

OPTIMAL CONTROL AND REDUCED ORDER MODELLING OF
FITZHUGH-NAGUMO EQUATION

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

BY

TUĞBA KÜÇÜKSEYHAN

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

MAY 2017

Approval of the thesis:

**OPTIMAL CONTROL AND REDUCED ORDER MODELLING OF
FITZHUGH–NAGUMO EQUATION**

submitted by **TUĞBA KÜÇÜKSEYHAN** 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, **Department of Mathematics & Institute of Applied
Mathematics, METU** _____

Assist. Prof. Dr. Murat Uzunca
Co-supervisor, **Department of Industrial Engineering, UTAA** _____

Examining Committee Members:

Prof. Dr. Nevzat G. Gençer
Department of Electrical and Electronics Engineering, METU _____

Prof. Dr. Bülent Karasözen
Department of Mathematics & Institute of Applied Mathematics,
METU _____

Assoc. Prof. Dr. Serdar Göktepe
Department of Civil Engineering, METU _____

Prof. Dr. Tamil Ergenç
Department of Mathematics, Atılım University _____

Assoc. Prof. Dr. Ayhan Aydın
Department of Mathematics, Atılım 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 KÜÇÜKSEYHAN

Signature :



ABSTRACT

OPTIMAL CONTROL AND REDUCED ORDER MODELLING OF FITZHUGH–NAGUMO EQUATION

Küçükseyhan, Tuğba

PhD, Department of Scientific Computing

Supervisor : Prof. Dr. Bülent Karasözen

Co-Supervisor : Assist. Prof. Dr. Murat Uzunca

May 2017, 112 pages

In this thesis, we investigate model order reduction and optimal control of FitzHugh-Nagumo equation (FHNE). FHNE is coupled partial differential equations (PDEs) of activator-inhibitor types. Diffusive FHNE is a model for the transmission of electrical impulses in a nerve axon, whereas the convective FHNE is a model for blood coagulation in a moving excitable media.

We discretize these state FHNEs using a symmetric interior penalty Galerkin (SIPG) method in space and an average vector field (AVF) method in time for diffusive FHNE. For time discretization of the convective FHNE, we use a backward Euler method. The diffusive FHNE has a skew-gradient structure. We show that the fully discrete energy of the diffusive FHNE satisfying the mini-maximizing property of the discrete energy of the skew-gradient system is preserved by SIPG-AVF discretization. Depending on the parameters and the non-linearity, specific patterns in one and two dimensional FHNEs occur like travelling waves and Turing patterns. Formation of fronts and pulses for the one dimensional (1D) diffusive FHNE, patterns and travelling waves for the two dimensional (2D) diffusive and convective FHNEs are studied numerically.

Because the computation of the pattern formations is very time consuming, we apply three different model order reduction (MOR) techniques; proper orthogonal decomposition (POD), discrete empirical interpolation (DEIM), and dynamic mode decomposition (DMD). All these MOR techniques are compared with the high fidelity fully

discrete SIPG-AVF solutions in terms of accuracy and computational time. Due to the local nature of the discontinuous Galerkin (DG) method, the nonlinear terms can be computed more efficiently by DEIM and DMD than for the continuous finite elements method (FEM). The numerical results reveal that the POD is the most accurate, the DMD the fastest, and the DEIM in between both.

We also investigate sparse and non-sparse optimal control problems governed by the travelling wave solutions of the convective FHNE. We also show numerically the validity of the second order optimality conditions for the local solutions of the sparse optimal control problem for vanishing Tikhonov regularization parameter. Further, we estimate the distance between the discrete control and associated local optima numerically by the help of the perturbation method and the smallest eigenvalue of the reduced Hessian. We use the DMD as an alternative method to DEIM in order to approximate the nonlinear term in the convective FHNE. Applying the POD-DMD Galerkin projection gives rise to a linear discrete equation for the activator, and the discrete optimal control problem becomes convex. FOM and sub-optimal control solutions with the above mentioned MOR techniques are compared for a variety of numerical examples.

Keywords : FitzHugh-Nagumo equations, Turing patterns, travelling waves, discontinuous Galerkin method, average vector field method, proper orthogonal decomposition, discrete empirical interpolation, dynamic mode decomposition, (sparse) optimal control

ÖZ

FITZHUGH-NAGUMO DENKLEMİNİN ENİYİLEMELİ KONTROLÜ VE İNDİRGENMİŞ DERECELİ MODELLEMESİ

Küçükseyhan, Tuğba

Doktora, Bilimsel Hesaplama Bölümü

Tez Yöneticisi : Prof. Dr. Bülent Karasözen

Ortak Tez Yöneticisi : Yrd. Doç. Dr. Murat Uzunca

Mayıs 2017, 112 sayfa

Bu tezde, FitzHugh-Nagumo denkleminin (FHNE) indirgenmiş dereceli modeli (MOR) ve eniyilemeli kontrolü araştırılmıştır. FHNE aktifleştirici-yavaşlatıcı tiplerinin birleştirilmiş kısmi diferansiyel denklemleridir (PDEs). Difüzyif FHNE bir sinir aksonunda elektrik dürtülerinin iletimi için bir model iken, konvektif FHNE hareketli uyarılabilir ortamdaki kan pıhtılaşması için bir modeldir.

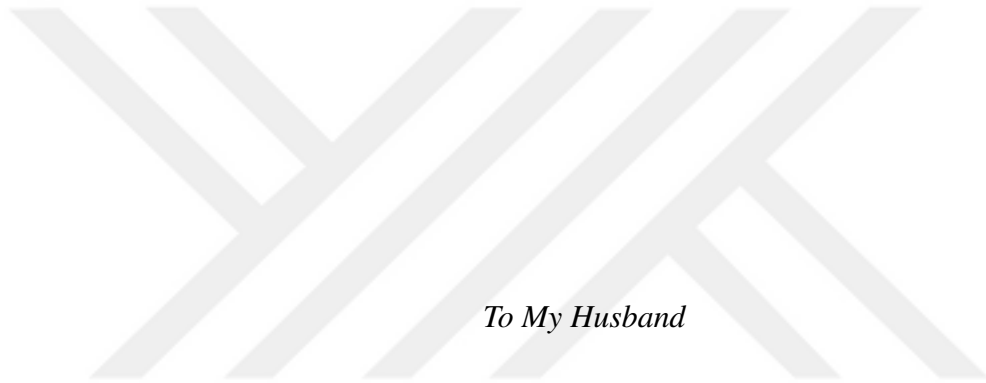
FHNE'ler uzayda simetrik süreksiz elemanlar Galerkin (SIPG) yöntemi ve zamanda difüzyif FHNE için ortalama vektör alanı (AVF) yöntemi kullanılarak ayrıştırıldı. Konvektif FHNE'yi zamanda ayrıştırmak için, geriye dönük Euler yöntemi kullanıldı. Difüzyif FHNE çarpıklık gradyanı yapısına sahiptir. Çarpıklık gradyan sisteminin tam olarak ayrıştırılmış enerjisinin mini-maksimize edici özelliğini karşılayan difüzyif FHNE'nin tamamen ayrık enerjisinin SIPG-AVF ayrıklaştırması ile korunduğu gösterildi. Reaksiyon ve difüzyon parametreleri arasındaki ilişkiye ve doğrusal olmayan terime bağlı olarak, bir ve iki boyutlu FHNE'lerde hareketli dalgalar ve Turing desenleri gibi özel desenler görülmektedir. Bir boyutta difüzyif FHNE için cephe ve darbe oluşumu, iki boyutta difüzyif ve konvektif FHNE'ler için desenler ve hareketli dalgalar sayısal olarak incelenmiştir.

Desen oluşumlarının hesaplanması çok zaman alıcı olduğundan, üç farklı MOR tekniği uygulandı: öz dik ayrışım (POD), ayrık ampirik enterpolasyon (DEIM) ve dinamik mod ayrıştırma (DMD). Tüm bu MOR teknikleri, doğruluk ve hesaplama süresi bakımın-

dan yüksek güvenilirliğe sahip tamamen ayrık SIPG-AVF çözümleri ile karşılaştırılmıştır. Süreksiz Galerkin (DG) yönteminin yerel doğası gereği doğrusal olmayan terimler, DEIM ve DMD tarafından sürekli sonlu elemanlar yöntemi (FEM) için olduğundan daha verimli hesaplanabilir. Sayısal sonuçlar, POD'yi en doğru, DMD'yi en hızlı ve DEIM'i her ikisi arasında göstermektedir.

Ayrıca, konvektif FHNE'nin hareketli dalga çözümleri tarafından yönetilen seyrek ve seyrek olmayan eniyilemeli kontrol problemlerini araştırılmıştır. Sıfıra yaklaşan Tikhonov normalleştirme parametresinin seyrek optimal kontrol probleminin yerel çözümleri için ikinci dereceden optimallik koşullarının sayısal olarak geçerliliğini de gösterildi. Dahası, pertürbasyon yöntemi ve indirgenmiş Hessian'ın en küçük özdeğeri ile ayrık kontrol ve ilişkili yerel optima arasındaki mesafe sayısal olarak tahmin edildi. Konvektif FHN denklemindeki doğrusal olmayan terime yaklaşmak için DEIM'e alternatif bir yöntem olarak DMD kullanıldı. POD-DMD Galerkin projeksiyonunun uygulanması, aktivatör için doğrusal bir ayrık denklem oluşmasına neden olur ve ayrık eniyilemeli kontrol problemi konveks olur. Yukarıda belirtilen MOR teknikleriyle FOM ve alt iyilemeli kontrol çözümleri çeşitli sayısal örnekler için karşılaştırılmıştır.

Anahtar Kelimeler: FitzHugh-Nagumo denklemleri, Turing desenleri, hareketli dalgalar, kesintili Galerkin yöntemleri, ortalama vektör alanı yöntemi, öz dik ayrışım, ayrık ampirik enterpolasyon, dinamik mod ayrıştırma, (seyrek) eniyilemeli kontrol problemleri



To My Husband



ACKNOWLEDGMENTS

It is a great pleasure to thank the many people who made this thesis possible.

I would like to begin by acknowledging my supervisor Prof. Dr. Bülent Karasözen for his guidance, patience, encouragement, and enormous support. He has given me numerous unvaluable oppurtunities which genuinely had any kind of effect in my scholarly life. I also want to express my gratitude to my co-supervisor Assist. Prof. Dr. Murat Uzunca for his invaluable help during my graduate education. Without help of them, this thesis would not have been possible.

I acknowledge Assist. Prof. Dr. Hamdullah Yücel for the discussion and help about optimal control problems.

I wish to thank Prof. Dr. Nevzat G. Gençer and Assoc. Prof. Dr. Serdar Göktepe for serving on my thesis committee and useful comments and suggestions about our studies.

I acknowledge all members of the Mathematics Department and Institute of Applied Mathematics at Middle East Technical University for providing me pleasant and lovely atmosphere.

I acknowledge Prof. Dr. Fredi Tröltzsch for kind hosting and helpful discussions on my studies in Technical University of Berlin. I also acknowledge Peter Nestler for their kind and helpful hosting in Berlin.

I also acknowledge the financial support of Turkish Scientific and Technical Research Council (TÜBİTAK) for its financial support during my PhD in Turkey and during a half year period in Berlin.

A very special thank goes to my family for their support. And last but not least, I express all my gratitude to my husband for his encouragement during my graduate education and for continually being a major part of my life.



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 FITZHUGH-NAGUMO EQUATIONS	7
2.1 Pattern Formation in Diffusive FitzHugh-Nagumo Equation	9
2.1.1 Pattern Formation in 1D	12
2.1.2 Pattern Formation in 2D	13
2.1.3 Three Component FitzHugh-Nagumo System	14
2.1.4 Skew-Gradient Structure of the Diffusive FHNE	16
2.2 Pattern Formation in Convective FitzHugh-Nagumo Equation	17
2.2.1 Parameter Analysis for Wave Type Solutions of the Convective FHNE	17
2.3 Space-Time Discretization	18

2.3.1	Space Discretization	19
2.3.2	Time Discretization: Average Vector Field Method	22
2.4	Preservation of the Discrete Energy for the Diffusive FHNE .	24
2.5	Numerical Results	27
2.5.1	Diffusive FHNE in 1D	27
2.5.1.1	Travelling Fronts	27
2.5.1.2	Travelling Pulses	29
2.5.2	Diffusive FHNE in 2D	30
2.5.2.1	Spot Patterns	30
2.5.2.2	Labyrinth-like Patterns	31
2.5.3	Three Component Diffusive FHNE in 1D	31
2.5.3.1	One-pulse	31
2.5.3.2	Two-pulse	33
2.5.3.3	Multi-pulse	34
2.5.4	Convective FHNE in 2D	35
3	REDUCED ORDER MODELLING	37
3.1	Proper Orthogonal Decomposition	38
3.2	Discrete Emprical Interpolation	43
3.3	Dynamic Mode Decomposition	47
3.3.1	DMD Modes with Optimal Amplitudes	50
3.3.2	POD with DMD	51
3.4	Randomized Singular Value Decomposition	52
3.5	Numerical Results	53

	3.5.1	Diffusive FHNE	54
	3.5.2	Parametrized Model Order Reduction for Diffusive FHNE	59
	3.5.3	Convective FHNE	61
4		OPTIMAL CONTROL AND SUB-OPTIMAL CONTROL USING REDUCED ORDER MODELLING	67
	4.1	Preliminaries	69
	4.2	Full Order Optimal Control	71
	4.2.1	Well-posedness of the State Equation	72
	4.2.2	First Order Optimality Conditions	77
	4.2.3	Second Order Optimality Conditions	79
	4.3	Space-Time Discretization of the Optimality System	81
	4.3.1	Space Discretization of the State and Adjoint Equations	81
	4.3.2	Fully Discrete Form of State and Adjoint Equations	83
	4.4	Reduced Order Optimal Control	84
	4.4.1	POD Approach for the State and Adjoint System .	84
	4.4.2	Discrete Empirical Interpolation	85
	4.4.3	Dynamic Mode Decomposition	85
	4.5	Numerical Results	86
	4.5.1	FOM Results for OCP of Convective FHNE	88
	4.5.1.1	Optimal Control in the Space-Time Domain	88
	4.5.1.2	Terminal Control	91
	4.5.2	ROM Results for OCP of Convective FHNE	93

