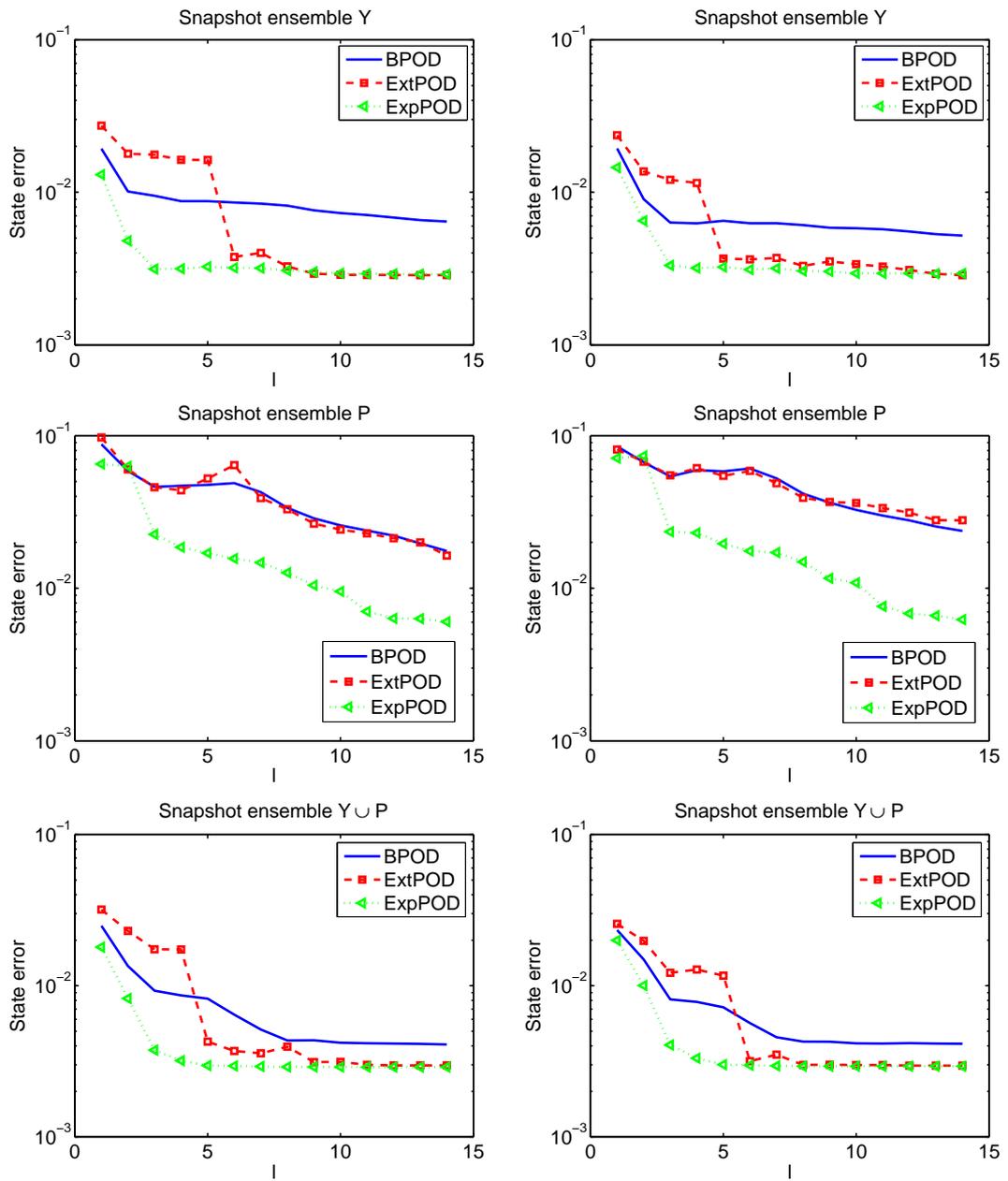


Figure 6.9: DCR Eqn - Case I: Error versus number of POD basis functions for state with $\epsilon = 1/80$ (first column) and $\epsilon = 1/120$ (second column)

Case II: Sensitivity with respect to the convective term β

The nominal value for the convective term is $\beta = \beta_0 = [y - 0.5, -x + p]$ with $p = 0.5$. We generate POD basis once using the snapshots associated to this nominal/baseline value and denote the corresponding results by BPOD in the figures. We choose the parameter range as $p = 0.1 * [1 : 1 : 9]$.

Sensitivities of the state are depicted in Figure 6.10 at $t = 0.2, 0.6, 1$, respectively. For comparison purposes, we present the results obtained from CSE and FD approximation and the results are almost the same. A small difference occur as $t \rightarrow 1$ in the state, while this difference is observe in the adjoint as $t \rightarrow 0$. The state solution is highly sensitive close the boundary $y = 1$ positively and close the boundary $y = 0$ negatively. It is because the fact that as we perturb the parameter p , the convective field shifts upward or downwards (see, Figure .6.1).

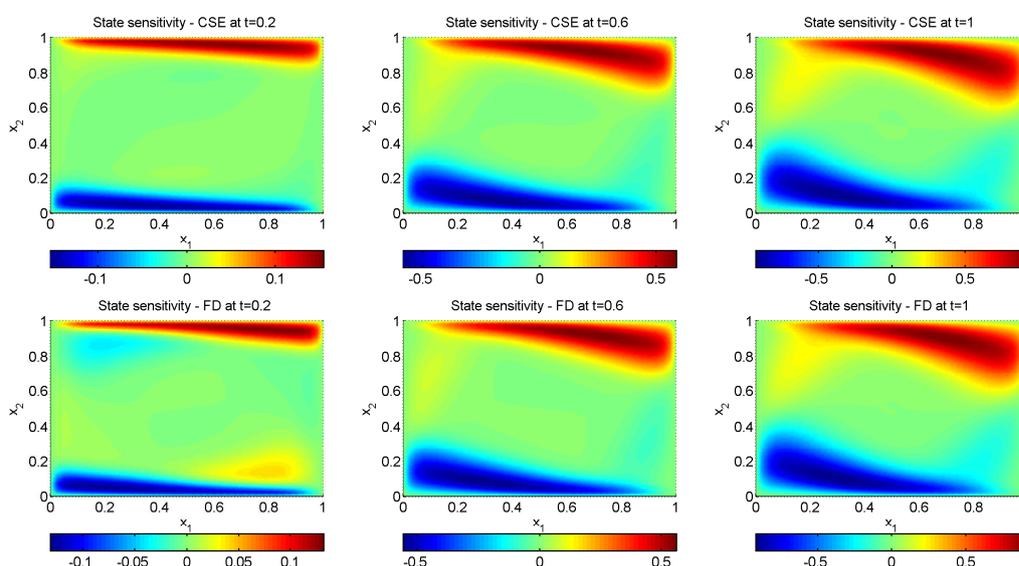


Figure 6.10: DCR Eqn - Case II: State sensitivities computed with CSE method (*first row*) and FD approximation (*second row*) at $t = 0.2, 0.6, 1$, respectively

Sensitivities of the adjoint are depicted in Figure 6.11 at $t = 0.8, 0.4, 0$, respectively. Since the convective term of the state is the negative of the convective term of the adjoint (see, Figure .6.2), the adjoint sensitivities are negative of the state sensitivities.

In Figure 6.12, we present the decay of the eigenvalue sensitivities. We observe that they follow the same pattern as eigenvalues do (see, Figure 5.4). They decrease rapidly showing that POD can be successfully applied. The sensitivities are computed by CSE method and FD approximation for comparison purposes. We observe that a small difference occur for the snapshot set P and it is reflected to the set $Y \cup P$.

In Figure 6.13, we present the error for the control with respect to the parameter perturbations using 7 and 14 POD basis functions. We realize that the sensitivity of the con-

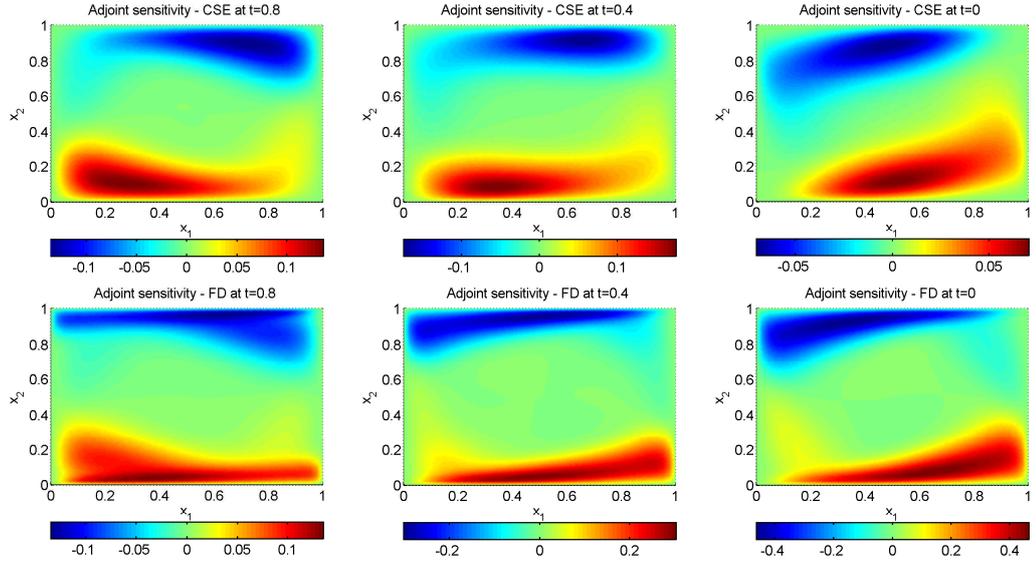


Figure 6.11: DCR Eqn - Case II: Adjoint sensitivities computed with CSE method (*first row*) and FD approximation (*second row*) at $t = 0.8, 0.4, 0$, respectively

trol with respect to p is observed in each figure, since error increases as we moved from the baseline value $p = 0.5$. The snapshot ensemble Y cannot approximate the control. After increasing the number of POD basis functions to 14, a small improvement is observed. The effect of expanding the POD basis appears in the figure positively. For the snapshot set P , increasing the number of POD basis functions results in a smaller error, especially around the nominal value $p = 0.5$. For the snapshot ensemble $Y \cup P$, increasing the number of basis functions, improves the results of ExtPOD most and it surpasses the BPOD. For each case, the best results are derived from ExpPOD. However, in terms of computational speed and the accuracy, ExtPOD obtained from the snapshot set P gives the best results.

In Figure 6.14, we present the error for the state with respect to the parameter perturbations using 7 and 14 POD basis functions. For each figure, the sensitivity of the error with respect to the parameter p is obvious. Among these 6 figures, the largest error is computed using the snapshot set P , since the adjoint and the state have different characteristics. The results obtained from the snapshot set Y and $Y \cup P$ are almost the same and ExpPOD leads to the best accuracy. As we increase the number of POD basis functions, ExtPOD and ExpPOD lead to the same results. Thus, the most accurate results are derived from the snapshot set Y with ExtPOD in a fast way.

In Figure 6.15, we present the results of the control error for negative and positive changes in p , particularly for $\epsilon = 0.1$ and $\epsilon = 0.9$. For each case, the error computed from BPOD is the same. For the snapshot set Y , ExtPOD cannot predict the dynamics of the control, because the state and its derivative are totally different from the control. For small values of basis functions, ExtPOD leads to worse results than BPOD for snapshots P and $Y \cup P$; associated errors decrease for larger POD bases. The most improvement is observed with ExpPOD. For this example, poor results of BPOD can

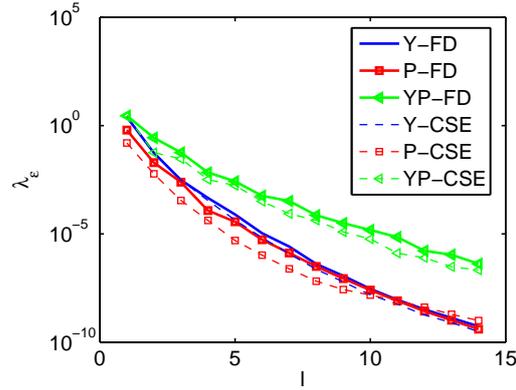


Figure 6.12: DCR Eqn - Case II: Eigenvalue sensitivities for snapshot ensemble Y , P and $Y \cup P$

be improved with the combination of the snapshot set Y and ExpPOD.

In Figure 6.16, we present the results of the state error for negative and positive changes in p , particularly for $\epsilon = 0.1$ and $\epsilon = 0.9$. The error of the BPOD is the same for each POD basis function. After the sixth POD basis function, the results of the BPOD can be improved with ExtPOD or ExpPOD. For larger POD basis set, ExtPOD and ExpPOD give almost the same results. The most accurate solution is derived from the snapshot set Y with ExtPOD in a fast way.