5	CONCLUSION AND DISCUSSION	99
	REFERENCES	101
	CURRICULUM VITAE	109



LIST OF FIGURES

Figure 1.1 Patterns in nature.	1
Figure 1.2 Blood coagulation.	4
Figure 2.1 Travelling front solutions for both state component u_1, u_2 (top), evolution of the discrete energy (bottom).	28
Figure 2.2 Travelling front solutions at different time instances.	28
Figure 2.3 Travelling pulse solutions for both state component u_1, u_2 (top), evolution of the discrete energy (bottom).	29
Figure 2.4 Travelling pulse solutions at different time instances.	29
Figure 2.5 Spot patterns solutions for both state component u_1, u_2 (top), evolution of the discrete energy (bottom).	30
Figure 2.6 Labyrinth-like patterns solutions for both state component u_1, u_2 (top), evolution of the discrete energy (bottom).	31
Figure 2.7 One-pulse solutions for three state components u_1, u_2, u_3 , evolution of the discrete energy.	32
Figure 2.8 One-pulse solutions at different time instances.	32
Figure 2.9 Two-pulse solutions for three state components u_1, u_2, u_3 , evolution of the discrete energy.	33
Figure 2.10 Two-pulse solutions at different time instances.	33
Figure 2.11 Multi-pulse solutions for three state components u_1, u_2, u_3 , evolution of the discrete energy.	34
Figure 2.12 Multi-pulse solutions at different time instances.	34
Figure 2.13 Wave-type pattern solutions for state u_1 at different time instances.	35
Figure 2.14 Wave-type pattern solutions for state u_2 at different time instances.	36
Figure 3.1 Connectivity of degrees of freedoms for linear basis	46
Figure 3.2 FOM solutions for the component u_1 (left), u_2 (right).	54

Figure 3.3	Decay of the singular values for u_1 , u_2 and the non-linearity f . . .	54
Figure 3.4	From top to bottom POD, DEIM, DMD solutions (left) and errors (right) for the component u_1	55
Figure 3.5	From top to bottom POD, DEIM, DMD solutions (left) and errors (right) for the component u_2	56
Figure 3.6	Energy plots.	57
Figure 3.7	Energy errors.	57
Figure 3.8	(Top) FOM-ROM $L^2(Q)$ errors for the states u_1 (left), u_2 (right) with increasing number of POD for POD solution, and increasing number of DEIM and DMD basis functions with 20 fixed POD basis functions for POD-DEIM and POD-DMD solutions; (bottom) corresponding CPU times.	58
Figure 3.9	(Top) FOM-ROM $L^2(Q)$ errors for the states u_1 (left), u_2 (right) with the same increasing number of POD, DEIM and DMD basis functions; (bottom) corresponding CPU times.	59
Figure 3.10	(Left) Decay of the singular values of solution snapshots \mathcal{U} , \mathcal{V} and of the nonlinear snapshots \mathcal{F} ; (right) CPU times for the computation of FOMs, POD and POD-EIM ROMs for the parameter value $\mu = 0.03$	60
Figure 3.11	Pattern solutions for u_1 at the final time $T = 1000$ with FOM (left), POD (middle) and POD-DEIM (right) for the parameter values $\mu \in \{-0.03, -0.01, 0.01, 0.03\}$ from top to bottom.	61
Figure 3.12	FOM solutions for the component u_1 (left), u_2 (right).	62
Figure 3.13	FOM solutions for the component u_1 (left), u_2 (right).	62
Figure 3.14	From top to bottom POD, DEIM, DMD solutions (left) and errors (right) for the component u_1	63
Figure 3.15	From top to bottom POD, DEIM, DMD solutions (left) and errors (right) for the component u_2	64
Figure 3.16	(Top) FOM-ROM $L^2(Q)$ errors for the states u_1 (left), u_2 (right) with increasing number of POD for POD solution, and increasing number of DEIM and DMD basis functions with 15 fixed POD basis functions for POD-DEIM and POD-DMD solutions; (bottom) corresponding CPU times.	65
Figure 3.17	(Top) FOM-ROM $L^2(Q)$ errors for the states u_1 (left), u_2 (right) with the same increasing number of POD, DEIM and DMD basis functions; (bottom) corresponding CPU times.	66
Figure 4.1	Optimal controls u (top) and associated states u_1 (bottom) at $t = 0.75$ without sparse control for $V_{\max}=16, 32, 64, 128$ (from left to right).	89

Figure 4.2	Optimal controls u (top) and associated states u_1 (bottom) for the sparsity parameters $\mu = 1/500, 1/100, 1/50, 1/35$ (from left to right) and $V_{\max} = 32$ at $t = 0.75$	90
Figure 4.3	$L^2(Q)$ errors for $\ \bar{u}_{1\lambda} - \bar{u}_{1\text{ref}}\ $, $\ \bar{u}_{2\lambda} - \bar{u}_{2\text{ref}}\ $, and $\ \bar{u}_\lambda - \bar{u}_{\text{ref}}\ $ for the sparsity parameter $\mu = 0$ (left), $\mu = 1/100$ (right).	91
Figure 4.4	Optimal controls u (top) and associated states u_1 (bottom) for the sparse parameters $\mu = 1/500, 1/100, 1/50, 1/35$ (from left to right) and $V_{\max} = 32$ at $t = 0.75$	92
Figure 4.5	Optimal states u_1 and u_2 (top), optimal control (bottom) for the full system at the final time.	94
Figure 4.6	Decay of the singular values for u_1, u_2 and the nonlinearity f	94
Figure 4.7	Errors of POD, POD-DEIM and POD-DMD with 8 basis functions (from top to bottom): optimal control errors u (left), associated states u_1 (middle), u_2 (right) errors at the final time.	95
Figure 4.8	Errors for states u_1, u_2 (top), and control u (bottom) in the whole domain with increasing number of modes.	96
Figure 4.9	Errors for states u_1, u_2 (top), and control u (bottom) at the final time with increasing number of modes.	97
Figure 4.10	Objective function values (left) and CPU times (right) with increasing number of modes.	97



LIST OF TABLES

Table 3.1 FOM-ROM errors & speed-up with 20 POD, 30 DEIM, and 70 DMD basis functions.	57
Table 3.2 The computation times (in sec), speed-up factors S_{POD} and S_{DEIM} , and the DEIM projection error bounds $\ (P^T W)^{-1}\ _2$	60
Table 3.3 FOM-ROM errors & speed-up with 15 POD, 15 DEIM, and 21 DMD basis functions.	64
Table 4.1 Cost functional J , number of nonlinear CG iterations, line searches, Newton steps without sparse control.	89
Table 4.2 Optimal values of I , μ^j , and number of iterations, line searches, and Newton steps for $V_{max} = 32$	90
Table 4.3 $L^2(Q)$ errors for $\ \bar{u}_{1\lambda} - u_{1Q}\ $ and $\ \bar{u}_{2\lambda} - u_{2Q}\ $	91
Table 4.4 Optimal value of I , μ^j , and number of nonlinear CG iterations, line searches, and Newton steps for $V_{max} = 64$	92
Table 4.5 Numerical errors of $\ u - u_h\ $ and error estimates $\frac{2}{\delta}\ \zeta\ $ for sparse controls with $\mu = 1/200$	93
Table 4.6 Optimal value of cost functional J , average number of nonlinear CG iterations, line searches, Newton steps, and speed-up's factors.	96



LIST OF ABBREVIATIONS

AVF	Average Vector Field
CG	Conjugate Gradient
DEIM	Discrete Empirical Interpolation Method
DG	Discontinuous Galerkin
DMD	Dynamic Mode Decomposition
DoF	Degree of Freedom
FEM	Finite Elements Method
FHNE	FitzHugh–Nagumo Equation
FOM	Full Order Model
MOR	Model Order Reduction
OCP	Optimal Control Problem
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
POD	Proper Orthogonal Decomposition
rSVD	Randomized Singular Value Decomposition
ROM	Reduced Order Model
SIPG	Symmetric Interior Penalty Galerkin
SSCs	Second-order Sufficient Optimality Conditions
SVD	Singular Value Decomposition



CHAPTER 1

INTRODUCTION

Spatially and temporally varying structures in form of patterns, travelling waves, fronts, periodic pulses are observed in variety of physical, chemical and biological systems far from equilibrium, for instance, pattern formation. In nature, pattern formation has a wide variety in many living and non-living systems. Just think about some examples of this formation like stripes on zebras, spots on leopards, patterns on fishes, or hexagons on honeycombs. There are also examples that can only be seen with a microscope like the patterns in the connectivity of neurons or fingerprints. Those examples are shown in Figure 1.1 (Source: <https://pixabay.com/tr>).

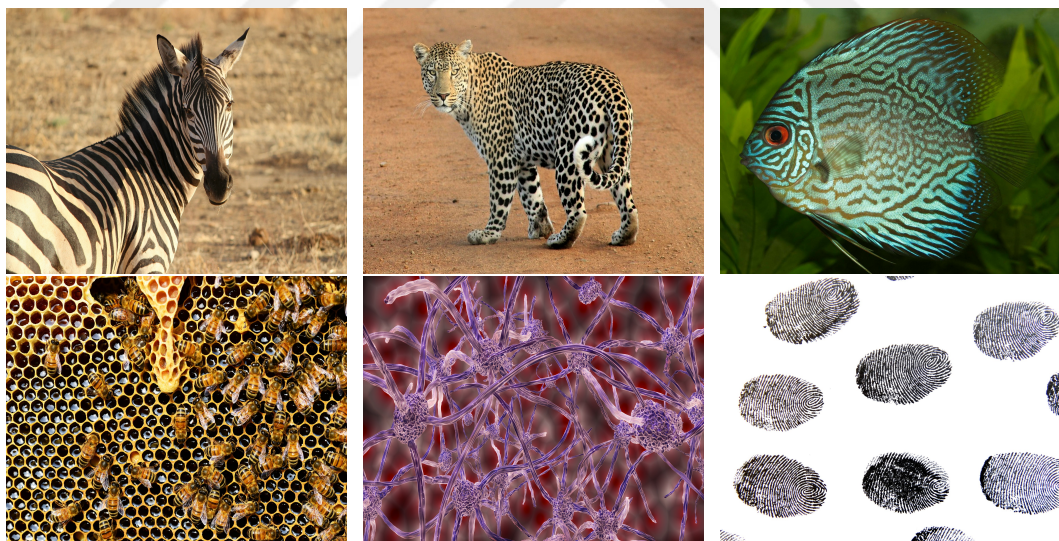


Figure 1.1: Patterns in nature.

Despite the great variability, they exhibit many common properties. Therefore, use of the simplest models plays a vital role in understanding the mechanisms of propagation in excitable, oscillatory, or bistable media. In 1952, a British mathematician Alan Turing explained this phenomena in his the only article, “The chemical basis of morphogenesis” [91]. According to this, the models explaining these phenomena are

usually in form of nonlinear system of reaction-diffusion equations:

$$\begin{aligned}\frac{\partial u_1}{\partial t} &= d_1 \Delta u_1 + f(u_1, u_2), \\ \frac{\partial u_2}{\partial t} &= d_2 \Delta u_2 + g(u_1, u_2).\end{aligned}\tag{1.1}$$

Alan Turing showed that under certain conditions any stable homogeneous equilibrium solution of (1.1) without diffusion terms can be converted into an unstable state of system (1.1). This phenomena is known as "*Turing instability*" or "*diffusion driven instability*". These conditions for system (1.1) are given by

$$\begin{aligned}\frac{\partial f}{\partial u_1} + \frac{\partial g}{\partial u_2} &< 0, \\ \frac{\partial f}{\partial u_1} \cdot \frac{\partial g}{\partial u_2} - \frac{\partial f}{\partial u_2} \cdot \frac{\partial g}{\partial u_1} &> 0, \\ d_2 \frac{\partial f}{\partial u_1} + d_1 \frac{\partial g}{\partial u_2} &> 0, \\ \left(d_2 \frac{\partial f}{\partial u_1} + d_1 \frac{\partial g}{\partial u_2} \right)^2 &> 4d_1 d_2 \left(\frac{\partial f}{\partial u_1} \cdot \frac{\partial g}{\partial u_2} - \frac{\partial f}{\partial u_2} \cdot \frac{\partial g}{\partial u_1} \right).\end{aligned}$$

One of the simplest and widely used models for pattern formation is FitzHugh-Nagumo equation (FHNE) [67]. The FHNE is an activator-inhibitor system, that is, it consists of two substances, an inhibitor and an activator. While the activator facilitates and enhances its own production, the inhibitor suppresses the production of the activator. Therefore, patterns of activator and inhibitor concentrations result from fast diffusion of the inhibitor or slow diffusion of the activator. In the literature, the most known type of FHNE is classical FHNE consisting of a partial differential equation (PDE) with a non-monotone nonlinear term for the activator u_1 and an ordinary differential equation (ODE) for the inhibitor u_2 [36, 69]:

$$\begin{aligned}u_{1t} &= d_1 \Delta u_1 - f(u_1) - v_2, \\ u_{2t} &= -\beta u_2 + \gamma u_1 - \delta,\end{aligned}\tag{1.2}$$

where d_1 is a diffusion coefficient of u_1 , γ , δ are constants, and $f(u_1) = k(u_1 - c_1)(u_1 - c_2)(u_1 - c_3)$ with $k > 0$ and $c_1 < c_2 < c_3$.