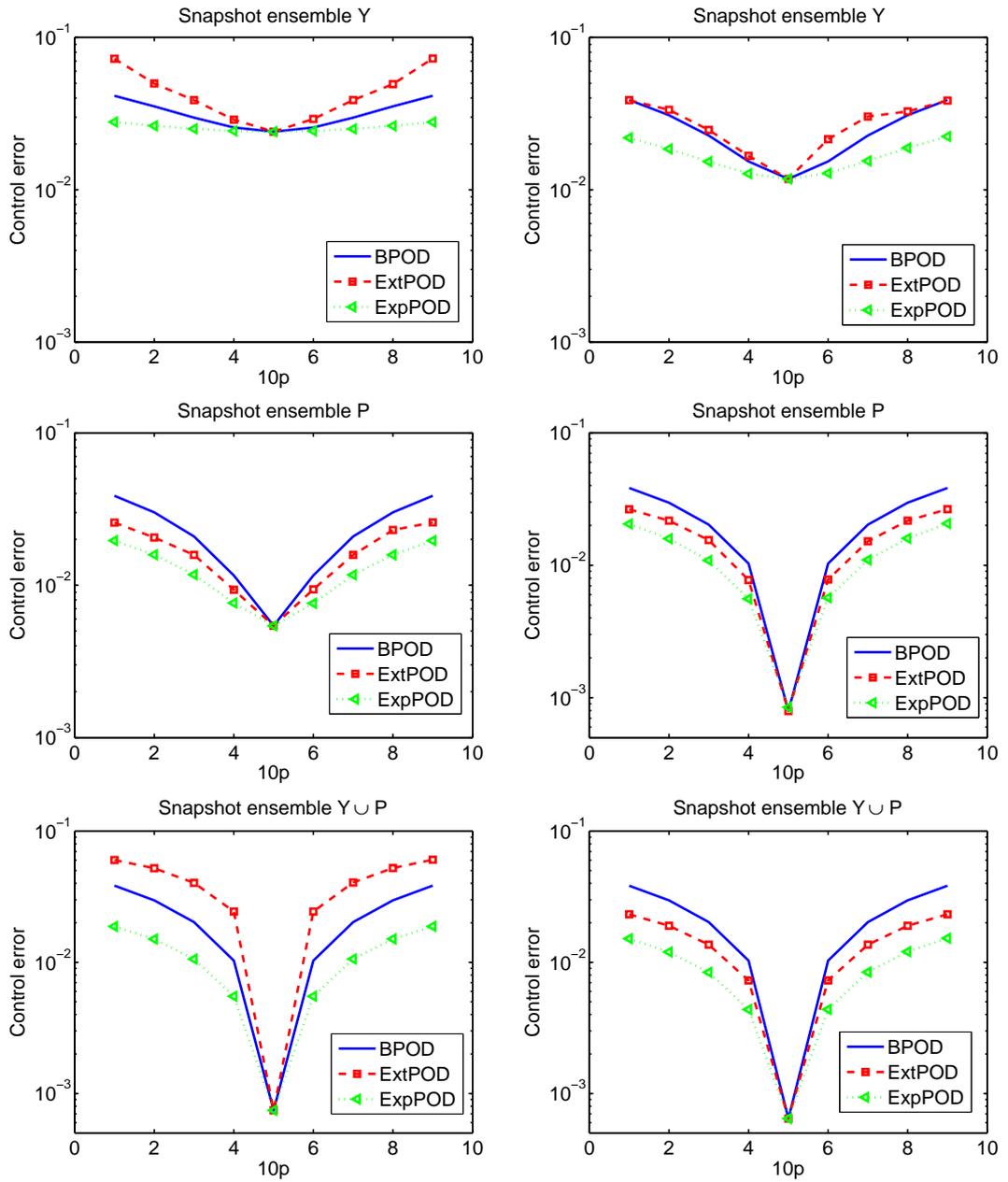


Figure 6.13: DCR Eqn - Case II: Local improvements in the error of the control with 7 (first column) and 14 (second column) POD basis functions

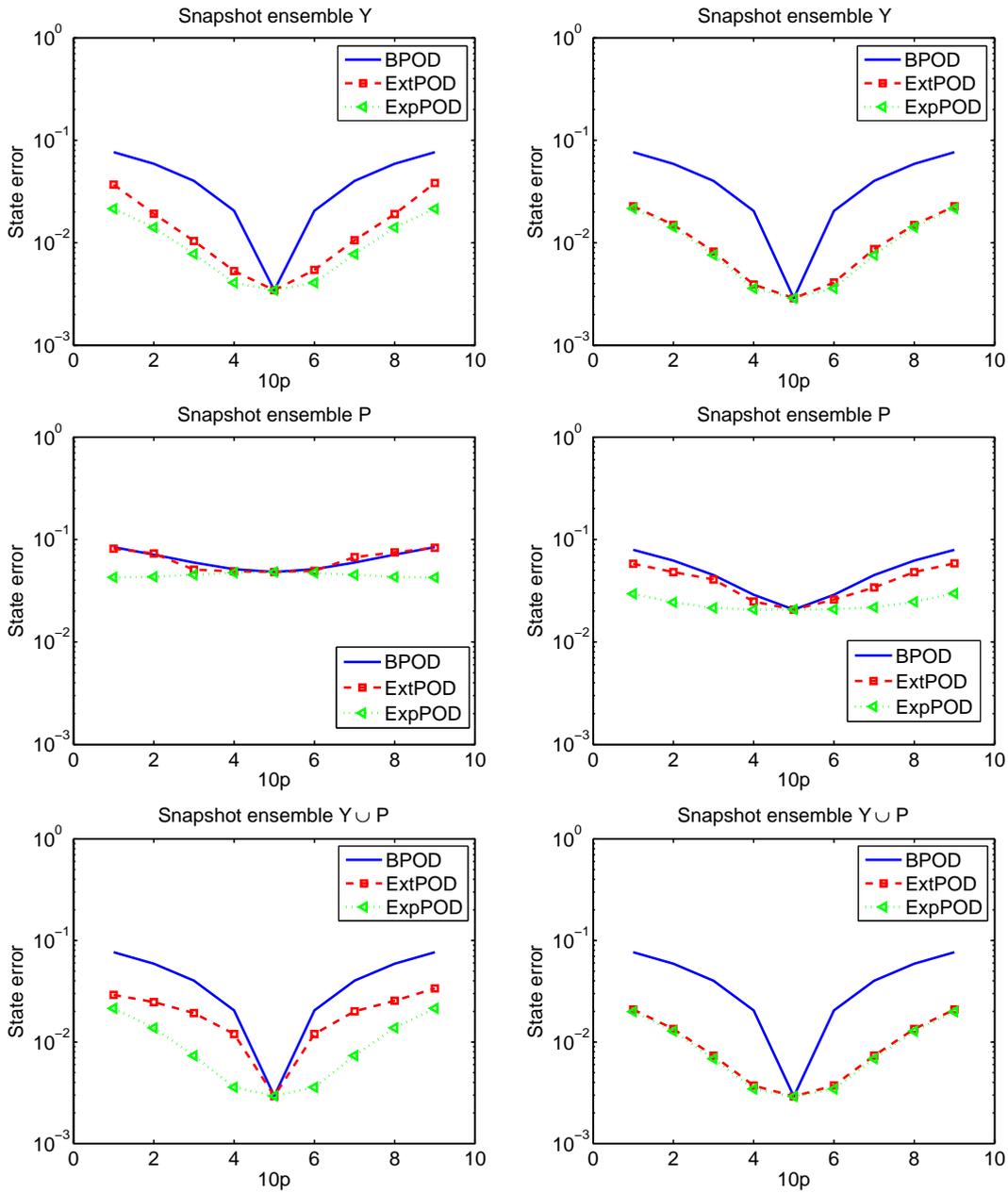


Figure 6.14: DCR Eqn - Case II: Local improvements in the error of the state with 7 (*first column*) and 14 (*second column*) POD basis functions

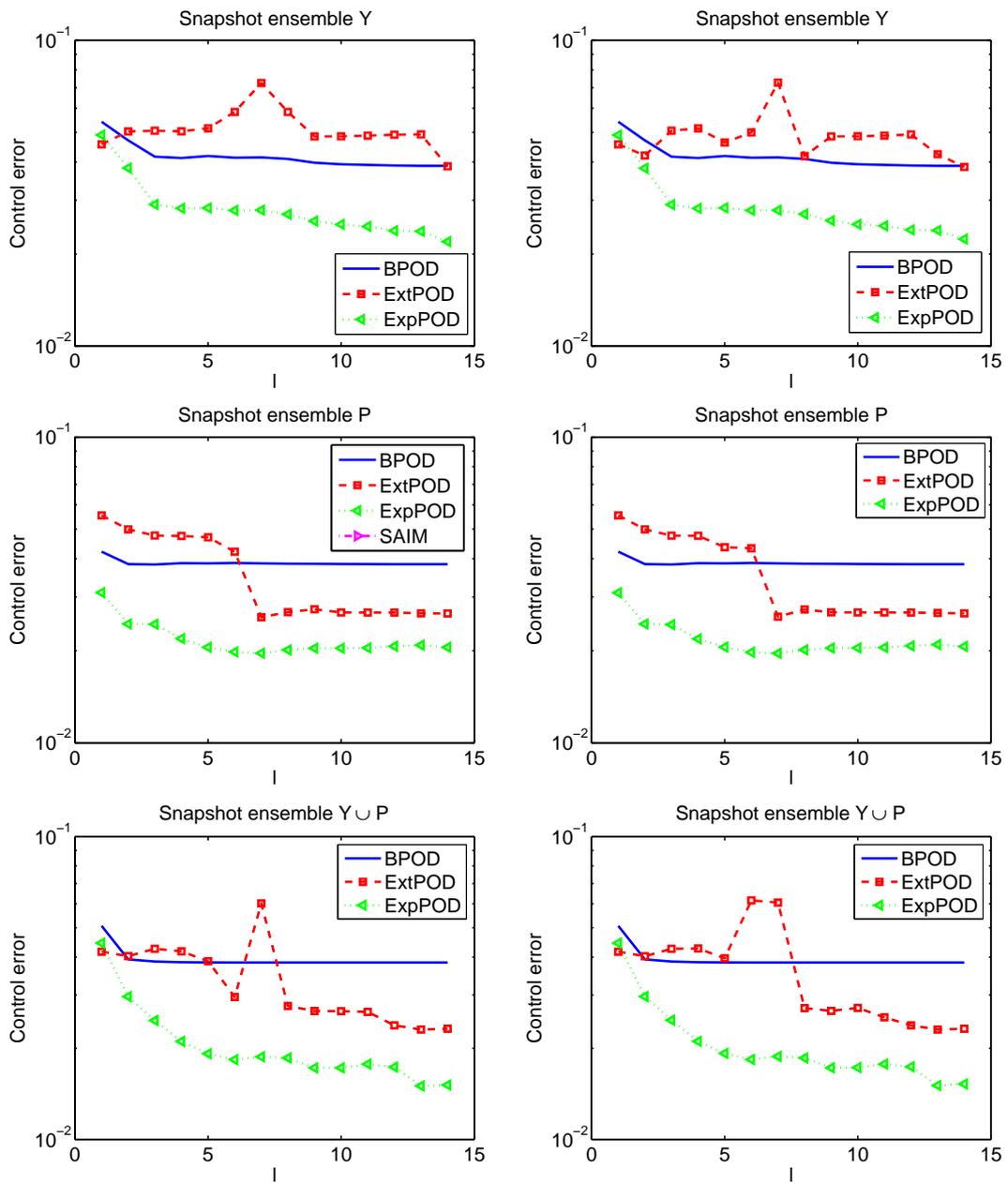


Figure 6.15: DCR Eqn - Case II: Error versus number of POD basis functions for control with $p = 0.1$ (first column) and $p = 0.9$ (second column)

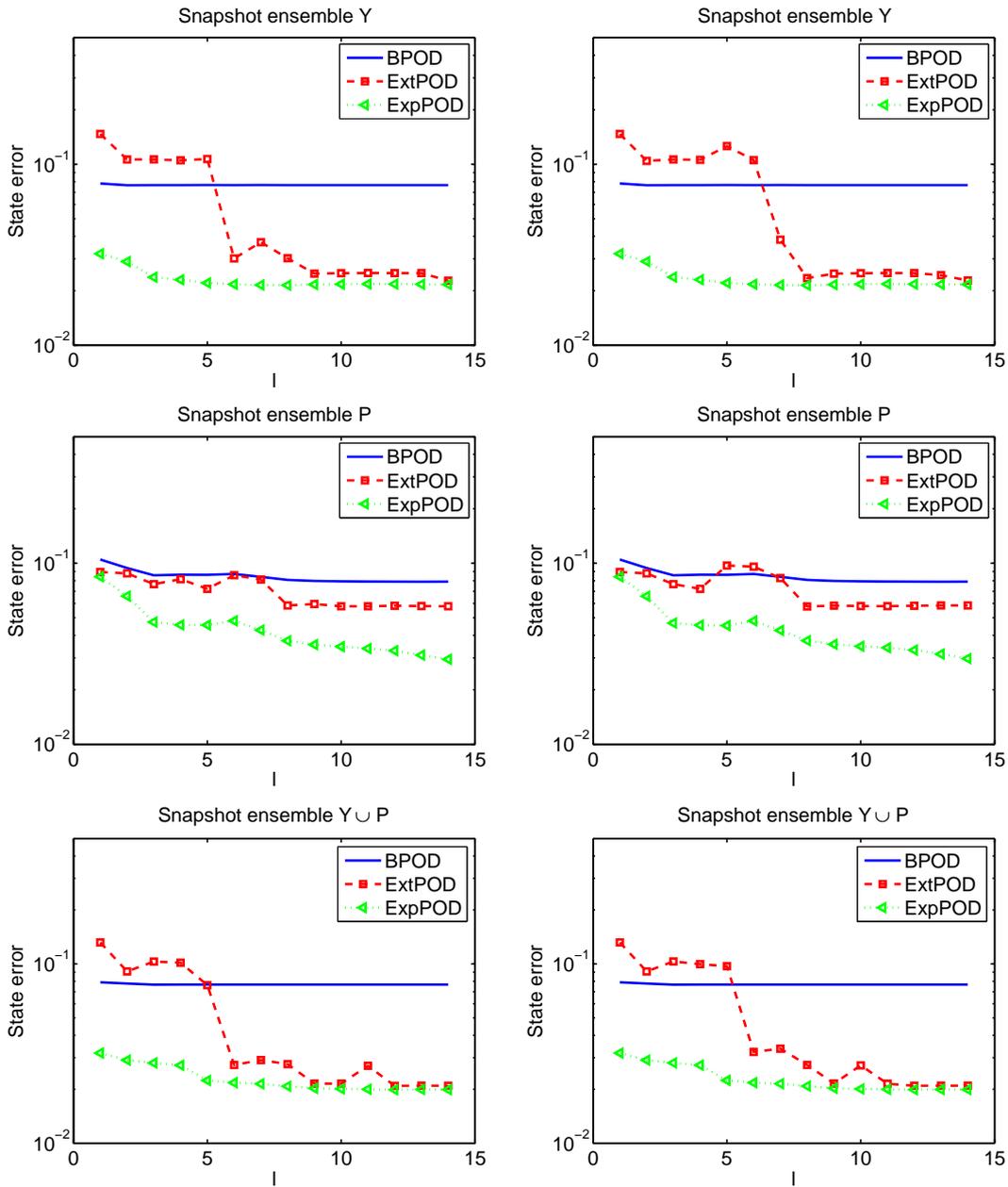


Figure 6.16: DCR Eqn - Case II: Error versus number of POD basis functions for state with $p = 0.1$ (first column) and $p = 0.9$ (second column)

Case III: Sensitivity with respect to the reaction term r

The nominal value for the reaction coefficient is $r = r_0 = 1$. We generate POD basis once using the snapshots associated to this nominal/baseline value and denote the corresponding results by BPOD in the figures. We choose the parameter range as $r = 0.1 * [6 : 1 : 14]$.