In this thesis, we study two different types of the classical FHNE (1.2), which are not intensively investigated in the literature. The first model which we are interested in is the diffusive FHNE consisting of two PDEs:

$$\begin{aligned}\tau_1 u_{1t} &= d_1 \Delta u_1 + f(u_1) - u_2 + \kappa, \\ \tau_2 u_{2t} &= d_2 \Delta u_2 + u_1 - \gamma u_2 + \epsilon\end{aligned}\tag{1.3}$$

with homogeneous (zero-flux) Neumann boundary conditions and suitable initial conditions in a smooth bounded domain Ω in $\mathbb{R}^n (n \leq 2)$, and on a time period $[0, T]$ with $u_1 = u_1(x, t)$ and $u_2 = u_2(x, t)$, $x \in \Omega$, $t \in [0, T]$. Here τ_1 , τ_2 and d_1 , d_2 correspond to time scales and diffusion coefficients of u_1 and u_2 , respectively. Here, κ , γ , ϵ are

constants and $f(u_1) = u(u - \beta)(1 - u)$ is a bistable nonlinearity with suitably chosen a constant β .

FHNE (1.2)-(1.3) was proposed for modelling the electrical impulses in a nerve axon [36, 69]. Neurons, the building blocks of the central nervous system, are highly complex dynamical systems. The diffusive FHNE as the simplified version of the Hodgkin–Huxley equation models in a detailed manner activation and deactivation dynamics of a spiking neurons. The FHN equation can describe the bifurcations with the variation of the key parameters for neuron dynamics. Within time FHNE became a favorite model for simulation of wave propagation in excitable media, such as heart tissue or nerve fiber. Understanding the complex behaviour of patterns in neuroscience can have a big impact for dealing and preventing with various diseases.

It is not possible to explain all patterns in nature by a system of reaction-diffusion equations, considering, for instance, blood coagulation process, shown in Figure 1.2 (Source: <http://www.dovemed.com/common-procedures/procedures-laboratory/coagulation-factors-blood-test/>). It is essential to use a convective term. So, the second model equation which we focus on in this thesis is the convective FHNE:

$$\begin{aligned}\tau_1 u_{1t} &= d_1 \Delta u_1 - \mathbf{V} \cdot \nabla u_1 + f(u_1) - u_2, \\ \tau_2 u_{2t} &= d_2 \Delta u_2 - \mathbf{V} \cdot \nabla u_2 + \epsilon(c_3 u_1 - u_2)\end{aligned}\quad (1.4)$$

on a space time cylinder $Q_T := \Omega \times (0, T)$, where $\Omega = (0, L) \times (0, H)$ with zero flux boundary condition except the left boundary, which is Dirichlet type as $u_1 = u_2 = 0.1$. As initial conditions we set $u_1 = 0.1$ inside a narrow triangle $0 \leq x \leq 0.5$, $0 \leq y \leq H$, and $u_1 = 0$ outside this rectangle; $u_2 = 0$ everywhere. Here, d_1 and d_2 correspond to diffusion coefficients of u_1 and u_2 , respectively. $f(u_1) = c_1 u_1 (u_1 - c_2)(1 - u_1)$ is a monostable nonlinearity and c_1, c_2, c_3 are constants. The divergence free velocity field $\mathbf{V} = (V_{x_1}, V_{x_2})$ is given along the x_1 -direction with a parabolic profile

$$V_{x_1}(x_2) = ax_2(H - x_2), \quad V_{max} = \frac{1}{4}aH^2, \quad a > 0, \quad V_{x_2} = 0, \quad (1.5)$$

where V_{max} denotes the maximum wave speed of the velocity field.

The convective FHNE (1.4) has been proposed as a model for a wave propagation in blood coagulation and bioreactor systems. The flow plays an important role by the regularization of the excitation threshold and travelling wave propagation. There exists a large number of proteins, special blood cells, and platelets in blood coagulation process. The complex process of coagulation consists of cascadic enzymatic reactions and feedback loops, differentiated into three stages, initiation (localized at the vascular damage site); propagation, or spatial expansion of coagulation wave into the vessel; termination of the biochemical reactions and clot enlargement stoppage [59]. These ultimately catalyze fibrinogen conversion into fibrin, which polymerizes to form a clot. These reactions allows autocatalytic thrombin generation far from the damage site. The most important property of blood coagulation process is the formation of autowaves with the velocity independent of the initial conditions [35, 34, 59]. The flow propagates within the impermeable channel walls (Neumann boundary conditions). Different type of standing and triggering waves occur depending on the constants of the convective

FHNE [35]. The waves became more curved when the maximum wave speed of the velocity field (1.5) increases.

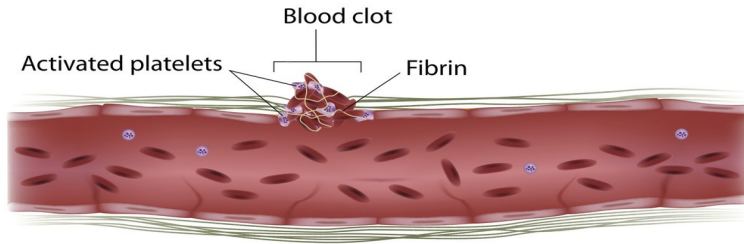


Figure 1.2: Blood coagulation.

Both types of the FHNE occur in pattern formation [61] and as waves in excitable media. [35]. So, it is quite natural and challengeable to investigate the behaviour of such systems.

Numerous numerical techniques have been developed for the space discretization of reaction-diffusion-convection equations, like finite difference, finite volume, spectral elements, continuous finite element, and discontinuous Galerkin (DG) finite element methods. We use DG finite element method of symmetric interior penalty Galerkin (SIPG) type [9, 74], using discontinuous polynomials for the discretization of PDEs. The DG discretization supports high order local approximations that can vary non-uniformly over the mesh. Therefore patterns, like travelling fronts, travelling pulses, and layers, can be captured better than the other methods. Besides, they are more stable than continuous finite element method (FEM) for convection dominated problems [11, 32].

As a special case of the diffusive FHNE, it is a skew-gradient system [101] coupling two energy dissipating systems. For time discretization we apply the energy preserving average vector field integrator (AVF) [27, 42]. The AVF method preserves the energy of conservative systems (e.g. the Hamiltonian) and the energy dissipation of gradient systems [27, 42]. We show that the AVF integrator combined with the SIPG space discretization preserves the mini-maximizing property of the skew-gradient systems [100] in the discrete form of the diffusive FHNE.

Discretization of PDEs with nonlinear terms like the FHNE lead to very large nonlinear system. The computation of pattern formations for different parameters are time consuming. In order to overcome this issue, we apply model order reduction (MOR) techniques; proper orthogonal decomposition (POD), with a discrete empirical interpolation (DEIM) [13, 28] and a dynamic mode decomposition (DMD) [60]. The goal in the MOR methods is to reduce the computational complexity and time of large-scale sized systems by approximating with lower dimension. The combination of DEIM as well as DMD with SIPG turns out to be very efficient approach for the computation of the nonlinear term in the reduced systems.

We also consider optimal control problems (OCPs) of the convective FHNE occurring in blood coagulation. Optimal control problems governing the classical FHNE and

its activator equation, called Schlögl or Nagumo equations, were investigated theoretically and numerically in [21, 23, 24, 77].

To control pattern solutions is really exciting, yet numerically it is difficult. Control problems require much more time than the numerical solution of the PDEs. Therefore, we apply the three MOR techniques POD, POD-DEIM and POD-DMD to optimal control of the convective FHNE. Reduced order modelling of linear and semi-linear PDE-constrained optimal control problems has been recently investigated in the literature intensively. [45, 56, 89, 49, 86, 53].

In this thesis, we obtain the following main results:

- The selection criteria for Turing pattern formation of the diffusive FHNE is given using the skew-gradient structure.
- The stability of the any steady state solutions of diffusive FHNE with skew-gradient system is proved using Yanagida's result [100], which states any steady state of a skew-gradient system is stable if and only if it is a mini-maximizer of the energy functional. According to this result, we proved that the space-time discrete form of skew-gradient diffusive FHNE by the SIPG-AVF methods preserves the mini-maximizing property of its discrete energy functional.
- We use the POD based and Galerkin projection MOR techniques with SIPG to solve the pattern formations of our model problems in an efficient way. To the best of our knowledge, besides the pattern formation in lambda-omega systems [65], there exist no other work concerning the reduced order modelling in pattern formation.
- The discontinuous Galerkin discretization is used first time in reduced order modelling as a computationally efficient method alternative to the classical FEM.
- We consider optimal control and sub-optimal control of convective FHNE occurring in blood coagulation. We compare three different reduced order approaches the POD, the POD with the DEIM, and the POD with DMD by comparing the accuracy and computational time of the reduced order sub-optimal solutions. To our best knowledge, reduced order OCPs using POD-DMD are not yet investigated in the literature.

The rest of the thesis is structured in the following form:

- In Chapter 2, we perform a parameter study using “*Turing instability*” analysis to determine formation of patterns for the diffusive FHNE. We prove that the AVF time integrator with the SIPG space discretization preserves the skew-gradient structure and the energy function of the diffusive FHNE. Numerical results show that the front and pulse solutions of the one dimensional diffusive FHNE with two and three components are computed accurately. For the 2D diffusive FHNE, the computed patterns are close to those in the literature.

- In Chapter 3, we present the three ROM techniques POD, POD-DEIM and POD-DMD with the SIPG discretization in space. The DEIM and DMD is the first time applied as a ROM technique in connection with the DG discretization. The SIPG discretization with the DEIM turns out to be more efficient than the classical continuous FEM, because the support of DG basis functions consist of only a single element, and vanishing outside that element. For the continuous FEM, DEIM points are connected by the neighbouring elements [5]. We also perform a parametric reduced order modelling for the diffusive FHNE. Comparison of the three ROM techniques show that the POD is the most accurate and DMD is the fastest.
- In Chapter 4, we derive optimality conditions for the OCP problem governed by a convective FHNE. The OCP problem is discretized using the so called *first optimize then discretize* approach using the backward Euler time discretization. We apply the nonlinear conjugate method [41] for solving the resulting discrete optimization problem. We present numerical results for sparse and non-sparse controls in the whole space-time domain and at the terminal time. We also investigate the effect of the Tikhonov parameter as it goes to zero as well as finding a bound for the distance between the local minima and discrete solution as done in [24, 77].

Finally, we apply POD, POD-DEIM, and POD-DMD for the sub-optimal control of the convective FHNE. The POD with the DMD is applied first time for reduced order OCPs. Again the speed-up and accuracy of the sub-optimal control results are compared for the POD, POD-DEIM, and POD-DMD.

The thesis ends with some conclusions and discussions for future work.

CHAPTER 2

FITZHUGH-NAGUMO EQUATIONS

Pattern formation is observed in many natural systems like biology, chemistry, and physics. Examples of patterns in real life are fingerprints, stripes on zebras, spots on leopards, pulses travelling inside nerve cells, spiral waves in sunflower and red cabbage, hexagons on honeycomb, geometric designs on seashells [67]. Reaction-diffusion equations are typical models for pattern formation. Alan Turing explained the formation of patterns resulting from the instability, called afterwards “*Turing instability*”, in reaction-diffusion equations in his seminal work on the chemical basis for morphogenesis in 1952 [91]. In general, the patterns are classified into three types according to the characteristics of media: excitable, Turing-Hopf (oscillatory), and bistable.

In this thesis, we study excitable and bistable pattern formation of FitzHugh-Nagumo equation (FHNE), which was first introduced by FitzHugh [36] and Nagumo et al. [69] for modelling the electrical impulses in a nerve axon as a simplified version of the Hodgkin-Huxley model. In spite of the fact that the FHNE is developed as a model in physiology, it is used as a generic model that exhibits many phenomena in excitable, oscillatory or bistable chemical media. The well-known classical FHNE consists of a partial differential equation (PDE) with a bistable cubic nonlinear term, known as activator equation, and an ordinary differential equation (ODE), known as inhibitor equation. We study here two other types of FHNE. The first one is diffusive FHNE consisting of an activator PDE with a bistable cubic-nonlinear term, and one or two linear PDEs as inhibitors [61, 93, 94]. Depending on the reaction and diffusion parameters and on the bistable nonlinearity, specific patterns occur in one dimensional FHNE, like travelling waves (e.g. fronts and pulses) [30], and in two dimensional FHNE, like spots or labyrinth-like patterns [61]. Fronts and pulses are the most well-known one dimensional waves in reaction-diffusion systems, connecting two different states of a reaction-diffusion system with bistable nonlinearity. In all these systems the time evolution of the energy plays an important role in formation of travelling waves, pulses and patterns. The other one is convective FHNE, proposed as a model for wave propagation in excitable medium blood coagulation and bioreactor systems [34, 35, 59]. Convection leads to complex wave phenomena, like pulses, triggering and autonomous waves in a moving excitable media.

For space discretization we use interior penalty discontinuous Galerkin finite element method [8, 74]. The original discontinuous Galerkin (DG) methods were first intro-

duced by Reed and Hill in order to solve first order linear hyperbolic neutron transport equation in 1973 [73] and developed in the following years by many authors [11, 12, 20, 32, 33, 72]. DG methods have a place between finite element and finite volume methods owning good features of both methods. Although continuous finite elements methods (FEMs) require continuity of the solution along element interfaces, DG methods do not require continuity of the solution along the interfaces. They can support high order local approximations that can vary non-uniformly over the mesh, by which fronts, pulses and layers can be captured better. In contrast to the stabilized continuous Galerkin FEM, DG methods produce stable discretization without the need for extra stabilization strategies for convection dominated problems, and damp the unphysical oscillations. Due to the local structure of DG discretization, the DG method is parallelizable and adaptive meshing techniques can be implemented efficiently. In addition, the boundary conditions are imposed weakly in DG methods.

For large class of reaction-diffusion systems there exist Lyapunov functions, so that all the solutions converge to equilibria. But it is difficult to derive a gradient structure for all reaction-diffusion systems, because the reaction terms contain quite general nonlinearities. For certain classes, gradient structures can be constructed. Among these are the activator-inhibitor type reaction-diffusion systems which are not order preserving and their linearized forms around the steady states are not self-adjoint. Yanagida [101] introduced the concept of skew-gradient systems to investigate the stability of these systems. It was shown in [101] under some restrictions on the parameters the diffusive FHNE exhibits a skew-gradient structure. In short, these activator-inhibitor diffusive FHNEs consist of two gradient systems. Nevertheless, for convective-diffusive equation, there do not exist Lyapunov function.

For time discretization of diffusive FHNE, we applied average vector field (AVF) method [27, 42] which preserves the energy for conservative systems, like Hamiltonian, and the energy dissipation for gradient systems. It is well suited for time discretization of the diffusive FHN with skew-symmetric gradient structure. It is second order in time and reduces to the mid-point rule for ODE with quadratic potentials. We show that the AVF integrator combined with the SIPG space discretization preserves the mini-maximizing property of the skew-gradient systems [100] in the discrete form of diffusive FHNE [50]. For convective FHNE, since it has no gradient structure we have used a semi-implicit Euler method as a time integrator.

In this chapter, we first give diffusive FHNE in one and two space dimensions, of two and three component type, and skew-gradient property in Section 2.1. In Section 2.2, convective FHNE is introduced. In Section 2.3, the space discretization by the DG method is presented. Then, the fully discrete form of the diffusive FHNE with the gradient stable AVF method in time and convective FHNE with the semi-implicit Euler method are derived. The energy analysis of the fully discrete scheme of the diffusive FHNE is given in Section 2.4. In Section 2.5, numerical simulations are shown for the diffusive and convective FHNEs exhibiting patterns in one or two dimensional domains.

2.1 Pattern Formation in Diffusive FitzHugh-Nagumo Equation

In this section, we study diffusive FHNE

$$\tau \mathbf{u}_t = \mathbf{D} \Delta \mathbf{u} + \mathbf{f}(\mathbf{u}), \quad (2.1)$$

where

$$\tau = \begin{bmatrix} \tau_1 & 0 \\ 0 & \tau_2 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}, \quad \mathbf{f}(\mathbf{u}) = \begin{bmatrix} f_1(u_1, u_2) \\ f_2(u_1, u_2) \end{bmatrix} \quad (2.2)$$

with homogeneous (zero-flux) Neumann boundary conditions and suitable initial conditions in a smooth bounded domain Ω in \mathbb{R}^n ($n \leq 2$) and on a time period $[0, T]$ with $\mathbf{u} = [u_1(x, t), u_2(x, t)]^T$, $x \in \Omega$, $t \in [0, T]$. Here τ_1 , τ_2 and d_1 , d_2 correspond to time scales and diffusion coefficients of u_1 and u_2 , respectively. $f_1(u_1, u_2)$ denotes a cubic nonlinear term. For instance, $f_1(u_1, u_2) = u_1(u_1 - \beta)(1 - u_1) - u_2 + \kappa$ with $0 < \beta < \frac{1}{2}$, and $f_2(u_1, u_2) = u_2 - \gamma u_1 + \epsilon$ for some suitable constants κ, γ, ϵ . The FHNE (2.1) is an activator-inhibitor system, where u_1 is the activator since it activates u_2 in the second component of (2.1), i.e. leads to an increase of u_2 . On the other hand, u_2 is the inhibitor since it leads to a decrease in u_1 and u_2 in both component of (2.1). In this thesis, we consider the diffusive FHNE (2.1) with a fast inhibitor diffusion, i.e. $d_2/d_1 > 1$, or it can be called as a slow activator diffusion.

A Reaction-diffusion system with pattern formation was first proposed by Turing in 1952 in his only paper on biological pattern formation [91]. In a single reaction-diffusion equation, the presence of the diffusion term does not change the stability of the steady state solution, in other words, the stable steady state solution can not be unstable when the diffusion term is added to the equation. On the contrary, the diffusion makes any stable steady state solution more stable. However, when two or more reaction-diffusion equations are considered, there can be a stable steady state solution turned into an unstable state under the presence of the diffusion term. Alan Turing first proposed this relationship between the parameters of reaction-diffusion equations and pattern formation. He explained that under certain conditions any stable steady state can be converted into an unstable state. This is known as “*diffusion driven instability*” or “*Turing instability*”. Some examples of pattern-generating reaction-diffusion equations are Gierer-Meinhardt [37], Gray-Scott [81], Lengyel-Epstein [58]

For “*Turing instability*”, the homogeneous steady state solution in the absence of diffusion terms must be linearly stable. Any homogeneous steady state solution $\mathbf{u} = (u_1, u_2)^T = (\phi, \psi)^T$ of equation (2.1) can be found as

$$\mathbf{f}(\mathbf{u}) = \mathbf{0}, \quad (2.3)$$

or equivalently,

$$\begin{aligned} f_1(\phi, \psi) &= 0, \\ f_2(\phi, \psi) &= 0. \end{aligned} \quad (2.4)$$

These homogeneous steady state solutions are linearly stable if and only if all eigenvalues of the stability matrix have negative real part. The stability matrix is

$$\mathbf{A} := \begin{bmatrix} (f_1)_1 & (f_1)_2 \\ (f_2)_1 & (f_2)_2 \end{bmatrix}_{(\phi, \psi)}, \quad (2.5)$$

where $(f_i)_j$ for $i, j \in \{1, 2\}$ are the partial derivatives of f_i , $i = 1, 2$ with respect to u_1 and u_2 . The eigenvalues of the stability matrix \mathbf{A} are found by

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} (f_1)_1 - \lambda & (f_1)_2 \\ (f_2)_1 & (f_2)_2 - \lambda \end{vmatrix}_{(\phi, \psi)} \\ \implies \lambda^2 - \lambda((f_1)_1 + (f_2)_2) + ((f_1)_1(f_2)_2 - (f_1)_2(f_2)_1) &|_{(\phi, \psi)} = 0 \\ \implies \lambda_{1,2} &= \frac{((f_1)_1 + (f_2)_2) \mp \sqrt{((f_1)_1 + (f_2)_2)^2 - 4((f_1)_1(f_2)_2 - (f_1)_2(f_2)_1)}}{2} \Big|_{(\phi, \psi)}. \end{aligned}$$

Hence, linear stability of homogeneous steady state solutions is guaranteed iff

$$\text{tr}(A) = (f_1)_1 + (f_2)_2 < 0, \quad (2.6)$$

$$|A| = (f_1)_1(f_2)_2 - (f_1)_2(f_2)_1 > 0. \quad (2.7)$$

In the presence of diffusion terms, stable steady states may be converted into unstable ones. We will derive other conditions, in addition to (2.6)-(2.7), for the “*Turing instability*”.

Linearizing equation (2.1) around the homogeneous steady state solution $\mathbf{u} = (\phi, \psi)^T$ leads to

$$\tau \mathbf{u}_t = \mathbf{D} \Delta \mathbf{u} + \mathbf{A} \mathbf{u}. \quad (2.8)$$

The general form of the solution for the diffusion equation $\mathbf{u}_t = \mathbf{D} \Delta \mathbf{u}$,

$$\mathbf{u}(x, t) = \sum_{j=0}^{\infty} a_j e^{\lambda_j t} e^{-ik_j x}$$

with corresponding wave modes k_j and eigenvalues $\lambda_j = \lambda(k_j)$, obtained by Fourier expansion. The general solution of (2.8) by substituting this formula into (2.8) gives

$$|\lambda_j \mathbf{I} - \mathbf{A} + k_j^2 \mathbf{D}| = 0,$$

for each k_j . Thus,

$$\begin{aligned} |\lambda \mathbf{I} - \mathbf{A} + \mathbf{D} k^2| &= \begin{vmatrix} \lambda - (f_1)_1 + k^2 d_1 & -(f_1)_2 \\ -(f_2)_1 & \lambda - (f_2)_2 + k^2 d_2 \end{vmatrix}_{(\phi, \psi)} \\ \implies \lambda^2 - \lambda b + c &|_{(\phi, \psi)} = 0, \\ \implies \lambda_{1,2} &= \frac{-b \mp \sqrt{b^2 - 4c}}{2} \Big|_{(\phi, \psi)}, \end{aligned}$$

where

$$\begin{aligned} b &= (f_1)_1 + (f_2)_2 - k^2(d_1 + d_2) \\ &= \text{tr}(A) - k^2(d_1 + d_2), \end{aligned}$$

$$\begin{aligned}
c &= (f_1)_1(f_2)_2 - (f_1)_2(f_2)_1 - k^2(d_2(f_1)_1 + d_1(f_2)_2) + k^4d_1d_2 \\
&= |A| - k^2(d_2(f_1)_1 + d_1(f_2)_2) + k^4d_1d_2.
\end{aligned}$$

For the “*Turing instability*” or “*diffusion driven instability*”, steady state solutions must be unstable. So some eigenvalues of the stability matrix must have positive real part. The condition (2.6), and $-k^2(d_1 + d_2) < 0$ imply $\lambda_1 + \lambda_2 = b < 0$, there is at least one eigenvalue with negative real part. However, in order to get an eigenvalue with positive real part, the other eigenvalues must have positive real part. Hence, either the inequality $\lambda_1\lambda_2 = c < 0$ must be satisfied or some value of c must be negative. So if $\lambda_1\lambda_2 = c < 0$ is satisfied, then a third condition for “*Turing instability*” becomes

$$d_2(f_1)_1 + d_1(f_2)_2 > 0 \quad (2.9)$$

with $d_1 \neq d_2$. Otherwise this contradicts with the condition (2.6). On the other hand, if c is negative for some values, then its minimum value must be negative. For this, one must check first and second derivatives of c

$$\begin{aligned}
c'(\tilde{k}) &= -(d_2(f_1)_1 + d_1(f_2)_2) + 2\tilde{k}d_1d_2, \\
c''(\tilde{k}) &= 2d_1d_2
\end{aligned}$$

with $\tilde{k} = k^2$.

Note that, c is concave up and its minimum occurs at the critical point

$$k^* = \sqrt{\tilde{k}} = \sqrt{\frac{d_2(f_1)_1 + d_1(f_2)_2}{2d_1d_2}}.$$

Hence, the minimum value is found as

$$c(k^*) = |A| - \frac{(d_2(f_1)_1 + d_1(f_2)_2)^2}{4d_1d_2}.$$

In order to make this value negative, we obtain the last “*Turing instability*” condition, which is given by

$$(f_1)_1(f_2)_2 - (f_1)_2(f_2)_1 < \frac{(d_2(f_1)_1 + d_1(f_2)_2)^2}{4d_1d_2}. \quad (2.10)$$

In summarize, “*Turing instability*” conditions for equation (2.1) can be written as

$$(f_1)_1 + (f_2)_2 < 0, \quad (2.11a)$$

$$(f_1)_1(f_2)_2 - (f_1)_2(f_2)_1 > 0, \quad (2.11b)$$

$$d_2(f_1)_1 + d_1(f_2)_2 > 0, \quad (2.11c)$$

$$(d_2(f_1)_1 + d_1(f_2)_2)^2 > 4d_1d_2((f_1)_1(f_2)_2 - (f_1)_2(f_2)_1), \quad (2.11d)$$

where all functions are evaluated at homogeneous stable steady state solution $\mathbf{u} = (\phi, \psi)^T$. Take in mind that either third or last condition is enough for “*Turing instability*”. For a detailed derivation of these conditions, see [68].

We firstly focus on one dimensional problems, although many real world patterns are two or three dimensional. We do not study three-dimensional problems.

2.1.1 Pattern Formation in 1D

Localized structures like fronts and pulses are the most well-known one dimensional patterns in reaction diffusion systems [30].

$$\begin{aligned}\tau_1 u_{1t} &= d_1 u_{1xx} + f(u_1) - u_2, \\ \tau_2 u_{2t} &= d_2 u_{2xx} + u_1 - \gamma u_2 + \epsilon,\end{aligned}\tag{2.12}$$

where $f(u_1) = u_1 - \beta u_1^3$ with $0 < \beta < \frac{1}{2}$, τ_1, τ_2 are time scales, and γ, ϵ are positive constants.

Existence of fronts and pulses for the diffusive FHNE are shown in [30]. In this section, we derive conditions for the parameters of the diffusive FHNE (2.12) with travelling fronts and pulses satisfying “*Turing instability*” conditions (2.11).

A travelling front is a travelling wave that connects two different stable states of a reaction-diffusion system with a bistable nonlinearity. Such trajectories are called heteroclinic orbits. Pulses exist far away from the homogeneous equilibrium and results from the balance between the dissipation and nonlinearity. Its trajectories are called homoclinic orbits since they start and end at the same steady state solutions.

Without diffusion terms, the homogeneous steady state solutions of (2.12) are determined by the intersection of the following equations

$$\begin{aligned}f(u_1) - u_2 &= 0, \\ u_1 - \gamma u_2 + \epsilon &= 0.\end{aligned}\tag{2.13}$$

Depending the number of solutions of the system (2.13), mono/bi-stability of the homogeneous steady state solutions of equation (2.12) are determined.

When all the parameters, other than γ , are fixed as [16]

$$(\tau_1, \tau_2, d_1, d_2, \beta, \epsilon) = \left(1, \frac{25}{2}, 1, \frac{5}{4}, \frac{1}{3}, \frac{7}{10}\right),$$

we obtain from (2.13) for u_2

$$u_2 = \frac{u_1 + \epsilon}{\gamma} = \frac{u_1 + \frac{7}{10}}{\gamma}.$$

Substituting this in the first equation of (2.13) gives

$$-\frac{1}{3}u_1^3 + \left(1 - \frac{1}{\gamma}\right)u_1 - \frac{7}{10\gamma} = 0.$$

For the bistable case, this equation must have three distinct real roots, such that the first derivative

$$-u_1^2 + \left(1 - \frac{1}{\gamma}\right) = 0$$

must have two distinct real roots. Therefore, for $\gamma > 1$, the equation (2.12) exhibits for bistable case travelling fronts. For the monostable case solution will be travelling

pulses. For example, we obtain for $\gamma = 8 > 1$ travelling front, and for $\gamma = 0.8 < 1$ travelling pulse solutions.