Sensitivities of the state are depicted in Figure 6.17 at $t = 0.2, 0.6, 1$, respectively. The state solution is sensitive close to the boundary of the domain; as $t \rightarrow 1$, the sensitivity of the solution starts to distribute into the interior of the domain.

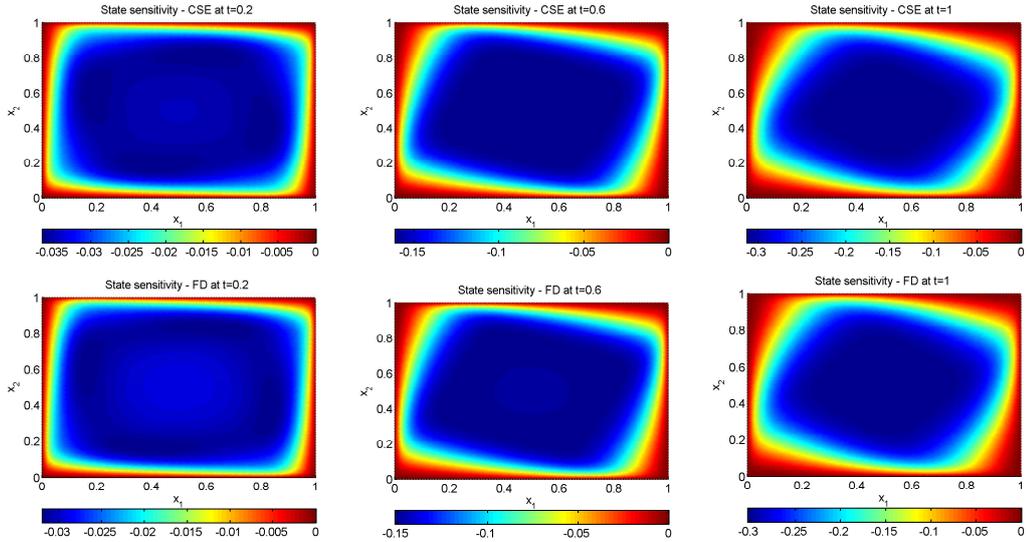


Figure 6.17: DCR Eqn - Case III: State sensitivities computed with CSE method (*first row*) and FD approximation (*second row*) at $t = 0.2, 0.6, 1$, respectively

The sensitivities of the adjoint are depicted in Figure 6.18 at $t = 0.8, 0.4, 0$, respectively. The adjoint is highly sensitive in the center of the domain and the solution becomes sensitive along the boundaries as $t \rightarrow 0$. We compare the results of CSE and FD approximation and observe that both results are almost the same.

In Figure 6.19, we present the decay of the eigenvalue sensitivities. We observe that they follow the same pattern as eigenvalues do (see, Figure 5.4). They decrease rapidly showing that POD can be successfully applied. The sensitivities are computed by CSE method and FD approximation are the same.

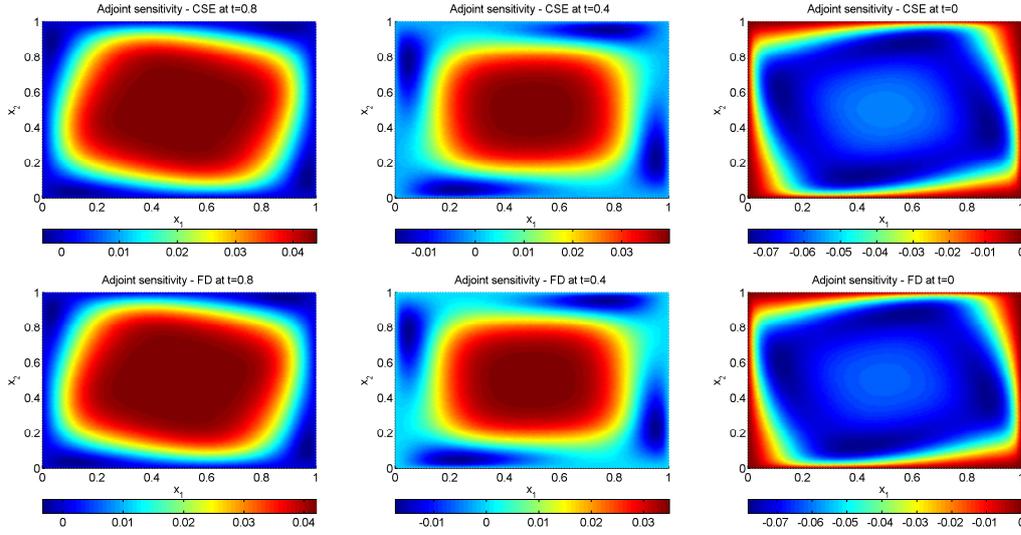


Figure 6.18: DCR Eqn - Case III: Adjoint sensitivities computed with CSE method (*first row*) and FD approximation (*second row*) at $t = 0.8, 0.4, 0$, respectively

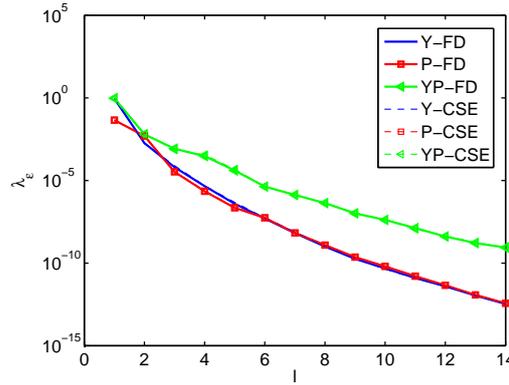


Figure 6.19: DCR Eqn - Case III: Eigenvalue sensitivities for snapshot ensemble Y , P and $Y \cup P$

In Figure 6.20, we present the error for the control with respect to the parameter perturbations using 7 and 14 POD basis functions. For this example, the sensitivity of the error with respect to the reaction term is not observed, because the error remains the same for each parameter value. The state snapshot Y leads to poor results and ExtPOD cannot beat BPOD. On the other hand, expanding the POD basis improves the results. For snapshot set P , increasing the number of POD basis functions cause the results of three bases become almost equal; while for the set $Y \cup P$, it does not affect the error and the results are the same with the ones obtained with the set P . Thus, the most effective combination is the snapshot set P with 14 POD basis functions.

In Figure 6.21, we present the error for the state with respect to the parameter perturbations using 7 and 14 POD basis functions. Snapshot ensemble P leads to poor results,

while the sets Y and $Y \cup P$ lead to the same error. The effect of expanding the basis can be observed for the snapshot set P . However, BPOD cannot be improved using ExtPOD or ExpPOD.

In Figure 6.22, we present the results of the control error for negative and positive changes in p , particularly for $\epsilon = 0.6$ and $\epsilon = 1.4$. For each case, the error computed from BPOD and ExtPOD are the same. The results of BPOD obtained from the set Y are improved with ExpPOD. The error computed with ExpPOD and the set P and $Y \cup P$ stuck. The smallest error is obtained with the set $Y \cup P$ where the results of ExtPOD and ExpPOD become equal as we increase the number of basis functions. Thus, for perturbed problem, ExpPOD computed from the snapshot Y improves the results of BPOD.

In Figure 6.23, we present the results of the state error for negative and positive changes in p , particularly for $\epsilon = 0.6$ and $\epsilon = 1.4$. For each case, the error computed from BPOD and ExtPOD are the same. This time, the results of BPOD obtained from the set P are improved with ExpPOD. The results of ExtPOD and ExpPOD computed from the set Y and $Y \cup P$ are almost the same as we increase the number of basis functions. Thus, for perturbed problem, ExpPOD computed from the snapshot P surpass BPOD.

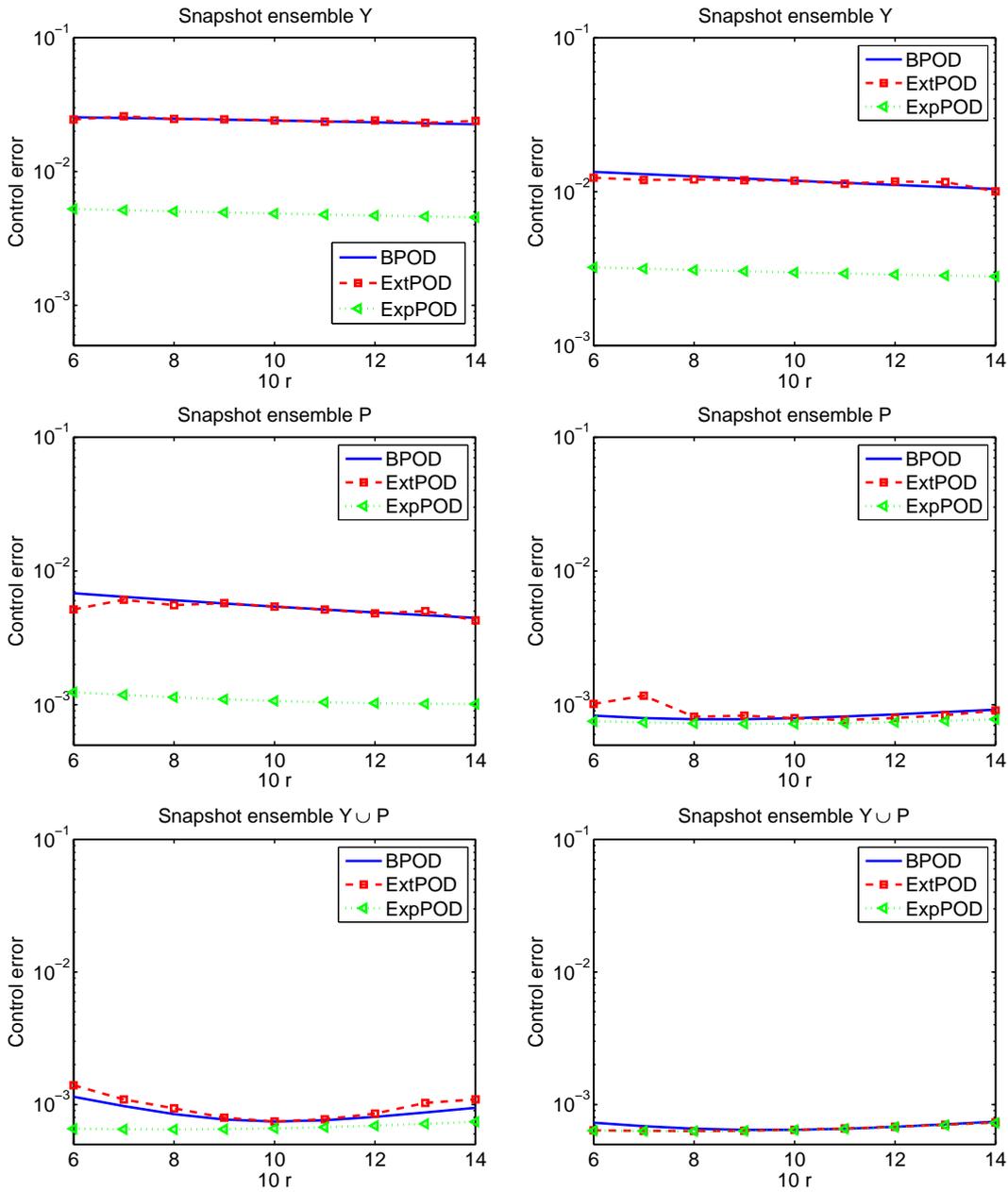


Figure 6.20: DCR Eqn - Case III: Local improvements in the error of the control with 7 (first column) and 14 (second column) POD basis functions

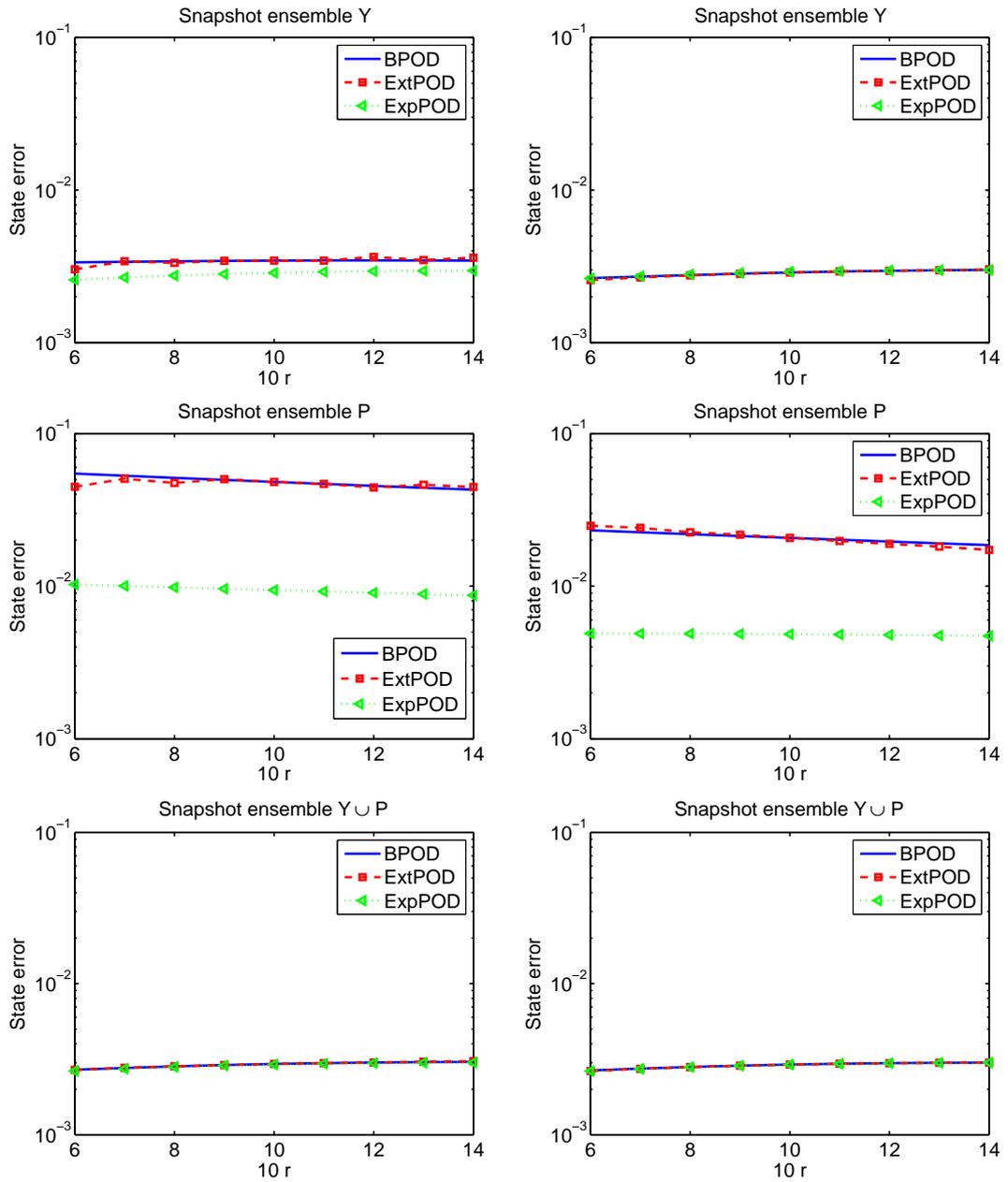


Figure 6.21: DCR Eqn - Case III: Local improvements in the error of the state with 7 (*first column*) and 14 (*second column*) POD basis functions

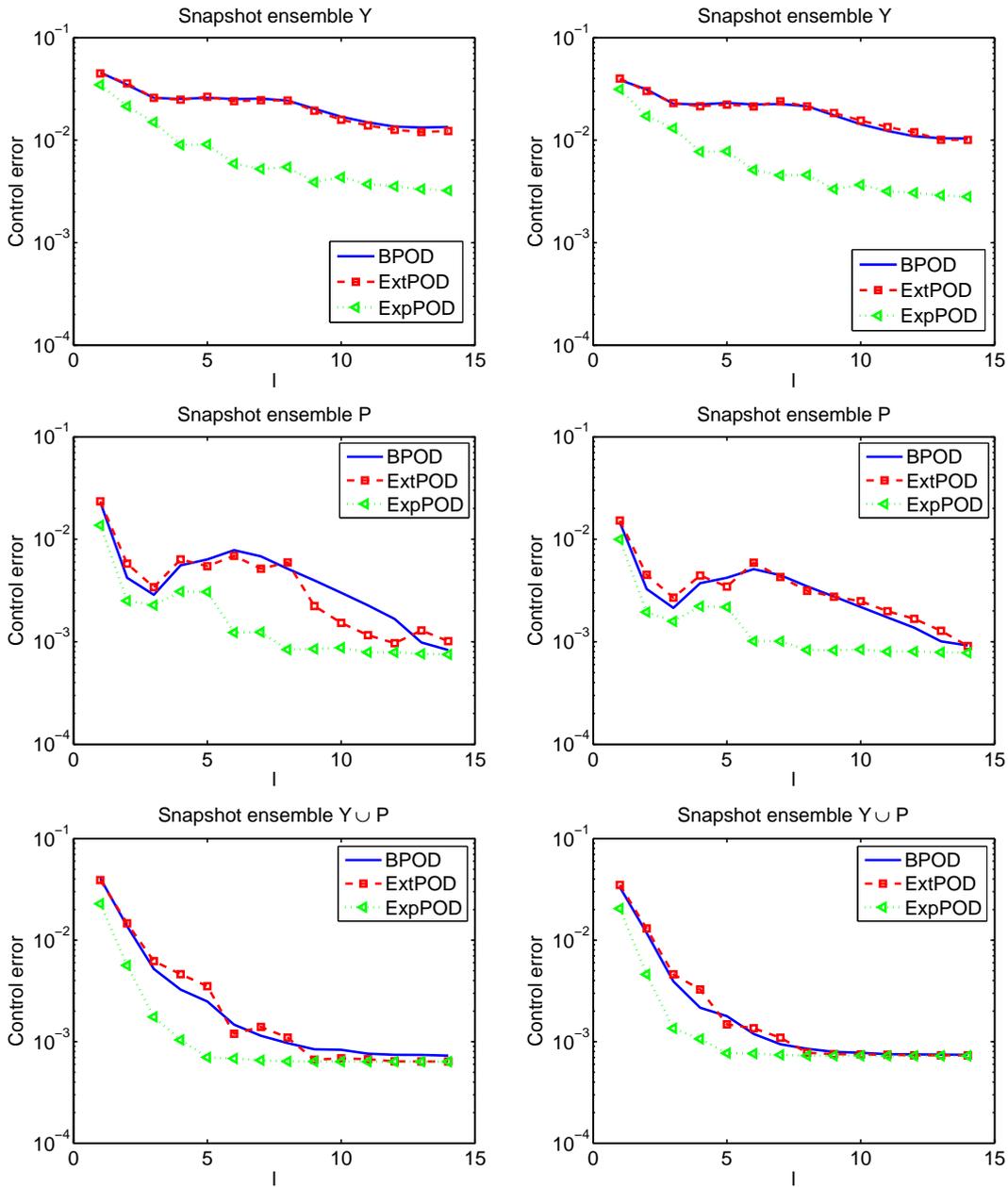


Figure 6.22: DCR Eqn - Case III: Error versus number of POD basis functions for control with $p = 0.6$ (first column) and $p = 1.4$ (second column)

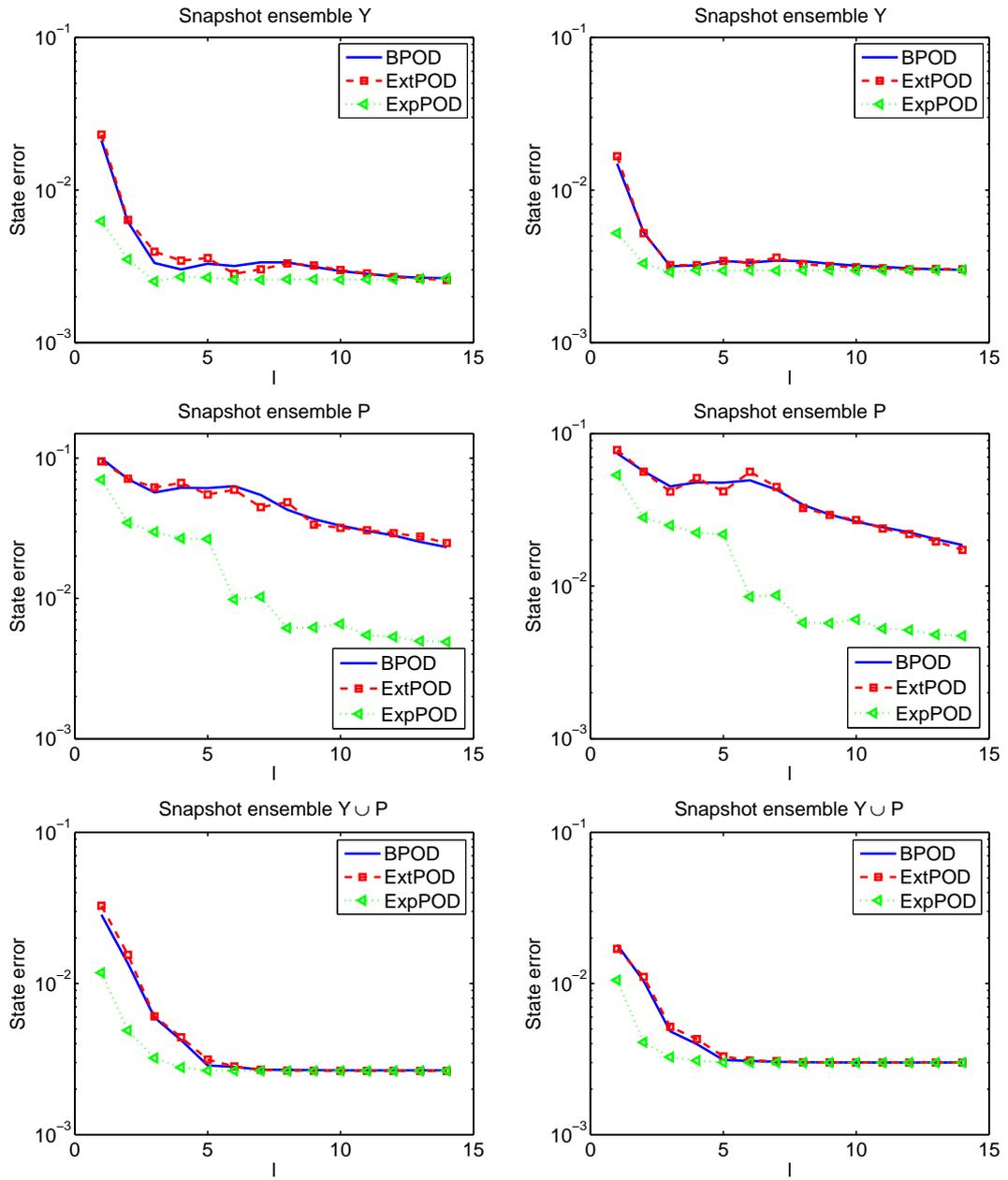


Figure 6.23: DCR Eqn - Case III: Error versus number of POD basis functions for state with $p = 0.6$ (first column) and $p = 1.4$ (second column)

6.3 Sensitivity Equations for Optimal Control of Burgers Equation

Now, we consider the Burgers equation. Sensitivity equations are obtained by differentiating the continuous state (1.2) and the adjoint equation associated to (2.9a) and the optimality condition (2.8c) with respect to ϵ . The corresponding optimality system with s_y , s_p and s_u is written as follows,

$$\begin{aligned} (\partial_t s_y, v) + a(s_y, v) + (s_y x + y s_x) &= (f_\epsilon + s_u, v), \quad \forall v \in V, \\ s_y(x, 0) &= (y_0)_\epsilon, \end{aligned} \quad (6.7a)$$

$$\begin{aligned} -(\partial_t s_\lambda, \psi) + a(\psi, s_\lambda) - (s_y \lambda_x + y (s_\lambda)_x) &= -(s_y - y_\epsilon^d, \psi), \quad \forall \psi \in V, \\ s_\lambda(x, T) &= 0, \end{aligned} \quad (6.7b)$$

$$\alpha s_u = s_\lambda. \quad (6.7c)$$

We note that the homogeneous Dirichlet boundary conditions are imposed to (6.7) after differentiating in the same way. The optimality system (6.7) is discretized using the same numerical method, i.e. space-time discontinuous Galerkin discretization, as for (2.8).

Sensitivity equations are always linear, so CSE would be especially promising for non-linear problems. On the other hand, FD approximation can also be used to find the sensitivities. It requires the evaluation of the OCP depending on the perturbed parameters. In particular, the sensitivity of the state can be computed via the centred difference as follows:

$$s_y(\epsilon_0) \approx \frac{y(\epsilon_0 + \Delta\epsilon) - y(\epsilon_0 - \Delta\epsilon)}{2\Delta\epsilon}. \quad (6.8)$$

The increment $\Delta\epsilon$ is chosen sufficiently small for an accurate FD approximation and it is chosen sufficiently large for the difference between two nearby POD vectors to be larger than the discretization error by one order of magnitude [64].

6.3.1 Numerical Results

In this section, we present some numerical results to investigate the performance of different bases. The Computation times are obtained on a 2.13 GHz desktop PC. The full-order problem is solved using piecewise linear discontinuous finite elements on a uniform mesh with $h = k = 1/200$. Three different snapshot sets for W are used to generate the POD basis functions, namely the state Y , the adjoint P and the combination of them $Y \cup P$, as in [72]. The sensitivities derived from CSE are calculated at the same time steps with FD approximation and we use the former in the bases generation step. The error between the DG and the reduced solution is measured with respect to $L^2(0, T; L^2(\Omega))$ norm.

We note that the nominal value for the diffusion term is $\epsilon_0 = 10^{-2}$. We generate POD basis once using the snapshots associated to this nominal/baseline value and denote

the corresponding results by BPOD in the figures. We choose the parameter range as $1/\epsilon = 80 : 5 : 120$.