We will show that, equation (2.12) satisfies “*Turing instability*” conditions (2.11) for the following choice of the parameters:

$$\mathbf{u}_1 = \begin{pmatrix} -1.66804 \\ -0.12101 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0.10039 \\ 0.10005 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 1.56766 \\ 0.28346 \end{pmatrix}.$$

The equilibrium solutions $\mathbf{u}_1, \mathbf{u}_3$ are stable and \mathbf{u}_2 is unstable. So, for \mathbf{u}_1 and \mathbf{u}_3 the stability matrices are given as

$$\mathbf{A}_{\mathbf{u}_1} = \begin{bmatrix} -1.78236 & -1 \\ 1 & -8 \end{bmatrix}, \quad \mathbf{A}_{\mathbf{u}_3} = \begin{bmatrix} -1.45756 & -1 \\ 1 & -8 \end{bmatrix}.$$

Except third condition, all conditions of (2.11) are satisfied for both \mathbf{u}_1 and \mathbf{u}_3 . Hence, front solutions of (2.12) are indeed Turing patterns.

Similarly for travelling pulse solutions, the only homogeneous steady state solution of equation (2.12) is given as

$$\mathbf{u} = \begin{pmatrix} -1.19941 \\ -0.62426 \end{pmatrix},$$

which is stable with

$$\mathbf{A}_{\mathbf{u}} = \begin{bmatrix} -0.43858 & -1 \\ 1 & -0.8 \end{bmatrix}.$$

Again, except third condition, all conditions of (2.11) are satisfied for the equilibrium solution \mathbf{u} . Hence, pulse solutions are also Turing patterns.

2.1.2 Pattern Formation in 2D

Now, we will investigate a pattern formation for the two-dimensional diffusive FHNE

$$\begin{aligned} \tau_1 u_{1t} &= d_1 \Delta u_1 + f(u_1) - u_2 + \kappa, \\ \tau_2 u_{2t} &= d_2 \Delta u_2 - u_2 + u_1, \end{aligned} \tag{2.14}$$

where $f(u_1) = u_1 - \beta u_1^3$ with $0 < \beta < \frac{1}{2}$, and κ is a positive constant. Turing patterns for 2D diffusive FHNE (2.14) occur in form of spots or labyrinth-like patterns, depending on the number of global minima of the Lyapunov energy functional [61]. Lyapunov energy functional will be introduced in the next subsection. If the Lyapunov energy functional of the system has a unique global minimum, spot will be formed. Otherwise, labyrinth-like patterns emerge. An analytical criteria is determined in [61] in order to determine the number of global minima of Lyapunov energy functional. The number of global minima can be determined by that the nonlinear reaction function contain quadratic terms or not. If the reaction function does not include even polynomial terms, then its continuous energy function contains only even polynomial terms. Hence, this energy functional will have two different points having the same minimum value. So, its pattern structure will be labyrinth-like patterns. Otherwise,

under the presence of a quadratic term in reaction term in (2.14), the pattern formation will appear as spots.

When we apply this criteria to our equation (2.14), the reaction term includes the only even order term κ . So, for $\kappa = 0$ the pattern formation will be labyrinth-like patterns, and for non-zero values of κ the pattern formation will be spots.

When we fix the parameters, other than d_2 as

$$(\tau_1, \tau_2, d_1, \kappa) = (1, 1, 0.00028, 0 \text{ or } 0.1),$$

arbitrarily and determine the values of d_2 for the “*Turing instability*” conditions (2.11), the only equilibrium solution of (2.14) is

$$\mathbf{u} = \begin{pmatrix} 0.46416 \\ 0.46416 \end{pmatrix}$$

with

$$\mathbf{A}_{\mathbf{u}} = \begin{bmatrix} 0.35367 & -1 \\ 1 & -1 \end{bmatrix}.$$

The first two conditions of (2.11) are clearly satisfied. So, for the “*Turing instability*” third or last conditions of (2.11) are satisfied by choosing $d_2 > 0.00079$ or $d_2 > 0.07666$. Hence, for “*Turing instability*” it is enough to take $d_2 > 0.07666$.

2.1.3 Three Component FitzHugh-Nagumo System

An example of a three component reaction-diffusion equation is a model for gas discharge dynamics [79]

$$\begin{aligned} \tau_1 u_{1t} &= u_{1xx} + u_1 - u_1^3 - \epsilon(\alpha u_2 + \beta u_3 + \gamma), \\ \tau_2 u_{2t} &= \frac{1}{\epsilon^2} u_{2xx} + u_1 - u_2, \\ \tau_3 u_{3t} &= \frac{d^2}{\epsilon^2} u_{3xx} + u_1 - u_3, \end{aligned} \tag{2.15}$$

where $0 < \epsilon \ll 1$, $d > 1$, and α and β denote the reacting rates and the constant γ is the source term. The equation (2.15) consists of one activator u_1 and two inhibitors u_2, u_3 components. It can be considered also as a generalized FHNE. First inhibitor u_2 diffuses more rapidly than the second inhibitor u_3 . The equation (2.15) became a standard model for studying the dynamics and interactions of spatially localized structures like multi-pulses and multi-fronts in one dimensional reaction-diffusion equations [93] through mathematical analysis and numerical simulations.

Different patterns are determined by looking at the homogeneous steady state solutions of (2.15) by the intersection of the following equations

$$\begin{aligned} u_1 - u_1^3 - \epsilon(\alpha u_2 + \beta u_3 + \gamma) &= 0, \\ u_1 - u_2 &= 0, \\ u_1 - u_3 &= 0. \end{aligned} \tag{2.16}$$

The last two equations of (2.16) imply that stationary solution components are equal. Hence, for $u_1 = u_2 = u_3 = u$ the solution of equation (2.16) reduces to

$$u^3 - u + \epsilon(\alpha u + \beta u + \gamma) = 0. \quad (2.17)$$

For the bistable case, the first order derivative of the equation (2.17) has two distinct roots

$$3u^2 - 1 + \epsilon(\alpha + \beta) = 0. \quad (2.18)$$

if its discriminant satisfies

$$\Delta = -12 \cdot (\epsilon(\alpha + \beta) - 1) > 0.$$

Hence for $\epsilon(\alpha + \beta) < 1$, (2.15) has bi-stable nonlinear term, and the three component FHNE (2.15) has front solutions. Depending on the initial conditions, (2.15) may have one-front, two-front, or multi-front solutions [93].

For the one and two front solutions, we fix the parameter values as

$$(\alpha, \beta, \epsilon, \gamma, d, \tau_1, \tau_2, \tau_3) = (3, 1, \frac{1}{100}, -\frac{1}{4}, 5, 1, \frac{100}{3}, 100),$$

in which some of the parameters are chosen from [93] and the rest we choose according to the skew-gradient structure, introduced in the following subsection 2.1.4, and “*Turing instability*” conditions.

We show that the equation (2.15) satisfies “*Turing instability*” conditions (2.11) for these parameters. The equilibrium solution is

$$\mathbf{u} = \begin{pmatrix} -0.0026 \\ -0.0026 \\ -0.0026 \end{pmatrix}$$

with

$$\mathbf{A}_{\mathbf{u}} = \begin{bmatrix} 0.99998 & -0.03 & -0.01 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}.$$

Since $\text{tr}(\mathbf{A}_{\mathbf{u}}) = -1.00002 < 0$ and $|\mathbf{A}_{\mathbf{u}}| = 0.95997972 > 0$, the first two “*Turing instability*” conditions (2.11) are satisfied. After a some calculation we obtain

$$c = k^6 d_1 d_2 d_3 + k^4 (d_1 d_3 + d_1 d_2) + k^2 (d_1 - 2 \cdot 0.99998 d_3 - 0.01 d_2) - 2 \cdot 0.99998 - 0.01.$$

The minimum value of c is -2705.11889 , which is negative and the last condition of (2.11) is also satisfied. Therefore, one-front and two-front solutions of (2.15) are Turing patterns. In a similar way, one can show that “*Turing instability*” conditions are clearly satisfied for multi-front solutions of (2.15) with chosen parameters

$$(\alpha, \beta, \epsilon, \gamma, d, \tau_1, \tau_2, \tau_3) = (100, 100, \frac{1}{100}, -\frac{1}{4}, 5, 1, 1, 1)$$

in the same way as the previous one and two front solutions.

2.1.4 Skew-Gradient Structure of the Diffusive FHNE

The two component diffusive FHNE

$$\begin{aligned}\tau_1 u_{1t} &= d_1 \Delta u_1 + f_1(u_1, u_2), \\ \tau_2 u_{2t} &= d_2 \Delta u_2 + f_2(u_1, u_2),\end{aligned}\tag{2.19}$$

has a skew-gradient structure if there exists a function F satisfying

$$\nabla F = (f_1(u_1, u_2), -f_2(u_1, u_2)),\tag{2.20}$$

where ∇ denotes the gradient operator. The alternative form of this definition is given as in [101]:

$$\frac{\partial}{\partial u_2} \left(\frac{f_1(u_1, u_2)}{\tau_1} \right) = - \frac{\partial}{\partial u_1} \left(\frac{f_2(u_1, u_2)}{\tau_2} \right).\tag{2.21}$$

This skew-gradient structure was first introduced by Yanagida in [101] for a generalized activator-inhibitor system in order to determine whether the steady states of (2.19) are stable or not. The diffusive FHNE satisfies the skew-gradient condition (2.20) or (2.21).

The steady state of (2.19) is a critical point of the energy functional

$$E(u_1, u_2) = \int_{\Omega} \left(\frac{d_1}{2} |\nabla u_1|^2 - \frac{d_2}{2} |\nabla u_2|^2 - F(u_1, u_2) \right) dx\tag{2.22}$$

with the corresponding potential function $F(u_1, u_2)$ satisfying

$$\frac{\partial F}{\partial u_1} = f_1(u_1, u_2) \quad \text{and} \quad - \frac{\partial F}{\partial u_2} = f_2(u_1, u_2).$$

For instance, the potential function for the diffusive FHNE, defined in 2.1.1, is

$$F(u_1, u_2) = -\frac{u_1^4}{4} + \frac{(1 + \beta)u_1^3}{3} - \frac{\beta u_1^2}{2} - u_1 u_2 + \frac{\gamma u_2^2}{2} - \epsilon u_2.$$

The first equation (activator) of (2.19) is a gradient flow with the potential $F(u_1, u_2)$, and the second (inhibitor) with the potential $-F(u_1, u_2)$ [100]. The energy functional $E(u_1, u_2)$ does not correspond to the Lyapunov functional, because it is not necessarily non-increasing or non-decreasing in time. However, it was proved in [100] that a steady state of a skew-gradient system is stable if and only if it is a mini-maximizer of the energy functional $E(u_1, u_2)$:

- If $u_1 = \xi$ is a local minimizer of $E(u_1, \eta)$ and $u_2 = \eta$ is a local maximizer of $E(\xi, u_2)$, then $(u_1, u_2) = (\xi, \eta)$ is a mini-maximizer of $E(u_1, u_2)$.
- If $E(\xi, \hat{u}_2) \leq E(\xi, \eta) \leq E(\hat{u}_1, \eta)$ for any neighborhoods \hat{u}_1, \hat{u}_2 of ξ and η , respectively, then we say that $(u_1, u_2) = (\xi, \eta)$ is a mini-maximizer of $E(u_1, u_2)$.

This result generalizes the stability criteria for gradient systems.

The n -component system with the skew-gradient structure has the form

$$f_i(u_1, \dots, u_n) = \begin{cases} \frac{\partial}{\partial u_i} F(u_1, \dots, u_n) & i = 1, \dots, \ell, \\ -\frac{\partial}{\partial u_i} F(u_1, \dots, u_n) & i = \ell + 1, \dots, n, \end{cases}$$

where F denotes the corresponding potential function. The first ℓ components of system $\{u_1, \dots, u_\ell\}$ are activators, whereas the remaining $\{u_{\ell+1}, \dots, u_n\}$ are inhibitors.

2.2 Pattern Formation in Convective FitzHugh-Nagumo Equation

In this section, we study a convective FHNE [35]

$$\begin{aligned} \tau_1 u_{1t} &= d_1 \Delta u_1 - \mathbf{V} \cdot \nabla u_1 + f_1(u_1, u_2), \\ \tau_2 u_{2t} &= d_2 \Delta u_2 - \mathbf{V} \cdot \nabla u_2 + f_2(u_1, u_2) \end{aligned} \quad (2.23)$$

on a space time cylinder $Q_T := \Omega \times (0, T)$, where $\Omega = (0, L) \times (0, H)$ with zero flux boundary condition except the left boundary, which is Dirichlet type as $u_1 = u_2 = 0.1$. As initial conditions we set $u_1 = 0.1$ inside a narrow triangle $0 \leq x \leq 0.5$, $0 \leq y \leq H$, and $u_1 = 0$ outside this rectangle; $u_2 = 0$ everywhere. Here, d_1 and d_2 correspond to diffusion coefficients of u_1 and u_2 , respectively. The first component of a reaction term $f_1(u_1, u_2)$ involves a cubic nonlinear term of u_1 .

In the absence of flow, i.e., $\mathbf{V} = 0$, equation (2.23) is the diffusive FHNE (2.1). There is also a three component convective FHNE like diffusive FHNE [34]. However, in this work we focus on only this two dimensional, coupled version of convective FHNE.

The convective FHNE (2.23) was developed as a model for wave propagation in blood coagulation and bioreactor systems [34, 35, 59]. Although the blood is a non-Newtonian fluid, this FHNE is described with the assumption that the blood is an incompressible fluid with divergence free velocity field, i.e. $\text{div} \mathbf{V} = 0$ [35, 59]. Among the state variables, u_1 represents the activator of coagulation (thrombin), and u_2 represents the inhibitor (usually protein C) [59], which are described more detail in [10].

The waves of the convective FHNE are autonomous, that is, changing the initial condition does not affect the form and velocity of occurring waves [35]. These waves can either spread over the whole spatial domain or localize in some part of the domain, called restrictions. In this thesis, we study formation of patterns in an excitable media, driven by a convective field in both activator and inhibitor components of a diffusive FHNE for two-dimensional space.