We consider the optimal control problem in Sec. 5.5.3 with

$$Q = (0, 1] \times \Omega, \quad \Omega = (0, 1), \quad \epsilon = 10^{-2}, \quad \text{and } \alpha = 0.05.$$

We take the source function $f = 0$, the desired state y_d and the initial condition y_0 are defined as

$$y_d(x, t) = \begin{cases} 1 & \text{in } (0, 1/2], \\ 0 & \text{otherwise.} \end{cases}$$

In Figure 6.24, we present the numerical solutions of the state and the adjoint sensitivities, respectively. We observe that the state is highly sensitive along the path, where the state is tried to be kept on the left-half of the domain, starting at $x = 1/2$. The adjoint, which is solved backward in time, becomes sensitive in a negative way as $t \rightarrow 0$ at the points close the left boundary.

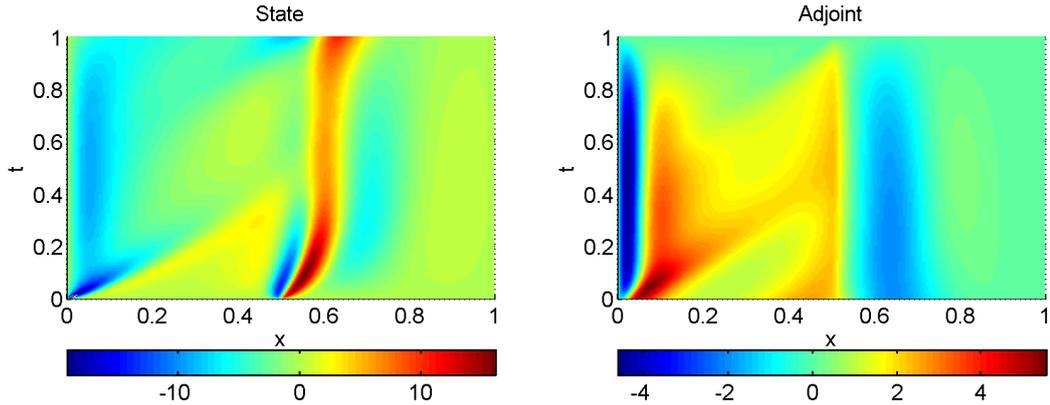


Figure 6.24: Burgers Eqn: State (*left*) and adjoint (*right*) sensitivities

In Figure 6.25, we present the decay of the eigenvalue sensitivities. They decrease rapidly showing that POD can be successfully applied. Sensitivities are computed by CSE and FD approximation for comparison purposes. We observe that both approaches give almost the same results. In addition, the eigenvalues (see, Figure 5.7) decay following the same pattern as the sensitivities do which means that the ordering will remain in case of parameter perturbations [64].

As we discussed in Sec. 5.4, we applied DEIM algorithm and choose 20 DEIM points, see Fig.6.26. We observe that the points are mostly located on the left-half of the domain where the most of the information about the state solution is contained.

We present the computational time for the full-problem, CSE and FD approximation in Table 6.4. In total, the full-problem and the sensitivity equations are solved in 220

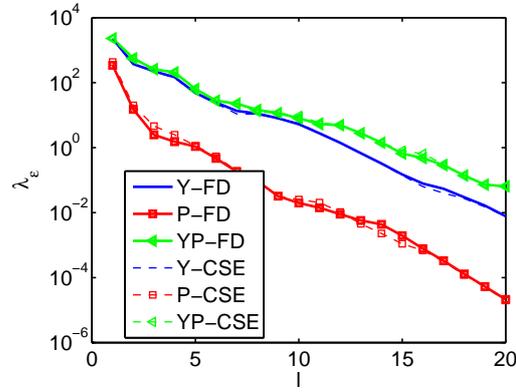


Figure 6.25: Burgers Eqn: Eigenvalue sensitivities for snapshot ensemble Y , P and $Y \cup P$

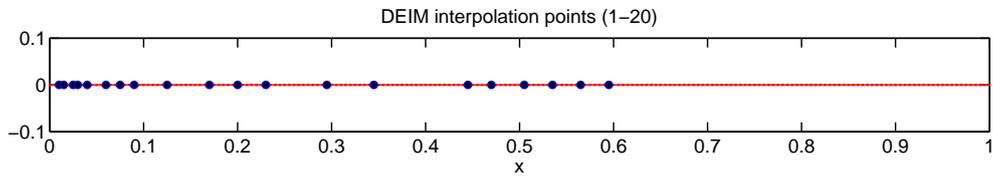


Figure 6.26: DEIM interpolation points

Table 6.4: Burgers Eqn: Computation times for the full-order problem

FE element solution	\approx	220 s
Sensitivity solution using CSE method	\approx	100 s
Sensitivity solution using FD approximation	\approx	2×220 s

and 100 seconds, respectively. The sensitivities are computed using FD approximation in 440 seconds. Here, the efficiency of CSE is revealed.

The total energy $\mathcal{E}(l)$ in the formula (5.22) is fixed up to $100(1 - \gamma)\%$ by setting $\gamma = 10^{-2}$ in order to keep the most energetic POD modes, namely the first 5 POD basis functions. In Table 6.5, the computational time for the POD basis functions are presented. The derivation of the POD basis sensitivities, except the DG approximation of the state and the adjoint sensitivities, is quite fast; so each new bases are computed in less than one second.

Table 6.5: Burgers Eqn: Computation times in seconds for the computation of POD basis functions

		Y	P	$Y \cup P$
BPOD	\approx	0.06	0.10	0.1
POD basis sensitivities	\approx	0.08	0.11	0.21
ExtPOD	\approx	0.14	0.21	0.31
ExpPOD	\approx	0.14	0.21	0.31

In Table 6.6, the speedup of the reduced problem in terms of different bases and the snapshot ensemble with/out DEIM is shown. We note that the speedup is defined as the ratio of the computational time of the full-order model to the computational time of the reduced-order model. It is observed that the reduced-order model is solved faster than the full-order model, as expected. DEIM method is required to speed up the process by computing the nonlinear term $\mathcal{N}^l(\mathbf{y}_{l,m})$ at interpolation points to approximate this term in an efficient way. This necessity is revealed in Table 6.6, because the speedup of the POD model obtained with DEIM is almost 3 times larger than the speedup of the POD model. There is not a significant difference between the snapshot ensembles, whereas the dimension of the reduced-basis doubles the computation time, namely for ExpPOD basis. We note that the speed of POD gains importance when we have to solve the full-problem several times in case of parameter perturbations.

Table 6.6: Burgers Eqn: Computational speedup for the reduced-order model with/out DEIM

		Y	P	$Y \cup P$	Y^D	P^D	$Y \cup P^D$
BPOD	\approx	4.10	4.17	4.00	12.22	12.48	11.58
ExtPOD	\approx	4.01	4.12	4.11	11.58	11.58	10.48
ExpPOD	\approx	2.22	2.29	2.20	6.29	6.30	5.95

In Fig. 6.27, we present the error for the state with respect to the diffusion term using five and ten POD basis functions. Using five POD bases, which are generated from three different snapshot ensembles, the sensitivity of the state with respect to the diffusion term cannot be captured and the error remains unchanged. Increasing the ExpPOD basis functions, which are derived from the snapshots set Y and $Y \cup P$, to ten; sensitivity of the state is revealed and the error is decreased. The state, which is approximated with ten POD bases generated from the adjoint solution, is also poor

because the characteristics of the state are totally different from the adjoint solution. On the other hand, the snapshot set Y and $Y \cup P$ outperform the snapshot set P . It is because the former set contains information about the state and the latter offers a good approximation to the control which acts on the right-hand side of the state equation (2.8a). The best result is achieved at the baseline parameter using ExpPOD generated with the snapshot set Y . As we move away from the nominal value, the error increases; but, the results of BPOD and ExtPOD are improved using ExpPOD without solving the full-problem depending on the perturbed parameters again.

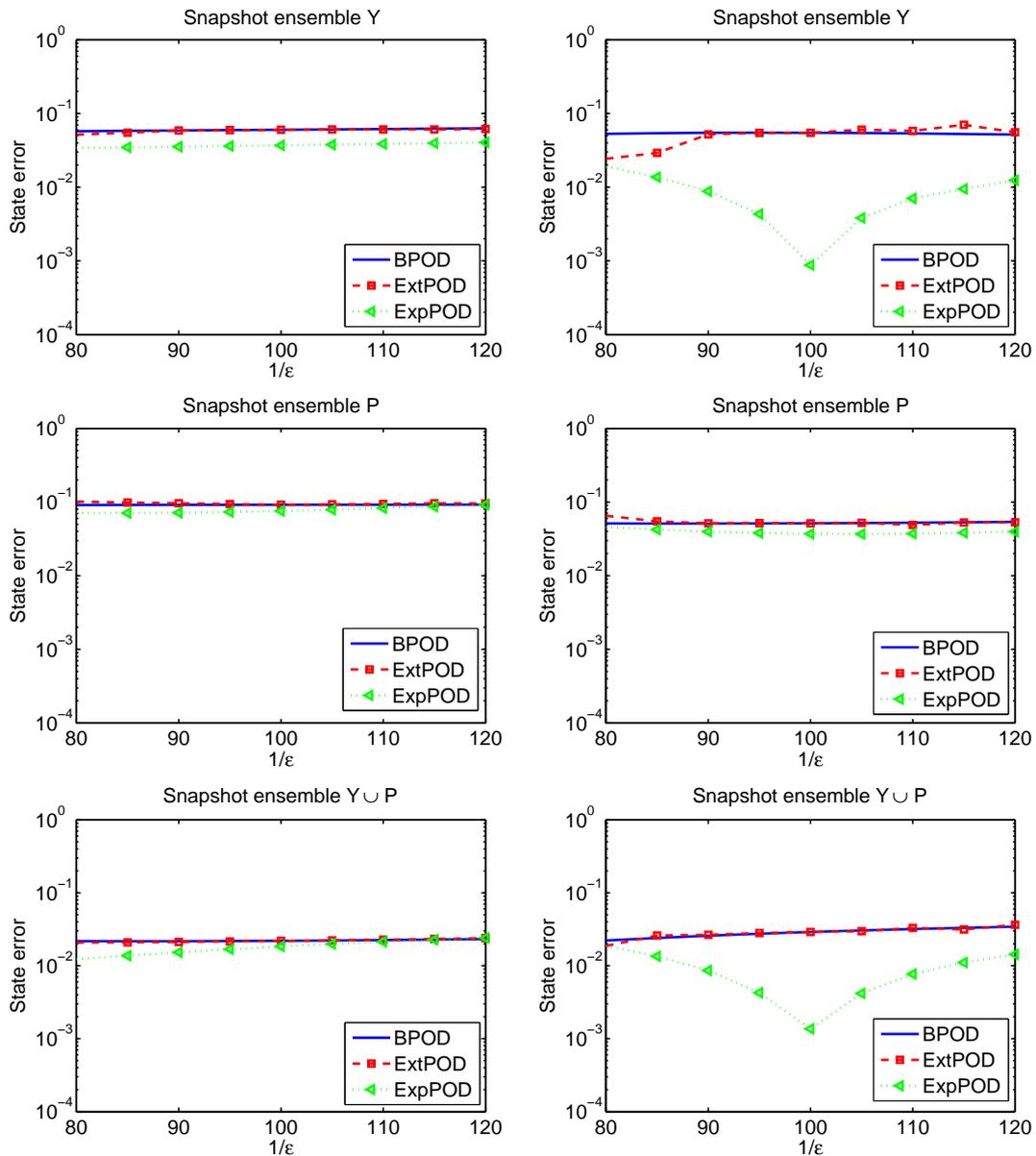


Figure 6.27: Burgers Eqn: Local improvements in the error of the state with 5 (*first column*) and 10 (*second column*) POD basis functions

In Fig. 6.28, we present the error for the control with respect to the diffusion term using five and ten POD basis functions. Using five ExpPOD bases which are generated from the snapshot set P approximates the problem best due to the relation (2.8c) between the adjoint and the control. Use of the state solution in the POD basis generation step leads to inaccurate solutions because the control and the state have different characteristics. Increasing the POD basis functions to ten, poor solutions arising from BPOD and ExtPOD cannot be improved; while ExpPOD which is generated from three different snapshot sets increases the accuracy. The most accurate result for the control is achieved using ExpPOD generated from the snapshot set P . The best result is achieved at the baseline parameter using ExpPOD and the error increases as we move away from the baseline value, as expected.

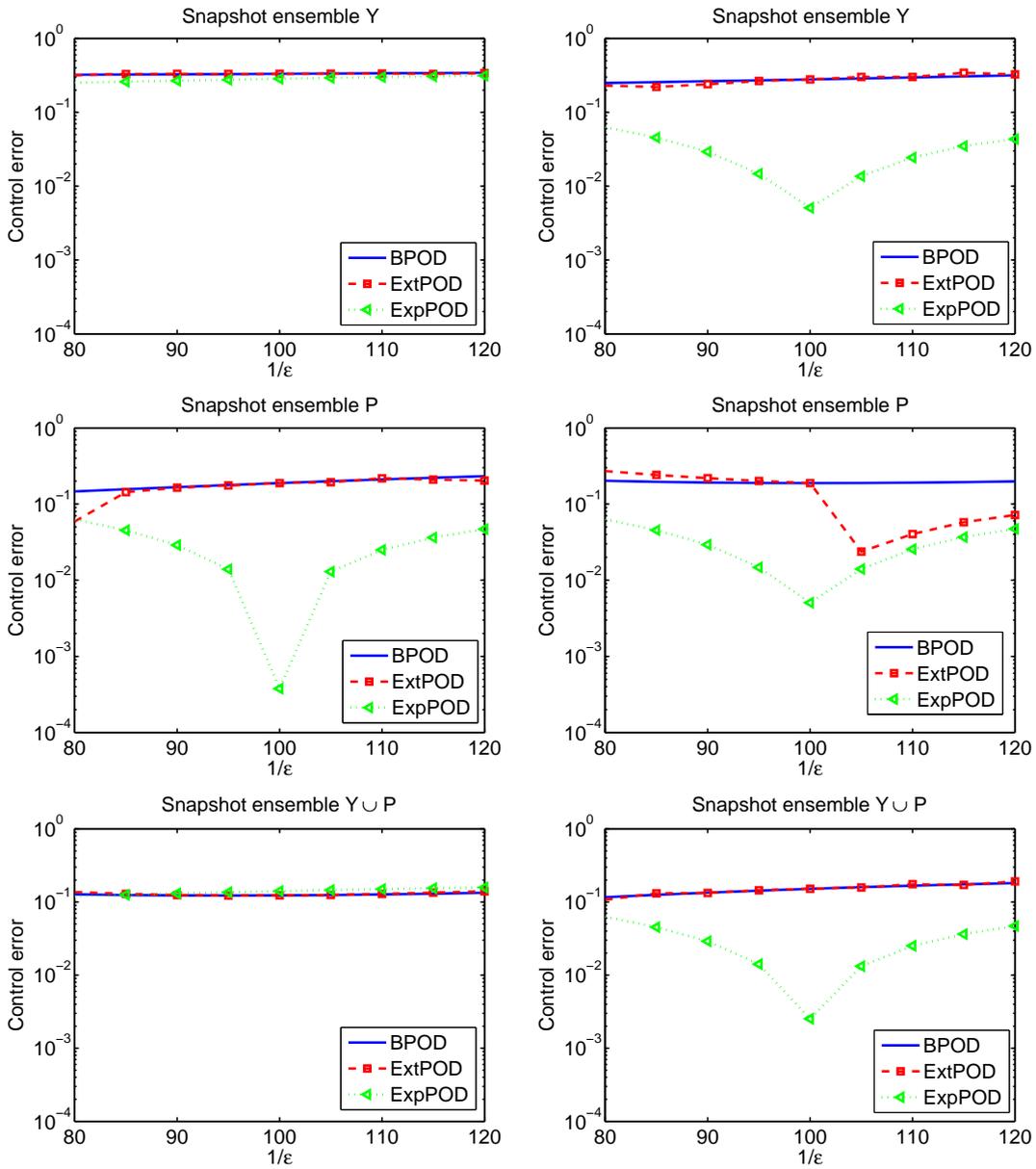


Figure 6.28: Burgers Eqn: Local improvements in the error of the control with 5 (*first column*) and 10 (*second column*) POD basis functions

CHAPTER 7

CONCLUSION

In this study, optimal control problem governed by diffusion-convection-reaction equation and Burgers equation are investigated and the problem is discretized by symmetric interior penalty Galerkin method in space and variational time discretization method. A priori error estimates for space-time discretization are derived. It is observed that the resulting error bound is suboptimal in space and optimal in time in $L^\infty(0, T; L^2(\Omega))$ norm. We compute the linear systems associated to the optimality system where we use Gauss-Radau points to approximate the resulting temporal integrals for dG method and we approximate these integrals using the trapezoidal rule for cGP method. DG method requires less regularity of the solution than the cG method. We observe that for OCP of DCR equation, resulting linear systems of the state and the adjoint equations have the same structure and it might be easy to use a preconditioner of the same form for each system. For Burgers equation, a priori error estimates are proven and the resulting error bound is suboptimal in space and optimal in time in $L^\infty(0, T; L^2(\Omega))$ norm. As opposed to DCR equation, the matrix system associate to linearised state equation and the adjoint equation are quite different from each other due to the nonlinearity in Burgers equation.

In addition to the full-order solution of the OCPs, we have computed the reduced-order solution using proper orthogonal decomposition method. For each case, the reduced-order solution is computed much faster than the full-order solution. Here, we use three different snapshot sets, namely the state Y , the adjoint P and a combination of them $Y \cup P$. A priori error estimates indicate that the error for the reduced solution is of first order accurate with respect to the POD truncation error Λ_ϵ . This result is indicated for the error in the adjoint and the control. The best rate associated to the snapshot set Y is less than the theoretical estimate, which means that the temporal and spatial terms pollute the solution and they are more dominant than Λ_ϵ and the error decays slower than the predicted rate. For optimal control of DCR equation, the snapshot set P fails to predict the state. For small values of POD basis functions, the snapshot set Y gives the best results. However, as we increase the number of basis functions, the results obtained through the snapshot set Y and $Y \cup P$ become almost the same. The snapshot set Y cannot predict the adjoint and the control well. Although the snapshot set P reveals the properties of the adjoint and the control for small number of basis functions, the snapshot set $Y \cup P$ leads to the best results as we increase the number of POD basis functions. Using the snapshot set P , the problem is solved faster than with the snapshot set Y . It is because a better approximation of the control is achieved

and led to fast convergence in the optimization step. On the other hand, the size of the set $Y \cup P$ is twice as large as Y or P . Therefore, it takes longer to compute the POD basis and the reduced solution. POD is not a robust method. Therefore, the nominal basis might not approximate the perturbed problem well. We use the sensitivity of the snapshot and compute the POD basis sensitivities. In terms of POD sensitivities, ExpPOD is slower than ExtPOD; because, its dimension is doubled. For this example, the best results for both of the state and the control are achieved through the snapshot set Y and P with ExtPOD, respectively.

For optimal control of Burgers equation, using small number of POD basis functions, the snapshot sets Y and $Y \cup P$ result in almost the same results of the state. As we increase the number of POD basis functions, more accurate results are obtained using the snapshot set $Y \cup P$. For the control, the snapshot set P gives the best results until 10 POD basis functions are used. For the snapshot set $Y \cup P$, the smallest error is achieved as we increase the number of POD basis functions. It is observed that the reduced-order model is solved faster than the full-order model, as expected. For nonlinear state equation, DEIM method is required to speed up the process. A priori error estimates indicate that the error for the reduced solution is of first order accurate with respect to the POD truncation error Λ_ϵ . The snapshot ensemble $Y \cup P$ approximates the state well and the theoretically estimated convergence rate is achieved for the state, the adjoint and the control. For small values of the POD basis functions, the sensitivity of the state with respect to the diffusion term cannot be captured and the error remains unchanged. Increasing the ExpPOD basis functions, which are derived from the snapshots set Y and $Y \cup P$, to ten; sensitivity of the state is revealed and the error is decreased. The best result is achieved at the baseline parameter using ExpPOD generated with the snapshot set Y . For the control, five ExpPOD bases which are generated from the snapshot set P approximates the problem best. Increasing the POD basis functions to ten, poor solutions arising from BPOD and ExtPOD cannot be improved; while ExpPOD which is generated from three different snapshot sets increases the accuracy. The most accurate result for the control is achieved using ExpPOD generated from the snapshot set P .

As a future work, other model-order reduction techniques can be used, compared with these results and other types of nonlinearities can be incorporated.

REFERENCES

- [1] D. Abbeloos, M. Diehl, M. Hinze, and S. Vandewalle, Nested multigrid methods for time-periodic, parabolic optimal control problems, *Comput. Vis. Sci.*, 14(1), pp. 27–38, 2011.
- [2] T. Akman, Local improvements to reduced-order approximations of optimal control problems governed by diffusion-convection-reaction equation, *Comput. Math. Appl.*, 70(2), pp. 104–131, 2015.
- [3] T. Akman and B. Karasözen, Variational time discretization methods for optimal control problems governed by diffusion–convection–reaction equations, *J. Comput. Appl. Math.*, 272, pp. 41–56, 2014.
- [4] T. Akman and B. Karasözen, Reduced order optimal control using proper orthogonal decomposition sensitivities, in A. Abdulle, S. Deparis, D. Kressner, F. Nobile, and M. Picasso, editors, *Numerical Mathematics and Advanced Applications - ENUMATH 2013*, Lecture Notes in Computational Science and Engineering, Vol. 103, pp. 409–417, Springer International Publishing Switzerland, 2015.
- [5] T. Akman, H. Yücel, and B. Karasözen, A priori error analysis of the upwind symmetric interior penalty galerkin (SIPG) method for the optimal control problems governed by unsteady convection diffusion equations, *Comput. Optim. Appl.*, 57, pp. 703–729, 2014.
- [6] A. Alla and M. Falcone, *An Adaptive POD Approximation Method for the Control of Advection-Diffusion Equations*, volume 164 of *International Series of Numerical Mathematics*, Springer, 2013.
- [7] D. Amsallem, J. Cortial, K. Carlberg, and C. Farhat, A method for interpolating on manifolds structural dynamics reduced-order models, *Internat. J. Numer. Methods Engrg.*, 80, pp. 1241–1258, 2009.
- [8] D. Amsallem and C. Farhat, Interpolation method for adapting reduced-order models and application to aeroelasticity, *AIAA J.*, 46, pp. 1803–1813, 2008.
- [9] H. Antil, M. Heinkenschloss, and R. H. W. Hoppe, Domain decomposition and balanced truncation model reduction for shape optimization of the Stokes system, *Optimization Methods and Software*, 26(4-5), pp. 643–669, 2011.
- [10] H. Antil, M. Heinkenschloss, R. H. W. Hoppe, and D. C. Sorensen, Domain decomposition and model reduction for the numerical solution of PDE constrained optimization problems with localized optimization variables, *Computing and Visualization in Science*, 13, pp. 249–264, 2010.

- [11] A. C. Antoulas, *Approximation of large-scale dynamical systems*, volume 6 of *Advances in Design and Control*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005, with a foreword by Jan C. Willems.
- [12] T. Apel and T. G. Flaig, Crank-Nicolson schemes for optimal control problems with evolution equations, *SIAM J. Numer. Anal.*, 50(3), pp. 1484–1512, 2012.
- [13] D. Arnold, F. Brezzi, B. Cockburn, and L. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *SIAM J. Numer. Anal.*, 39(5), pp. 1749–1779, 2001/02.
- [14] B. Ayuso and L. D. Marini, Discontinuous Galerkin methods for advection-diffusion-reaction problems, *SIAM J. Numer. Anal.*, 47(2), pp. 1391–1420, 2009.
- [15] C. A. Beattie, J. Borggaard, S. Gugercin, and T. Iliescu, A domain decomposition approach to POD, in *Proceedings of the 45th IEEE Conference on Decision and Control, 2006*, Dec. 2006.
- [16] R. Becker, D. Meidner, and B. Vexler, Efficient numerical solution of parabolic optimization problems by finite element methods, *Optimization Methods and Software*, 22(5), pp. 813–833, 2007.
- [17] R. Becker and B. Vexler, Optimal control of the convection-diffusion equation using stabilized finite element methods, *Numer. Math.*, 106(3), pp. 349–367, 2007.
- [18] P. Benner, V. Mehrmann, and D. C. Sorensen, editors, *Dimension reduction of large-scale systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, Springer, Berlin, 2005.
- [19] P. Benner, E. Sachs, and S. Volkwein, Model order reduction for PDE constrained optimization, in *Trends in PDE Constrained Optimization*, International Series of Numerical Mathematics, Vol. 165, Birkhäuser Basel, 2014, to appear.
- [20] A. Björck and H. Golub, Gene, Numerical methods for computing angles between linear subspaces, *Math. Comp.*, 27, pp. 579–594, 1973.
- [21] A. Borzi, Multigrid methods for parabolic distributed optimal control problems, *J. Comput. Appl. Math.*, 157(2), pp. 365–382, 2003.
- [22] T. Bui-Thanh, K. Willcox, O. Ghattas, and B. van Bloemen Waanders, Goal-oriented, model-constrained optimization for reduction of large-scale systems, *J. Comput. Phys.*, 224(2), pp. 880–896, 2007.
- [23] E. Burman, Crank-Nicolson finite element methods using symmetric stabilization with an application to optimal control problems subject to transient advection-diffusion equations, *Commun. Math. Sci.*, 9(1), pp. 319–329, 2011.
- [24] E. Burman and B. Stamm, Bubble stabilized discontinuous Galerkin methods on conforming and non-conforming meshes, *Calcolo*, 48, pp. 189–209, 2021.

- [25] A. Cangiani, E. H. Georgoulis, and S. Metcalfe, Adaptive discontinuous Galerkin methods for nonstationary convection-diffusion problems, *IMA J. Numer. Anal.*, 34(4), pp. 1578–1597, 2014.
- [26] J. Česenek and M. Feistauer, Theory of the space-time discontinuous Galerkin method for nonstationary parabolic problems with nonlinear convection and diffusion, *SIAM J. Numer. Anal.*, 50(3), pp. 1181–1206, 2012.
- [27] S. Chaturantabut and D. C. Sorensen, Nonlinear model reduction via discrete empirical interpolation, *SIAM Journal on Scientific Computing*, 32(5), pp. 2737–2764, 2010.
- [28] G. Chen and S. S. Collis, Optimal control for Burgers flow using the discontinuous Galerkin method, in *AIAA Region IV Student Paper Conference*, 2003.
- [29] K. Chrysafinos, Discontinuous Galerkin approximations for distributed optimal control problems constrained by parabolic PDE's, *Int. J. Numer. Anal. Model.*, 4(3-4), pp. 690–712, 2007.
- [30] K. Chrysafinos, Analysis and finite element approximations for distributed optimal control problems for implicit parabolic equations, *J. Comput. Appl. Math.*, 231(1), pp. 327–348, 2009.
- [31] K. Chrysafinos and N. J. Walkington, Error estimates for the discontinuous Galerkin methods for parabolic equations, *SIAM J. Numer. Anal.*, 44(1), pp. 349–366 (electronic), 2006.
- [32] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, New York, 1978.
- [33] B. Cockburn and C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, *SIAM J. Numer. Anal.*, 35, pp. 2440–2463, 1998.
- [34] S. S. Collis and M. Heinkenschloss, Analysis of the streamline upwind/Petrov Galerkin method applied to the solution of optimal control problems, Technical Report TR02–01, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005–1892, 2002.
- [35] J. C. De los Reyes and K. Kunisch, A comparison of algorithms for control constrained optimal control of the Burgers equation, *Calcolo*, 41, pp. 203–225, 2004.
- [36] L. Dedè, Reduced basis method and a posteriori error estimation for parametrized linear-quadratic optimal control problems, *SIAM Journal on Scientific Computing*, 32(2), pp. 997–1019, 2010.
- [37] L. Dedè, Reduced basis method and error estimation for parametrized optimal control problems with control constraints, *J. Sci. Comput.*, 50(2), pp. 287–305, 2012.