2.2.1 Parameter Analysis for Wave Type Solutions of the Convective FHNE

For the convective FHNE (2.23), the reaction terms are chosen as

$$f_1(u_1, u_2) = c_1 u_1 (u - c_2) (1 - u_1) - u_2, \quad f_2(u_1, u_2) = \epsilon (c_3 u_1 - u_2)$$

for some positive real constants c_1, c_2, c_3 [35]. The homogeneous steady state solution $\mathbf{u} = (u_1, u_2)^T = (\phi, \psi)^T$ of equation (2.23) becomes

$$\begin{aligned} c_1\phi(\phi - c_2)(1 - \phi) - \psi &= 0, \\ \epsilon(c_3\phi - \psi) &= 0. \end{aligned} \quad (2.24)$$

The equation (2.24) is reduced to

$$\phi(c_1\phi^2 - c_1(1 + c_2)\phi - (c_1c_2 + c_3)) = 0. \quad (2.25)$$

In an excitable medium, (2.25) is monostable, i.e., it has only one equilibrium solution, which is $(\phi, \psi) = (0, 0)$. Therefore, the discriminant of $c_1\phi^2 - c_1(1 + c_2)\phi + (c_1c_2 + c_3)$ is less than zero:

$$\Delta = c_1^2(1 + c_2)^2 - 4c_1(c_1c_2 + c_3) < 0. \quad (2.26)$$

If we choose c_2 and c_3 as 0.02 and 5, respectively, then (2.26) is simplified as

$$0.9614c_1^2 - 20c_1 < 0. \quad (2.27)$$

Hence, for $0 < c_1 < 20.803$ the excitable medium is monostable.

2.3 Space-Time Discretization

In this section, we give a space-time discretization of 2D diffusive and convective FHNE with homogeneous Neumann and non-homogeneous Dirichlet boundary conditions

$$\begin{aligned} \tau_1 u_{1t} - d_1 \Delta u_1 + \mathbf{V} \cdot \nabla u_1 - f_1(u_1, u_2) &= 0, & \text{in } Q_T, \\ \tau_2 u_{2t} - d_2 \Delta u_2 + \mathbf{V} \cdot \nabla u_2 - f_2(u_1, u_2) &= 0, & \text{in } Q_T, \\ \frac{\partial u_1}{\partial n} = \frac{\partial u_2}{\partial n} &= 0, & \text{on } \Sigma_T^N, \\ u_1 = u_{1D}, \quad u_2 = u_{2D}, & & \text{on } \Sigma_T^D, \\ u_1(x, 0) = u_{10}, \quad u_2(x, 0) = u_{20}, & & \text{in } \Omega \end{aligned} \quad (2.28)$$

on a space-time domain $Q_T := \Omega \times (0, T)$ with $\Omega = [0, L] \times [0, H]$. The lateral surface is denoted by $\Sigma = \Gamma \times (0, T)$. We use the notation $\Sigma_T^D := \Gamma_D \times (0, T)$ for Dirichlet and $\Sigma_T^N := \Gamma_N \times (0, T)$ for Neumann boundaries. Dirichlet boundary conditions are given by $u_{1D}, u_{2D} \in H^{3/2}(\Gamma_D)$. Moreover, the initial functions are given as $u_{10}, u_{20} \in L^\infty(\Omega)$. We denote the outward unit normal vector and the associated outward normal derivative on $\partial\Omega$ by \mathbf{n} and ∂_n , respectively. Further, the function $f_1(u_1, u_2)$ denotes a cubic polynomial nonlinearity and $f_2(u_1, u_2)$ is a linear polynomial. For both diffusive and convective FHNEs (2.1)-(2.23), reaction terms $f_i(u_1, u_2)$ for $i = 1, 2$ can be written as a summation of separable functions of u_1 and u_2 :

$$\begin{aligned} f_1(u_1, u_2) &= f(u_1) + \ell_1(u_2), \\ f_2(u_1, u_2) &= \ell_2(u_1) + \ell_3(u_2), \end{aligned} \quad (2.29)$$

where f denotes cubic polynomial nonlinearity, ℓ_1, ℓ_2, ℓ_3 denote linear polynomials. The divergence free velocity field $\mathbf{V} = (V_{x_1}, V_{x_2})$ is given along the x_1 -direction with a parabolic profile

$$V_{x_1}(x_2) = ax_2(H - x_2), \quad V_{max} = \frac{1}{4}aH^2, \quad a > 0, \quad V_{x_2} = 0, \quad (2.30)$$

where V_{max} denotes the maximum wave speed of the velocity field.

For space discretization of equation (2.28), we use the symmetric interior penalty Galerkin (SIPG) method and as a time discretization we use the average vector field (AVF) method for the diffusive FHNE, the semi-implicit Euler method for the convective FHNE.

2.3.1 Space Discretization

Let $\{\mathcal{T}_h\}_h$ be a family of shape-regular meshes of a polygonal domain Ω such that each mesh \mathcal{T}_h consists of closed triangular elements, i.e., $\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} \overline{K}$. We assume that the mesh is a conforming mesh, i.e., the intersection of two elements $K_i \cap K_j$ is either empty, a vertex, an edge, or a face. A face occurs only if $i = j$. The diameter of an element K is denoted by h_K , and the length of an edge E is denoted by h_E . In addition, the maximum diameter is $h = \max h_K$. Also we assume that the mesh is shape-regular, i.e., there is a constant $C > 0$ such that

$$\frac{h_K}{\rho_K} \leq C, \quad \forall K \in \mathcal{T}_h$$

with the maximum diameter of ball ρ_K , enclosed by the element K .

We split the set of all edges \mathcal{E}_h into the set of interior edges \mathcal{E}_h^0 , Dirichlet boundary edges \mathcal{E}_h^D , and Neumann boundary edges \mathcal{E}_h^N so that $\mathcal{E}_h = \mathcal{E}_h^B \cup \mathcal{E}_h^0$ with $\mathcal{E}_h^B = \mathcal{E}_h^D \cup \mathcal{E}_h^N$. Let \mathbf{n} denote the unit outward normal to $\partial\Omega$. For the activator u_1 and the inhibitor u_2 , we define the inflow and outflow boundaries of $\partial\Omega$ by

$$\Gamma^- = \{x \in \partial\Omega : \mathbf{V}(x) \cdot \mathbf{n}(x) < 0\}, \quad \Gamma^+ = \partial\Omega \setminus \partial\Gamma^-.$$

In a similar way, we can define the inflow and outflow boundaries of an element K as

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

where \mathbf{n}_K denotes 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 u , there are two traces of u along E , denoted by $u|_E$ from inside K and $u^e|_E$ from inside K^e . Then, the jump and average of u across the edge E are defined by

$$[u] = u|_E \mathbf{n}_K + u^e|_E \mathbf{n}_{K^e}, \quad \{u\} = \frac{1}{2}(u|_E + u^e|_E). \quad (2.31)$$

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

$$[\nabla u] = \nabla u|_E \cdot \mathbf{n}_K + \nabla u^e|_E \cdot \mathbf{n}_{K^e}, \quad \{\nabla u\} = \frac{1}{2}(\nabla u|_E + \nabla u^e|_E). \quad (2.32)$$

For a boundary edge $E \in K \cap \partial\Omega$, we set $\{u\} = u$ and $[u] = u\mathbf{n}$.

For DG methods, we do not impose continuity constraints on the trial and test functions across the element interfaces. Therefore, the weak formulation of (2.28) must include jump terms across interfaces. In order to control these jump terms, penalty terms are added. The discontinuous discrete space is defined as

$$W_h = \{w \in L^2(\Omega) : w|_K \in \mathbb{P}^k(K), \quad \forall K \in \mathcal{T}_h\}, \quad (2.33)$$

where $\mathbb{P}^k(K)$ is the set of all polynomials defined on $K \in \mathcal{T}_h$ of degree at most k . We note that the space of discrete states and the space of test functions are identical due to the weak treatment of boundary conditions for DG methods.

Then by using the identities (2.29), a semi-discrete formulation of (2.28) for $\forall w \in W_h$ and $t \in (0, T]$ becomes

$$\begin{aligned} \left(\tau_1 \frac{du_{1h}}{dt}, w \right) + a_h(d_1; u_{1h}, w) - (f(u_{1h}), w) - (\ell_1(u_{2h}), w) &= r_{h,u_1}(w), \\ (u_{1h}(\cdot, 0), w) &= (u_{10}, w), \\ \left(\tau_2 \frac{du_{2h}}{dt}, w \right) + a_h(d_2; u_{2h}, w) - (\ell_2(u_{1h}), w) - (\ell_3(u_{2h}), w) &= r_{h,u_2}(w), \\ (u_{2h}(\cdot, 0), w) &= (u_{20}, w), \end{aligned} \quad (2.34)$$

where the (bi)-linear terms are defined for $i = 1, 2, \forall w \in W_h$,

$$\begin{aligned} a_h(d_i; u_{ih}, w) &= \sum_{K \in \mathcal{T}_h} \int_K d_i \nabla u_{ih} \cdot \nabla w \, dx \\ &\quad - \sum_{E \in \mathcal{E}_h^0 \cup \mathcal{E}_h^D} \int_E (\{d_i \nabla u_{ih}\} \cdot [w]) \, ds + \kappa \sum_{E \in \mathcal{E}_h^0 \cup \mathcal{E}_h^D} \int_E (\{d_i \nabla w\} \cdot [u_{ih}]) \, ds \\ &\quad + \sum_{E \in \mathcal{E}_h^0 \cup \mathcal{E}_h^D} \frac{\sigma d_i}{h_E} \int_E [u_{ih}] \cdot [w] \, ds + \sum_{K \in \mathcal{T}_h} \int_K \mathbf{v} \cdot \nabla u_{ih} w \, dx \\ &\quad + \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \setminus \partial\Omega} \mathbf{v} \cdot \mathbf{n} (u_{ih}^e - u_{ih}) w \, ds - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \mathbf{v} \cdot \mathbf{n} u_{ih} w \, ds, \\ r_{h,u_i}(w) &= \sum_{E \in \mathcal{E}_h^D} \int_E u_{iD} \left(\frac{\sigma d_i}{h_E} \mathbf{n} \cdot [w] - \{d_i \nabla w\} \right) \, ds \\ &\quad - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \mathbf{v} \cdot \mathbf{n} u_{iD} w \, ds, \end{aligned} \quad (2.35)$$

where the parameter $\sigma \in \mathbb{R}_0^+$ is called a penalty parameter which should be sufficiently large to ensure the stability of the DG discretization; independent of the mesh size h and of the diffusion coefficients d_i as described in [74] with a lower bound depending only on the polynomial degree.

Depending on the choice of κ , there exists three different types of DG methods

- For $\kappa = 1$, the resulting method is called the symmetric interior penalty Galerkin (SIPG) method [38].
- For $\kappa = -1$, the resulting method is called non-symmetric interior penalty Galerkin (NIPG) method [98, 8].
- For $\kappa = 0$, the resulting method is called incomplete interior penalty Galerkin method.

In this thesis, we use the SIPG method. For each time step, we can expand the discrete solutions of the activator and inhibitor variables u_{1h}, u_{2h} as

$$u_{1h}(t) = \sum_{i=1}^{n_e} \sum_{j=1}^{n_k} u_{1j}^i(t) \phi_j^i, \quad u_{2h}(t) = \sum_{i=1}^{n_e} \sum_{j=1}^{n_k} u_{2j}^i(t) \phi_j^i, \quad (2.36)$$

where u_{1j}^i and u_{2j}^i are the unknown coefficients and ϕ_j^i are the basis functions for $j = 1, 2, \dots, n_k$ and $i = 1, 2, \dots, n_e$. The number n_e denotes the number of DG elements and $n_k = (k+1)(k+2)/2$ is the local dimension of each DG element depending on the order k of the polynomial basis. The degree of freedoms (DoFs) for DG methods is given by $N = n_e \times n_k$.

Plugging (2.36) into the scheme (2.34), we get $2 \times N$ dimensional ODEs for the unknown vectors \vec{u}_1 and \vec{u}_2

$$\begin{aligned} \vec{u}_1 &= (u_{11}^1, \dots, u_{1n_k}^1, \dots, u_{11}^{n_e}, \dots, u_{1n_k}^{n_e}), \\ \vec{u}_2 &= (u_{21}^1, \dots, u_{2n_k}^1, \dots, u_{21}^{n_e}, \dots, u_{2n_k}^{n_e}), \end{aligned}$$

as following

$$\begin{aligned} \tau_1 \mathbf{M} \frac{d\vec{u}_1}{dt} + \mathbf{A}_{u_1} \vec{u}_1 - \mathbf{f}(\vec{u}_1) - \mathbf{M} \ell_1(\vec{u}_2) &= \mathbf{r}_{u_1}, \\ \tau_2 \mathbf{M} \frac{d\vec{u}_2}{dt} + \mathbf{A}_{u_2} \vec{u}_2 - \mathbf{M} \ell_2(\vec{u}_1) - \mathbf{M} \ell_3(\vec{u}_2) &= \mathbf{r}_{u_2}. \end{aligned} \quad (2.37)$$

Here, \mathbf{M} is the mass matrix, \mathbf{A}_{u_i} is the stiffness matrices corresponding to $a_h(d_i; u_{ih}, w)$; $\mathbf{f}(\vec{u}_1)$ is the vector corresponding to nonlinear term $(f(u_{1h}), w)$; \mathbf{r}_{u_i} are vectors corresponding to linear terms $r_{h,u_i}(w)$ for $i \in \{1, 2\}$.

The semi-discrete form of the diffusive FHNE is almost the same except the convection term in the bilinear form (2.35).

2.3.2 Time Discretization: Average Vector Field Method

Energy stable time discretization methods preserve the dissipative structure of the numerical solution of gradient flow equations and skew-gradient systems like the diffusive FHNE (2.1). After the spatial discretization, the small values of the diffusion parameters d_1 and d_2 lead to stiff systems. Therefore, we need to choose an appropriate time integrator technique, especially for diffusive FHNE (2.1).