- [38] L. Dedé and A. Quarteroni, Optimal control and numerical adaptivity for advection-diffusion equations, *ESAIM: Mathematical Modelling and Numerical Analysis*, 39, pp. 1019–1040, 2005.
- [39] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Nonlinear Equations and Unconstrained Optimization*, SIAM, Philadelphia, 1996.
- [40] D. A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, volume 69 of *Mathématiques et Applications*, Springer, New York, 2012.
- [41] V. Dolejší and M. Feistauer, Error estimates of the discontinuous Galerkin method for nonlinear nonstationary convection-diffusion problems, *Numer. Funct. Anal. Optim.*, 26(3), pp. 349–383, 2005.
- [42] V. Dolejší, M. Feistauer, V. Kučera, and V. Sobotíková, An optimal $L^\infty(L^2)$ -error estimate for the discontinuous Galerkin approximation of a nonlinear nonstationary convection-diffusion problem, *IMA J. Numer. Anal.*, 28(3), pp. 496–521, 2008.
- [43] V. Dolejší, M. Feistauer, and C. Schwab, A finite volume discontinuous Galerkin scheme for nonlinear convection-diffusion problems, *Calcolo*, 39, pp. 1–40, 2002.
- [44] V. Dolejší, M. Feistauer, and V. Sobotíková, Analysis of the discontinuous Galerkin method for nonlinear convection-diffusion problems, *Comput. Methods Appl. Mech. Engrg.*, 194(25-26), pp. 2709–2733, 2005.
- [45] K. C. E. Casas, A discontinuous Galerkin time-stepping scheme for the velocity tracking problem, *SIAM J. Numer. Anal.*, 50(5), pp. 2281–2306, 2012.
- [46] Y. Epshteyn and B. Rivière, Estimation of penalty parameters for symmetric interior penalty Galerkin methods, *Journal of Computational and Applied Mathematics*, 206(2), pp. 843–872, 2006.
- [47] K. Eriksson and C. Johnson, Error estimates and automatic time step control for nonlinear parabolic problems. I, *SIAM J. Numer. Anal.*, 24(1), pp. 12–23, 1987.
- [48] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. I. A linear model problem, *SIAM J. Numer. Anal.*, 28(1), pp. 43–77, 1991.
- [49] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$, *SIAM J. Numer. Anal.*, 32(3), pp. 706–740, 1995.
- [50] D. J. Estep and R. W. Freund, Using Krylov-subspace iterations in discontinuous Galerkin methods for nonlinear reaction-diffusion systems, in *Discontinuous Galerkin methods (Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pp. 327–335, Springer, Berlin, 2000.
- [51] M. Fahl, *Trust-region Methods for Flow Control based on Reduced Order Modelling*, Ph.D. thesis, Fachbereich IV der Universität Trier, Trier, 2000.

- [52] M. Feistauer, J. Hájek, and K. Svadlenka, Space-time discontinuous Galerkin method for solving nonstationary convection-diffusion-reaction problems, *Appl. Math.*, 52(3), pp. 197–233, 2007.
- [53] M. Feistauer, V. Kučera, K. Najzar, and J. Prokopová, Analysis of space-time discontinuous Galerkin method for nonlinear convection-diffusion problems, *Numer. Math.*, 117(2), pp. 251–288, 2011.
- [54] M. Feistauer and K. Švadlenka, Discontinuous Galerkin method of lines for solving nonstationary singularly perturbed linear problems, *J. Numer. Math.*, 12(2), pp. 97–117, 2004.
- [55] R. W. Freund, Model reduction methods based on Krylov subspaces, *Acta Numer.*, 12, pp. 267–319, 2003.
- [56] H. Fu, A characteristic finite element method for optimal control problems governed by convection-diffusion equations, *J. Comput. Appl. Math.*, 235(3), pp. 825–836, 2010.
- [57] H. Fu and H. Rui, A priori error estimates for optimal control problems governed by transient advection-diffusion equations, *J. Sci. Comput.*, 38(3), pp. 290–315, 2009.
- [58] H. Fu and H. Rui, Adaptive characteristic finite element approximation of convection-diffusion optimal control problems, *Numer. Methods Partial Differential Equations*, 29(3), pp. 979–998, 2013.
- [59] M. Gubisch and S. Volkwein, POD a-posteriori error analysis for optimal control problems with mixed control-state constraints, *Comput. Optim. Appl.*, 2014.
- [60] S. Gugercin, An iterative SVD-Krylov based method for model reduction of large-scale dynamical systems, *Linear Algebra Appl.*, 428(8-9), pp. 1964–1986, 2008.
- [61] S. Gugercin and A. C. Antoulas, A survey of model reduction by balanced truncation and some new results, *Internat. J. Control*, 77(8), pp. 748–766, 2004.
- [62] A. Hay, I. Akhtar, and J. T. Borggaard, On the use of sensitivity analysis in model reduction to predict flows for varying inflow conditions, *Internat. J. Numer. Methods Fluids*, 68(1), pp. 122–134, 2012.
- [63] A. Hay, J. Borggaard, I. Akhtar, and D. Pelletier, Reduced-order models for parameter dependent geometries based on shape sensitivity analysis, *J. Comput. Phys.*, 229(4), pp. 1327–1352, 2010.
- [64] A. Hay, J. T. Borggaard, and D. Pelletier, Local improvements to reduced-order models using sensitivity analysis of the proper orthogonal decomposition, *J. Fluid Mech.*, 629, pp. 41–72, 2009.
- [65] M. Heinkenschloss, Numerical solution of implicitly constrained optimization problems, Technical Report TR08–05, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005–1892, 2008.

- [66] M. Heinkenschloss, T. Reis, and A. Antoulas, Balanced truncation model reduction for systems with inhomogeneous initial conditions, *Automatica*, 47(3), pp. 559 – 564, 2011.
- [67] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods: Analysis, Algorithms, and Applications*, Springer-Verlag, Berlin, 2008.
- [68] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case, *J. Computational Optimization and Applications*, 30, pp. 45–63, 2005.
- [69] M. Hinze and K. Kunisch, Three control methods for time-dependent fluid flow, *Flow, Turbulence and Combustion. An International Journal*, 65(3-4), pp. 273–298, 2000.
- [70] M. Hinze and M. Kunkel, Discrete empirical interpolation in POD model order reduction of drift-diffusion equations in electrical networks, in *Scientific Computing in Electrical Engineering (SCEE 2010)*, volume 16 of *Mathematics in Industry*, pp. 423–431, Springer-Verlag, Berlin Heidelberg, 2012.
- [71] M. Hinze and S. Volkwein, Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control, in P. Benner, V. Mehrmann, and D. C. Sorensen, editors, *Dimension Reduction of Large-Scale Systems*, Lecture Notes in Computational Science and Engineering, Vol. 45, pp. 261–306, Springer-Verlag, Heidelberg, 2005.
- [72] M. Hinze and S. Volkwein, Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition, *Comput. Optim. Appl.*, 39(3), pp. 319–345, 2008.
- [73] M. Hinze, N. Yan, and Z. Zhou, Variational discretization for optimal control governed by convection dominated diffusion equations, *J. Comput. Math.*, 27(2-3), pp. 237–253, 2009.
- [74] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge Monographs on Mechanics, Cambridge University Press, Cambridge, New-York, 1996.
- [75] P. Houston, C. Schwab, and E. Süli, Discontinuous hp -finite element methods for advection-diffusion-reaction problems, *SIAM J. Numer. Anal.*, 39(6), pp. 2133–2163 (electronic), 2002.
- [76] T. Iliescu and Z. Wang, Variational multiscale proper orthogonal decomposition: Navier-stokes equations, *Numer. Methods Partial Differential Eq.*, 30(2), pp. 641—663, 2014.
- [77] P. Jamet, Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain, *SIAM J. Numer. Anal.*, 15(5), pp. 912–928, 1978.
- [78] J.-D. Jansen, O. H. Bosgrac, and P. M. J. V. den Hof, Model-based control of multiphase flow in subsurface oil reservoirs, *Journal of Process Control*, 18(9), pp. 846 – 855, 2009.

- [79] M. Kärcher and M. A. Grepl, *A Posteriori* Error Estimation for Reduced Order Solutions of Parametrized Parabolic Optimal Control Problems, *ESAIM Math. Model. Numer. Anal.*, 48(6), pp. 1615–1638, 2014.
- [80] C. T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.
- [81] C. M. Klaij, J. J. W. van der Vegt, and H. van der Ven, Space-time discontinuous Galerkin method for the compressible Navier-Stokes equations, *J. Comput. Phys.*, 217(2), pp. 589–611, 2006.
- [82] K. Kunisch and S. Volkwein, Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition, *J. Optim. Theory Appl.*, 102(2), pp. 345–371, 1999.
- [83] K. Kunisch and S. Volkwein, Galerkin proper orthogonal decomposition methods for parabolic problems, *Numerische Mathematik*, 90, pp. 117–148, 2001.
- [84] K. Kunisch and S. Volkwein, Proper orthogonal decomposition for optimality systems, *ESIAM M2AN*, 42(1), pp. 1–23, 2008.
- [85] K. Kunisch and S. Volkwein, Optimal snapshot location for computing POD basis functions, *M2AN Math. Model. Numer. Anal.*, 44(3), pp. 509–529, 2010.
- [86] K. Kunisch, S. Volkwein, and L. Xie, HJB-POD-based feedback design for the optimal control of evolution problems, *SIAM J. Appl. Dyn. Syst.*, 3(4), pp. 701–722, 2004.
- [87] K. Kunisch and L. Xie, POD-based feedback control of the Burgers equation by solving the evolutionary HJB equation, *Comput. Math. Appl.*, 49(7-8), pp. 1113–1126, 2005.
- [88] O. Lass and S. Volkwein, POD Galerkin schemes for nonlinear elliptic-parabolic systems, *SIAM J. Sci. Comput.*, 35(3), pp. A1271–A1298, 2013.
- [89] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza, Model order reduction in fluid dynamics: challenges and perspectives, in *Reduced order methods for modeling and computational reduction*, volume 9 of *MS&A. Model. Simul. Appl.*, pp. 235–273, Springer, Cham, 2014.
- [90] A. W. Leung, Optimal harvesting-coefficient control of steady-state prey-predator diffusive Volterra-Lotka systems, *Appl. Math. Optim.*, 31(2), pp. 219–241, 1995.
- [91] D. Leykekhman, Investigation of commutative properties of discontinuous Galerkin methods in PDE constrained optimal control problems, *J. Sci. Comput.*, 53(3), pp. 483–511, 2012.
- [92] D. Leykekhman and M. Heinkenschloss, Local error analysis of discontinuous Galerkin methods for advection-dominated elliptic linear-quadratic optimal control problems, *SIAM J. Numer. Anal.*, 50(4), pp. 2012–2038, 2012.

- [93] D. Leykekhman and B. Vexler, Optimal a priori error estimates of parabolic optimal control problems with pointwise control, *SIAM J. Numer. Anal.*, 51(5), pp. 2797–2821, 2013.
- [94] B. Q. Li, *Discontinuous Finite Elements in Fluid Dynamics and Heat Transfer*, Computational Fluid and Solid Mechanics Series, Springer-Verlag, London, 2006.
- [95] M. Li and P. D. Christofides, Optimal control of diffusion-convection-reaction processes using reduced-order models, *Computers and Chemical Engineering*, 32, pp. 2123–2135, 2008.
- [96] F. Lörcher, G. Gassner, and C.-D. Munz, An explicit discontinuous Galerkin scheme with local time-stepping for general unsteady diffusion equations, *J. Comput. Phys.*, 227(11), pp. 5649–5670, 2008.
- [97] X. Luo, Y. Chen, Y. Huang, and T. Hou, Some error estimates of finite volume element method for parabolic optimal control problems, *Optimal Control Appl. Methods*, 35(2), pp. 145–165, 2014.
- [98] Z. Luo, J. Chen, I. M. Navon, and X. Yang, Mixed finite element formulation and error estimates based on proper orthogonal decomposition for the non-stationary Navier-Stokes equations, *SIAM J. Numer. Anal.*, 47(1), pp. 1–19, 2008/09.
- [99] G. Matthies and F. Schieweck, Higher order variational time discretizations for nonlinear systems of ordinary differential equations, 2011, preprint 23/2011, Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg.
- [100] D. Meidner and B. Vexler, Adaptive space-time finite element methods for parabolic optimization problems, *SIAM J. Control Optim.*, 46(1), pp. 116–142 (electronic), 2007.
- [101] D. Meidner and B. Vexler, A priori error estimates for space-time finite element discretization of parabolic optimal control problems. I. Problems without control constraints, *SIAM J. Control Optim.*, 47(3), pp. 1150–1177, 2008.
- [102] D. Meidner and B. Vexler, A priori error estimates for space-time finite element discretization of parabolic optimal control problems. II. Problems with control constraints, *SIAM J. Control Optim.*, 47(3), pp. 1301–1329, 2008.
- [103] F. Negri, G. Rozza, A. Manzoni, and A. Quarteroni, Reduced basis method for parametrized elliptic optimal control problems, *SIAM J. Sci. Comput.*, 35(5), pp. A2316–A2340, 2013.
- [104] A. Noack and A. Walther, Adjoint concepts for the optimal control of Burgers equation, *Comput. Optim. Appl.*, 36, pp. 109–133, 2007.
- [105] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Verlag, Berlin, Heidelberg, New York, second edition, 2006.