Explicit methods, like forward Euler's method, were developed to calculate the state of the system at a future time without solving a nonlinear equation. Then implicit-explicit methods were developed, like semi-implicit Euler method, where the linear stiff part is treated implicitly, and the nonlinear part is explicitly. So to calculate the state of the system, again it is not needed to solve nonlinear equation. These methods do not preserve the energy dissipation (2.1). Implicit methods are more stable for solving stiff systems with a larger time step size. So we prefer to use energy stable implicit time discretization methods, which are implicit (backward) Euler's and average vector field (AVF) methods. These methods are robust with small diffusion parameters. The AVF method is the only second order implicit energy stable method and it preserves energy decreasing property for the gradient systems, like the Allen-Cahn and the Schlögl equations. In this part, we apply the AVF method [27, 42] to solve the system of ODEs arising from the semi-discretization of the diffusive FHNE (2.37).

We split the time interval $[0, T]$ into J equally-length subintervals $(t_{k-1}, t_k]$ with $0 = t_0 < t_1 < \dots < t_J = T$ with a uniform step-size $\Delta t = t_n - t_{n-1}$, $n = 1, 2, \dots, J$. The AVF method for an arbitrary ODE $\dot{u} = f(u)$ is given by

$$\frac{u^{n+1} - u^n}{\Delta t} = \int_0^1 f(\xi u^{n+1} + (1 - \xi)u^n) d\xi, \quad (2.38)$$

where $u^n \approx u(t_n)$ for $n = 1, \dots, J - 1$ and $u^0 = u(t_0)$.

For $n = 0, 1, \dots, J - 1$ the AVF method applied to the semi-discretized form (2.37) reads as

$$\begin{aligned} \tau_1 \mathbf{M} \frac{\vec{u}_1^{n+1} - \vec{u}_1^n}{\Delta t} &= \underbrace{\int_0^1 [\mathbf{r}_{u_1}^{n+1} - \mathbf{A}_{u_1}(\xi \vec{u}_1^{n+1} + (1 - \xi)\vec{u}_1^n)] d\xi}_{\text{linear}} \\ &+ \underbrace{\int_0^1 \mathbf{f}(\xi \vec{u}_1^{n+1} + (1 - \xi)\vec{u}_1^n) d\xi}_{\text{nonlinear}} + \underbrace{\int_0^1 \mathbf{M} \ell_1(\xi \vec{u}_2^{n+1} + (1 - \xi)\vec{u}_2^n) d\xi}_{\text{linear}}, \end{aligned}$$

$$\begin{aligned}
\tau_2 \mathbf{M} \frac{\vec{u}_2^{n+1} - \vec{u}_2^n}{\Delta t} &= \underbrace{\int_0^1 [\mathbf{r}_{u_2}^{n+1} - \mathbf{A}_{u_2}(\xi \vec{u}_2^{n+1} + (1 - \xi) \vec{u}_2^n)] d\xi}_{\text{linear}} \\
&\quad + \underbrace{\int_0^1 [\mathbf{M}\ell_2(\xi \vec{u}_1^{n+1} + (1 - \xi) \vec{u}_1^n) + \mathbf{M}\ell_3(\xi \vec{u}_2^{n+1} + (1 - \xi) \vec{u}_2^n)] d\xi}_{\text{linear}},
\end{aligned}$$

where we use the separable property of reaction terms as given in equation (2.29).

Since the AVF method coincides with the midpoint rule for the linear terms, this discrete formulation can be simplified as

$$\begin{aligned}
\left(\tau_1 \mathbf{M} + \frac{\Delta t}{2} \mathbf{A}_{u_1} \right) \vec{u}_1^{n+1} &= \left(\tau_1 \mathbf{M} - \frac{\Delta t}{2} \mathbf{A}_{u_1} \right) \vec{u}_1^n \\
&\quad + \Delta t \int_0^1 \mathbf{f}(\xi \vec{u}_1^{n+1} + (1 - \xi) \vec{u}_1^n) d\xi \\
&\quad + \frac{\Delta t}{2} \mathbf{M}\ell_1(\vec{u}_2^{n+1} + \vec{u}_2^n), \\
\left(\tau_2 \mathbf{M} + \frac{\Delta t}{2} \mathbf{A}_{u_2} \right) \vec{u}_2^{n+1} &= \left(\tau_2 \mathbf{M} - \frac{\Delta t}{2} \mathbf{A}_{u_2} \right) \vec{u}_2^n + \frac{\Delta t}{2} \mathbf{M}\ell_2(\vec{u}_1^{n+1} + \vec{u}_1^n) \\
&\quad + \frac{\Delta t}{2} \mathbf{M}\ell_3(\vec{u}_2^{n+1} + \vec{u}_2^n).
\end{aligned} \tag{2.39}$$

The fully discrete system of nonlinear equations (2.39) are solved by Newton's method on each time-interval $(t_{n-1}, t_n]$, $n = 0, 1, \dots, J - 1$.

We set the functions of system residuals for (2.39)

$$\begin{aligned}
R_1(\vec{u}_1^{n+1}, \vec{u}_2^{n+1}) &= \left(\tau_1 \mathbf{M} + \frac{\Delta t}{2} \mathbf{A}_{u_1} \right) \vec{u}_1^{n+1} - \left(\tau_1 \mathbf{M} - \frac{\Delta t}{2} \mathbf{A}_{u_1} \right) \vec{u}_1^n \\
&\quad + \Delta t \int_0^1 \mathbf{f}_1(\xi \vec{u}_1^{n+1} + (1 - \xi) \vec{u}_1^n) d\xi + \frac{\Delta t}{2} \mathbf{M}\ell_1(\vec{u}_2^{n+1} + \vec{u}_2^n), \\
R_2(\vec{u}_1^{n+1}, \vec{u}_2^{n+1}) &= \left(\tau_2 \mathbf{M} + \frac{\Delta t}{2} \mathbf{A}_{u_2} \right) \vec{u}_2^{n+1} - \left(\tau_2 \mathbf{M} - \frac{\Delta t}{2} \mathbf{A}_{u_2} \right) \vec{u}_2^n \\
&\quad + \frac{\Delta t}{2} \mathbf{M}\ell_2(\vec{u}_1^{n+1} + \vec{u}_1^n) + \frac{\Delta t}{2} \mathbf{M}\ell_3(\vec{u}_2^{n+1} + \vec{u}_2^n).
\end{aligned} \tag{2.40}$$

To solve the coupled nonlinear equations (2.40), we start with initial guesses $\vec{u}_{1(0)}^{n+1}$,

$\vec{u}_{2(0)}^{n+1}$. Then, the k^{th} Newton iteration step reads as

$$\begin{aligned} J^{(k)}[w_1^{(k)}; w_2^{(k)}] &= -[R_1(\vec{u}_{1(k)}^{n+1}, \vec{u}_{2(k)}^{n+1}); R_2(\vec{u}_{1(k)}^{n+1}, \vec{u}_{2(k)}^{n+1})], \\ \vec{u}_{1(k+1)}^{n+1} &= \vec{u}_{1(k)}^{n+1} + w_1^{(k)}, \\ \vec{u}_{2(k+1)}^{n+1} &= \vec{u}_{2(k)}^{n+1} + w_2^{(k)}, \quad k = 0, 1, 2, \dots, \end{aligned}$$

where the matrix J stands for the Jacobian matrix of the coupled system (2.40) evaluated at $[\vec{u}_{1(k)}^{n+1}; \vec{u}_{2(k)}^{n+1}]$. We further notice that the integral terms in (2.40) are approximated using the fourth order Gaussian quadrature rule so that the cubic nonlinear terms are evaluated exactly.

2.4 Preservation of the Discrete Energy for the Diffusive FHNE

In this section, we will show that the AVF time integrator combined with the SIPG space discretization method preserves the mini-maximizing property of the skew-gradient system [100] in the discrete form for diffusive FHNE (2.1). We will examine the relation between the stability of a steady state solution (ξ, η) of (2.1) and the mini-maximizing property of the critical point of the energy functional (2.22) for the SIPG-AVF fully discrete system for diffusive FHNE (2.1).

The SIPG discretized energy functional at a time $t_n = n\Delta t$ is given as

$$\begin{aligned} E(u_{1h}^n, u_{2h}^n) &= \frac{d_1}{2} \|\nabla u_{1h}^n\|_{L^2(\Omega)}^2 - \frac{d_2}{2} \|\nabla u_{2h}^n\|_{L^2(\Omega)}^2 + (F(u_{1h}^n, u_{2h}^n), 1)_\Omega \\ &\quad + \sum_{E \in \mathcal{E}_h^0} \left(-(\{d_1 \partial_n u_{1h}^n\}, [u_{1h}^n])_E + \frac{\sigma d_1}{2h_e} ([u_{1h}^n], [u_{1h}^n])_E \right) \\ &\quad - \sum_{E \in \mathcal{E}_h^0} \left(-(\{d_2 \partial_n u_{2h}^n\}, [u_{2h}^n])_E + \frac{\sigma d_2}{2h_e} ([u_{2h}^n], [u_{2h}^n])_E \right), \end{aligned} \quad (2.41)$$

where $(\cdot, \cdot)_\Omega$ is the L^2 inner product over the domain Ω and $F(\cdot, \cdot)$ is the potential function, see subsection 2.1.4.

When we integrate the semi-discrete system (2.34) in time by the AVF method, $\forall w_h \in W_h$

$$\begin{aligned} &\frac{1}{\Delta t} (u_{1h}^{n+1} - u_{1h}^n, w_h) + \frac{1}{2} a_h(d_1; u_{1h}^{n+1} + u_{1h}^n, w_h) \\ &\quad - \int_0^1 (f(u_{1h}), w_h) d\xi - \int_0^1 \ell_1(u_{2h}, w_h) d\xi = 0, \\ &\frac{1}{\Delta t} (u_{2h}^{n+1} - u_{2h}^n, w_h) + \frac{1}{2} a_h(d_2; u_{2h}^{n+1} + u_{2h}^n, w_h) \\ &\quad - \int_0^1 (\ell_2(u_{1h}), w_h) d\xi - \int_0^1 (\ell_3(u_{2h}), w_h) d\xi = 0, \end{aligned} \quad (2.42)$$

where the bilinear form is given by for $i = 1, 2$

$$\begin{aligned}
a_h(d_i; u_{ih}, w_h) &= \sum_{K \in \mathcal{T}_h} \int_K d_i \nabla u_{ih} \cdot \nabla w_h \, dx \\
&\quad - \sum_{E \in \mathcal{E}_h^0} \int_E (\{d_i \nabla u_{ih}\} \cdot [w_h]) \, ds - \sum_{E \in \mathcal{E}_h^0} \int_E (\{d_i \nabla w_h\} \cdot [u_{ih}]) \, ds \\
&\quad + \sum_{E \in \mathcal{E}_h^0} \frac{\sigma d_i}{h_E} \int_E [u_{ih}] \cdot [w_h] \, ds.
\end{aligned}$$

Choosing $w_h \in \{u_{1h}^{n+1} - u_{1h}^n, u_{2h}^{n+1} - u_{2h}^n\}$, respectively in (2.42), and using the identity $(a + b, a - b) = (a^2 - b^2, 1)$ with the bilinear form of a_h lead:

$$\begin{aligned}
&\frac{1}{\Delta t} (u_{1h}^{n+1} - u_{1h}^n, u_{1h}^{n+1} - u_{1h}^n) + \frac{1}{2} a_h(d_1; u_{1h}^{n+1}, w_h) + \frac{1}{2} a_h(d_1; u_{1h}^n, w_h) \\
&\quad - \int_{\Omega} \left[\int_0^1 (f(\xi u_{1h}^{n+1} + (1 - \xi) u_{1h}^n) (u_{1h}^{n+1} - u_{1h}^n) d\xi \right] dx \\
&\quad - \int_{\Omega} \left[\int_0^1 (\ell_1(\xi u_{2h}^{n+1} + (1 - \xi) u_{2h}^n) (u_{1h}^{n+1} - u_{1h}^n) d\xi \right] dx = 0, \\
&\frac{1}{\Delta t} (u_{2h}^{n+1} - u_{2h}^n, u_{2h}^{n+1} - u_{2h}^n) + \frac{1}{2} a_h(d_2; u_{2h}^{n+1}, w_h) + \frac{1}{2} a_h(d_2; u_{2h}^n, w_h) \\
&\quad - \int_{\Omega} \left[\int_0^1 (\ell_2(\xi u_{1h}^{n+1} + (1 - \xi) u_{1h}^n) (u_{2h}^{n+1} - u_{2h}^n) d\xi \right] dx \\
&\quad - \int_{\Omega} \left[\int_0^1 (\ell_3(\xi u_{2h}^{n+1} + (1 - \xi) u_{2h}^n) (u_{2h}^{n+1} - u_{2h}^n) d\xi \right] dx = 0.
\end{aligned} \tag{2.43}$$

In order to simplify the integral terms in (2.43), we make the following change of variables

$$\begin{aligned}
z_{1h} &= \xi u_{1h}^{n+1} + (1 - \xi) u_{1h}^n, \\
z_{2h} &= \xi u_{2h}^{n+1} + (1 - \xi) u_{2h}^n.
\end{aligned}$$

Then the integral terms in (2.43) can be simplified as

$$\begin{aligned}
\int_0^1 (f(\xi u_1^{n+1} + (1-\xi)u_1^n)(u_1^{n+1} - u_1^n) d\xi &= \int_{u_1^n}^{u_1^{n+1}} f(z_{1h}) dz_{1h}, \\
\int_0^1 (\ell_1(\xi u_2^{n+1} + (1-\xi)u_2^n)(u_1^{n+1} - u_1^n) d\xi &= \int_{u_1^n}^{u_1^{n+1}} \ell_1(z_{2h}) dz_{1h}, \\
\int_0^1 (\ell_2(\xi u_1^{n+1} + (1-\xi)u_1^n)(u_2^{n+1} - u_2^n) d\xi &= \int_{u_1^n}^{u_1^{n+1}} \ell_2(z_{1h}) dz_{2h}, \\
\int_0^1 (\ell_3(\xi u_2^{n+1} + (1-\xi)u_2^n)(u_2^{n+1} - u_2^n) d\xi &= \int_{u_1^n}^{u_1^{n+1}} \ell_3(z_{2h}) dz_{2h}.
\end{aligned} \tag{2.44}$$

If we add the first two equations of (2.44), we obtain

$$\begin{aligned}
\int_{u_1^n}^{u_1^{n+1}} f(z_{1h}) dz_{1h} + \int_{u_1^n}^{u_1^{n+1}} \ell_1(z_{2h}) dz_{1h} &= \int_{u_1^n}^{u_1^{n+1}} (f(z_{1h}) + \ell_1(z_{1h})) dz_{1h} \\
&= \int_{u_1^n}^{u_1^{n+1}} (f_1(z_{1h}, z_{2h})) dz_{1h} = \frac{\partial F}{\partial u_{1h}}
\end{aligned} \tag{2.45}$$

using the property of skew-gradient, see subsection 2.21.