- [106] G. Nupur and N. Neela, An hp -discontinuous Galerkin method for the optimal control problem of laser surface hardening of steel, *ESAIM Math. Model. Numer. Anal.*, 45(6), pp. 1081–1113, 2011.
- [107] A. Patera and G. Rozza, *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*, MIT Pappalardo Graduate Monographs in Mechanical Engineering, Cambridge, MA, 2007, available from http://augustine.mit.edu/methodology/methodology_book.htm.
- [108] J. Peraire and P.-O. Persson, The compact discontinuous Galerkin (CDG) method for elliptic problems, *SIAM J. Sci. Comput.*, 30(4), pp. 1806–1824, 2008.
- [109] A. Quarteroni, G. Rozza, and A. Manzoni, Certified reduced basis approximation for parametrized partial differential equations and applications, *J. Math. Ind.*, 1, pp. Art. 3, 44, 2011.
- [110] W. H. Reed and T. R. Hill, Triangular mesh methods for the neutron transport equation, Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.
- [111] T. Richter, A. Springer, and B. Vexler, Efficient numerical realization of discontinuous Galerkin methods for temporal discretization of parabolic problems, *Numer. Math.*, 124(1), pp. 151–182, 2013.
- [112] B. Rivière, *Discontinuous Galerkin methods for solving elliptic and parabolic equations*, volume 35 of *Frontiers in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008, ISBN 978-0-898716-56-6, theory and implementation.
- [113] A. Rösch, Error estimates for parabolic optimal control problems with control constraints, *Z. Anal. Anwendungen*, 23(2), pp. 353–376, 2004.
- [114] M. Schmich and B. Vexler, Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations, *SIAM J. Sci. Comput.*, 30(1), pp. 369–393, 2007/08.
- [115] D. Schötzau and C. Schwab, Time discretization of parabolic problems by the hp -version of the discontinuous Galerkin finite element method, *SIAM J. Numer. Anal.*, 38(3), pp. 837–875, 2000.
- [116] M. A. Singer and W. H. Green, Using adaptive proper orthogonal decomposition to solve the reaction-diffusion equation, *Appl. Numer. Math.*, 59(2), pp. 272–279, 2009.
- [117] A. Springer and B. Vexler, Third order convergent time discretization for parabolic optimal control problems with control constraints, *Comput. Optim. Appl.*, 57(1), pp. 205–240, 2014.
- [118] A. Studinger and S. Volkwein, *Numerical Analysis of POD A-posteriori Error Estimation for Optimal Control*, volume 164 of *International Series of Numerical Mathematics*, Springer, 2013.

- [119] T. Sun, Discontinuous Galerkin finite element method with interior penalties for convection diffusion optimal control problem, *Int. J. Numer. Anal. Model.*, 7(1), pp. 87–107, 2010.
- [120] V. Thomée, *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics*, Springer-Verlag, Berlin, second edition, 2006.
- [121] P. Tiso and D. J. Rixen, Discrete empirical interpolation method for finite element structural dynamics, in *Topics in Nonlinear Dynamics I*, volume 35 of *Conference Proceedings of the Society for Experimental Mechanics Series*, pp. 203–212, Springer, New York Heidelberg Dordrecht London, 2013.
- [122] F. Tröltzsch, *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, 2010, theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.
- [123] F. Tröltzsch and S. Volkwein, The SQP method for control constrained optimal control of the Burgers equation, *ESAIM. Control, Optimisation and Calculus of Variations. European Series in Applied and Industrial Mathematics*, 6, pp. 649–674 (electronic), 2001.
- [124] F. Tröltzsch and S. Volkwein, POD a-posteriori error estimates for linear-quadratic optimal control problems, *Comput. Optim. Appl.*, 44(1), pp. 83–115, 2009.
- [125] D. Uciński, Optimal sensor location for parameter estimation of distributed processes, *Internat. J. Control*, 73(13), pp. 1235–1248, 2000.
- [126] M. Ulbrich, Constrained optimal control of Navier-Stokes flow by semismooth Newton methods, *Systems Control Lett.*, 48(3-4), pp. 297–311, 2003, optimization and control of distributed systems.
- [127] J. J. W. van der Vegt and J. J. Sudirham, A space-time discontinuous Galerkin method for the time-dependent Oseen equations, *Appl. Numer. Math.*, 58(12), pp. 1892–1917, 2008.
- [128] F. Vetrano, L. C. Garrec, G. D. Mortchelewicz, and R. Ohayon, Assessment of strategies for interpolating pod based reduced order model and application to aeroelasticity, *Journal of Aeroelasticity and Structural Dynamics*, 2, pp. 85–104, 2011.
- [129] M. Vlasák and V. Dolejší, Implicit-explicit runge-kutta discontinuous galerkin finite element method for convection-diffusion problems, in K. Kunisch, G. Of, and O. Steinbach, editors, *Numerical Mathematics and Advanced Applications. ENUMATH 2007*, pp. 355–362, Springer-Verlag, Heidelberg, 2008.
- [130] M. Vlasák, V. Dolejší, and J. Hájek, A priori error estimates of an extrapolated space-time discontinuous Galerkin method for nonlinear convection-diffusion problems, *Numer. Methods Partial Differential Equations*, 27(6), pp. 1456–1482, 2011.

- [131] S. Volkwein, Application of the augmented Lagrangian-SQP method to optimal control problems for the stationary Burgers equation, *Comput. Optim. Appl.*, 16(1), pp. 57–81, 2000.
- [132] S. Volkwein, Distributed control problems for the Burgers equation, *Comput. Optim. Appl.*, 18(2), pp. 115–140, 2001.
- [133] S. Volkwein, Proper orthogonal decomposition: Applications in optimization and control, <http://www.math.uni-konstanz.de/numerik/personen/volkwein/teaching/Lecture-Notes-Volkwein.pdf>, 2008, lecture Notes, CEA-EDF-INRIA Summer School Numerical Analysis Summer School Model Reduction and Reduced Basis methods: Application in Optimization. Saint-Lambert-des-Bois, France, June 23 to July 4, 2008, 74 pages.
- [134] T. Warburton and J. Hesthaven, On the constants in hp -finite element trace inverse inequalities, *Comput. Methods Appl. Mech. Engrg.*, 192, pp. 2765–2773, 2003.
- [135] S. Weller and S. Basting, Efficient preconditioning of variational time discretization methods for parabolic Partial Differential Equations, *ESIAM M2AN*, 49(2), pp. 331–347, 2015.
- [136] T. Werder, K. Gerdes, D. Schötzau, and C. Schwab, hp -discontinuous Galerkin time stepping for parabolic problems, *Comput. Methods Appl. Mech. Engrg.*, 190(49-50), pp. 6685–6708, 2001.
- [137] N. Yan and Z. Zhou, A priori and a posteriori error estimates of streamline diffusion finite element method for optimal control problem governed by convection dominated diffusion equation, *Numer. Math. Theor. Meth. Appl.*, 1(3), pp. 297–320, 2008.
- [138] M. Yano, A. T. Patera, and K. Urban, A space-time hp -interpolation-based certified reduced basis method for Burgers’ equation, *Math. Models Methods Appl. Sci.*, 24(9), pp. 1903–1935, 2014.
- [139] H. Yücel, M. Heinkenschloss, and B. Karasözen, Distributed optimal control of diffusion-convection-reaction equations using discontinuous Galerkin methods, in A. Cangiani, R. L. Davidchack, E. Georgoulis, A. N. Gorban, J. Levesley, and M. V. Tretyakov, editors, *Numerical Mathematics and Advanced Applications 2011*, pp. 389–397, Springer Berlin Heidelberg, 2013.
- [140] H. Yücel and B. Karasözen, Adaptive symmetric interior penalty Galerkin (SIPG) method for optimal control of convection diffusion equations with control constraints, *Optimization*, 2013.
- [141] Z. Zhou and N. Yan, The local discontinuous Galerkin method for optimal control problem governed by convection diffusion equations, *Int. J. Numer. Anal. Model.*, 7(4), pp. 681–699, 2010.

- [142] Z. Zhou, X. Yu, and N. Yan, Local discontinuous Galerkin approximation of convection-dominated diffusion optimal control problems with control constraints, *Numer. Methods Partial Differential Equations*, 30(1), pp. 339–360, 2014.
- [143] J.-M. Zokagoa and A. Soulaïmani, Low-order modelling of shallow water equations for sensitivity analysis using proper orthogonal decomposition, *Int. J. Comput. Fluid Dyn.*, 26(5), pp. 275–295, 2012.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Tuğba Akman
Nationality: Turkish
E-mail tr.tugba.akman@gmail.com

EDUCATION

Degree	Institution	Year of Graduation
M.S.	Scientific Computing, IAM, METU	2011
B.S.	Dept. of Mathematics, METU	2009
High School	İnebolu High School	2004

AWARDS

- 2013 METU Graduate Courses Performance Award (*The most successful student in the PhD Program of the Dept. of Scientific Computing in 2011-2012 Academic Year*)
- 2011 METU Graduate Courses Performance Award (*The most successful student in the MS Program of the Dept. of Scientific Computing in 2009-2010 Academic Year*)

Language skills

Turkish (native), English (fluently)

Computer skills

Proficient in MATLAB, L^AT_EX

Work experience

September 2009 - Present Research Assistant, Dept of Mathematics, METU

Teaching experience

- MATH 118 Calculus II, Spring 2015
- MATH 117 Calculus I, Fall 2014
- MATH 119 Calculus With Analytic Geometry, Fall 2009, 2010, 2011, 2012, Spring 2012, Fall 2013
- Math 120 Calculus for Functions of Several Variables, Spring 2009, 2010, 2011, 2013

PUBLICATIONS

Refereed journal articles (*indexed in SCI*)

- T. Akman *Local improvements to reduced-order approximations of optimal control problems governed by diffusion-convection-reaction equation*, Comput. Math. Appl., 70(2) (2015) 104-131
- T. Akman and B. Karasözen, *Variational time discretization methods for optimal control problems governed by diffusion convection reaction equations*, J. Comput. Appl. Math., 272 (2014) 41–56.
- T. Akman, H. Yücel and B. Karasözen, *A priori error analysis of the upwind symmetric interior penalty Galerkin (SIPG) method for the optimal control problems governed by unsteady convection diffusion equations*, Comput. Optim. Appl. 57 (2014) 703–729.

Accepted book articles (*Refereed*)

- T. Akman and B. Karasözen, *Space-Time Discontinuous Galerkin Methods for Optimal Control Problems Governed by Time Dependent Diffusion-Convection-Reaction Equations*, In Multiple Shooting and Time Domain Decomposition Methods, Contributions in Mathematical and Computational Sciences, T. Carraro, M. Geiger, S. Körkel, R. Rannacher (editors), Springer Verlag, (2015) *To appear*.

Refereed international conference proceedings

- T. Akman and B. Karasözen, *Reduced order optimal control using proper orthogonal decomposition sensitivities*, Numerical Mathematics and Advanced Applications - ENUMATH 2013, Lecture Notes in Computational Science and Engineering, Vol. 103, Springer International Publishing Switzerland, 409–417, 2015.

Papers submitted to international journals

- T. Akman, *Sensitivity analysis approach for reduced-order approximations of optimal control problems governed by Burgers equation*, 2014.
- T. Akman, *Error estimates for space-time discontinuous Galerkin formulation based on proper orthogonal decomposition*, 2014.

INTERNATIONAL SCIENTIFIC MEETINGS

- Conference on Computational and Experimental Science and Engineering, 25-29 October 2014, Antalya, TURKEY
Contributed talk: Variational Time Discretization Methods for Semilinear Parabolic Optimal Control Problems
- Conference on Computational and Experimental Science and Engineering, 25-29 October 2014, Antalya, TURKEY
Contributed talk: Proper Orthogonal Decomposition Sensitivities for Schlögl Model
- 3rd International Eurasian Conference on Mathematical Sciences and Applications, 25-28 August 2014, Vienna, AUSTRIA
Contributed talk: Local Improvements to Reduced-Order Approximations for Optimal Control of Burgers Equation
- 3rd International Eurasian Conference on Mathematical Sciences and Applications, 25-28 August 2014, Vienna, AUSTRIA
Contributed talk: Variational Time Discretization Methods for Linear-Quadratic Optimal Control Problems
- The European Conference on Numerical Mathematics and Advanced Applications, 26-30 August 2013, Lausanne, SWITZERLAND
Contributed talk: Reduced Order Optimal Control of Diffusion Convection Reaction Equation Using Proper Orthogonal Decomposition
- 3rd European Conference on Computational Optimization, 17-19 July 2013, Chemnitz, GERMANY
Contributed talk: Space-Time Discontinuous Galerkin Solution of Convection Dominated Optimal Control Problems
- Workshop on "Multiple Shooting and Time Domain Decomposition Methods", 6-8 May 2013, Heidelberg, GERMANY
Contributed talk: Optimal Control of Diffusion-Convection-Reaction Equations Using Space-Time Discontinuous Galerkin Methods
- International Conference on Applied and Computational Mathematics(ICACM), 3-6 October 2012, Ankara, TURKEY
Contributed talk: All-at-once Solution of Convection Dominated Time Dependent Optimal Control Problems by Discontinuous Galerkin Methods

- Workshop on "Adaptivity and Model Order Reduction in PDE Constrained Optimization", 23-27 July 2012, Hamburg, GERMANY
Poster presentation
- Summer school on "Optimal Control of Partial Differential Equations", 12-17 July 2010, Cortona, ITALY