In a similar way, the last two equation of (2.44) results in

$$\begin{aligned}
\int_{u_1^n}^{u_1^{n+1}} \ell_2(z_{1h}) dz_{1h} + \int_{u_1^n}^{u_1^{n+1}} \ell_3(z_{2h}) dz_{1h} &= \int_{u_1^n}^{u_1^{n+1}} (\ell_2(z_{1h}) + \ell_3(z_{1h})) dz_{1h} \\
&= \int_{u_1^n}^{u_1^{n+1}} (f_2(z_{1h}, z_{2h})) dz_{1h} = -\frac{\partial F}{\partial u_{2h}}.
\end{aligned} \tag{2.46}$$

Therefore, when we substitute (2.45) and (2.46) into (2.43), we obtain

$$\begin{aligned}
&\left(\frac{1}{\Delta t} \|u_1^{n+1} - u_1^n\|_{L^2(\Omega)}^2 + \frac{1}{2} a_h(d_1; u_1^{n+1}, u_1^{n+1}) - \frac{1}{2} a_h(d_1; u_1^n, u_1^n) \right) \\
&- \left(\frac{1}{\Delta t} \|u_2^{n+1} - u_2^n\|_{L^2(\Omega)}^2 + \frac{1}{2} a_h(d_2; u_2^{n+1}, u_2^{n+1}) - \frac{1}{2} a_h(d_2; u_2^n, u_2^n) \right) \\
&\approx ((F(u_1^{n+1}, u_2^{n+1}), 1)_\Omega - (F(u_1^n, u_2^n), 1)_\Omega)
\end{aligned} \tag{2.47}$$

This is equivalent to

$$\begin{aligned} E(u_{1h}^{n+1}, u_{2h}^{n+1}) - E(u_{1h}^n, u_{2h}^n) &\approx -\frac{1}{\Delta t} \|u_{1h}^{n+1} - u_{1h}^n\|_{L^2(\Omega)}^2 \\ &+ \frac{1}{\Delta t} \|u_{2h}^{n+1} - u_{2h}^n\|_{L^2(\Omega)}^2. \end{aligned} \quad (2.48)$$

Now, at the steady state point $(u_{1h}^n, u_{2h}^n) = (\xi, \eta)$, if u_{2h}^n is fixed to $\eta(x)$, then

$$\frac{1}{\Delta t} \|u_{2h}^{n+1} - u_{2h}^n\|_{L^2(\Omega)}^2 = 0.$$

So, from (2.48)

$$E(\xi_{n+1}, \eta) \leq E(\xi_n, \eta).$$

Similarly, if u_{1h}^n is fixed to $\xi(x)$, then

$$E(\xi, \eta_{n+1}) \geq E(\xi, \eta_n).$$

Thus, the first equation of (2.1) describes a gradient flow with $E(u_{1h}^n, \eta)$ and the second equation of (2.1) describes a gradient flow with $-E(\xi, u_{2h}^n)$, meaning that $(u_{1h}^n, u_{2h}^n) = (\xi, \eta)$ is stable as a steady state of (2.1) since it is a mini-maximizer of $E(u_{1h}^n, u_{2h}^n)$ [100].

2.5 Numerical Results

In this section, we provide numerical results for 1D-2D diffusive and 2D convective FHNEs.

2.5.1 Diffusive FHNE in 1D

As described in subsection 2.1.1, travelling fronts and pulses occur for 1D diffusive FHNE.

2.5.1.1 Travelling Fronts

We consider the diffusive FHNE (2.12) on a spatial domain $\Omega = [-60, 60]$ with the spatial mesh size $\Delta x = 0.1$ and the temporal step size $\Delta t = 0.5$. The parameter set is chosen as

$$(\tau_1, \tau_2, d_1, d_2, \beta, \epsilon, \gamma) = \left(1, \frac{25}{2}, 1, \frac{5}{4}, \frac{1}{3}, \frac{7}{10}, 8\right).$$

We set the initial conditions for travelling fronts

$$\begin{aligned} u_1(x, 0) &= \tanh(x), \\ u_2(x, 0) &= 1 - \tanh(x). \end{aligned}$$

In the following Figure 2.1, travelling fronts are shown in the whole space-time domain. As it is seen, fronts are moving in time. To understand better, Figure 2.2 displays the fronts at some different times. The results we obtained here are very similar to our reference solutions in [29, 17].

The discrete energy is also shown in Figure 2.1. It is decaying very slowly after a sharp increase at the beginning.

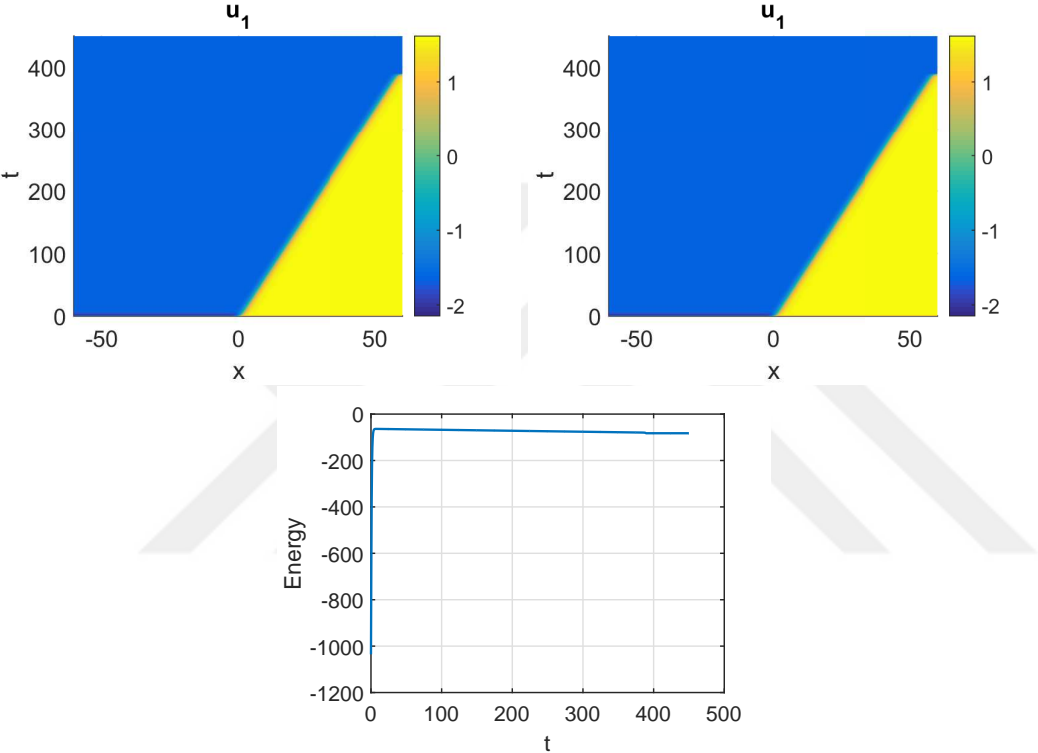


Figure 2.1: Travelling front solutions for both state component u_1, u_2 (top), evolution of the discrete energy (bottom).

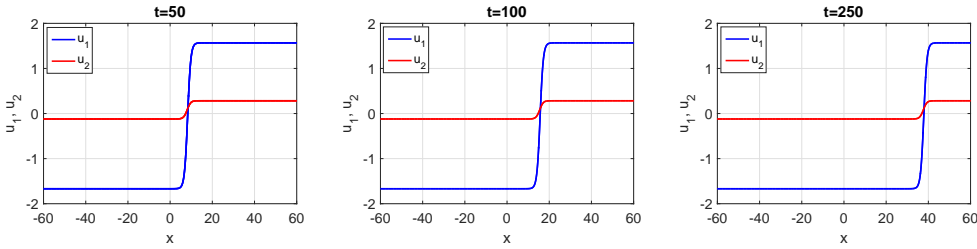


Figure 2.2: Travelling front solutions at different time instances.

2.5.1.2 Travelling Pulses

For travelling pulses, all parameters other than $\gamma = 0.8$ are the same with travelling fronts. Initial conditions are taken as

$$\begin{aligned} u_1(x, 0) &= \tanh(x), \\ u_2(x, 0) &= -0.6. \end{aligned}$$

The travelling pulse solutions obtained for $\Delta x = 0.1$ and $\Delta t = 0.1$ are shown in Figures 2.3- 2.4. As it is seen, the travelling pulse solutions are moving in time. The discrete energy remains constant after an oscillation. Again, the results are good agreement with those in [29, 17].

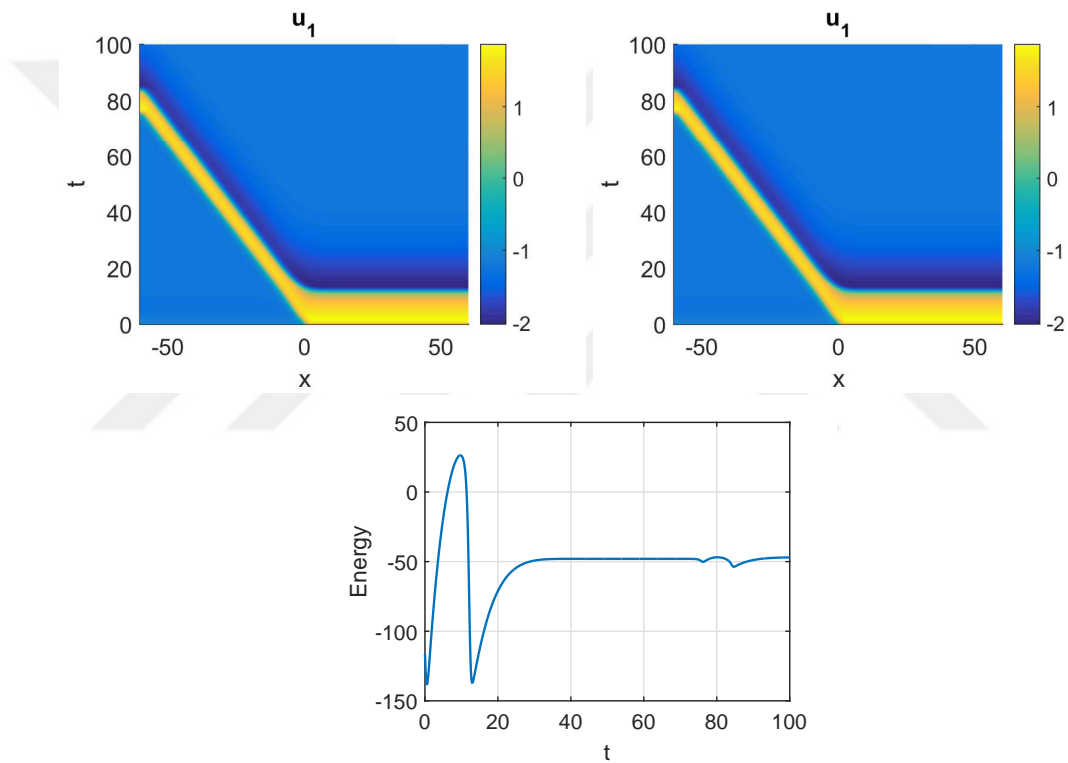


Figure 2.3: Travelling pulse solutions for both state component u_1, u_2 (top), evolution of the discrete energy (bottom).

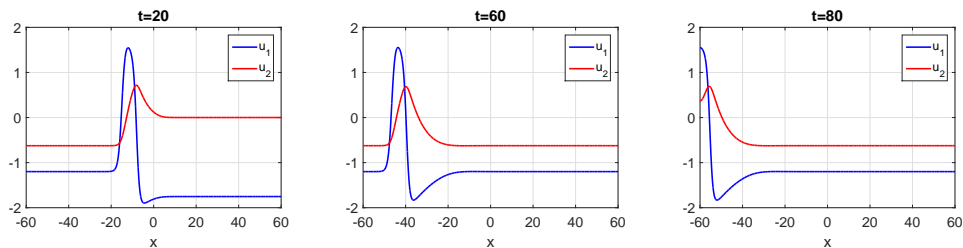


Figure 2.4: Travelling pulse solutions at different time instances.

2.5.2 Diffusive FHNE in 2D

As described in subsection 2.1.2, we have spot and labyrinth-like pattern solutions depending on the value of κ for the 2D diffusive FHNE (2.14). We consider the equation (2.14) on the spatial domain $\Omega = [-1, 1]^2$ with the spatial mesh size $\Delta x = \Delta y = 1/32$ and the temporal step size $\Delta t = 1/10$. We fix the following set of parameters for both pattern solutions

$$(\tau_1, \tau_2, d_1, d_2, \kappa) = (1, 1, 0.00028, 0.005, 0 \text{ or } 0.1)$$

satisfying the skew-gradient condition (2.21) and “*Turing instability*” conditions (2.11). The initial conditions are uniformly distributed randomly numbers between -1 and 1 .

The results agree with the reference solutions in [61]. The energy plots show that the spot and labyrinth-like patterns reach the steady state after $t \approx 150$ and $t \approx 20$, respectively.

2.5.2.1 Spot Patterns

For $\kappa = 0.1$, the pattern formation of diffusive FHNE (2.14) will be spots as it is seen in Figure 2.5.

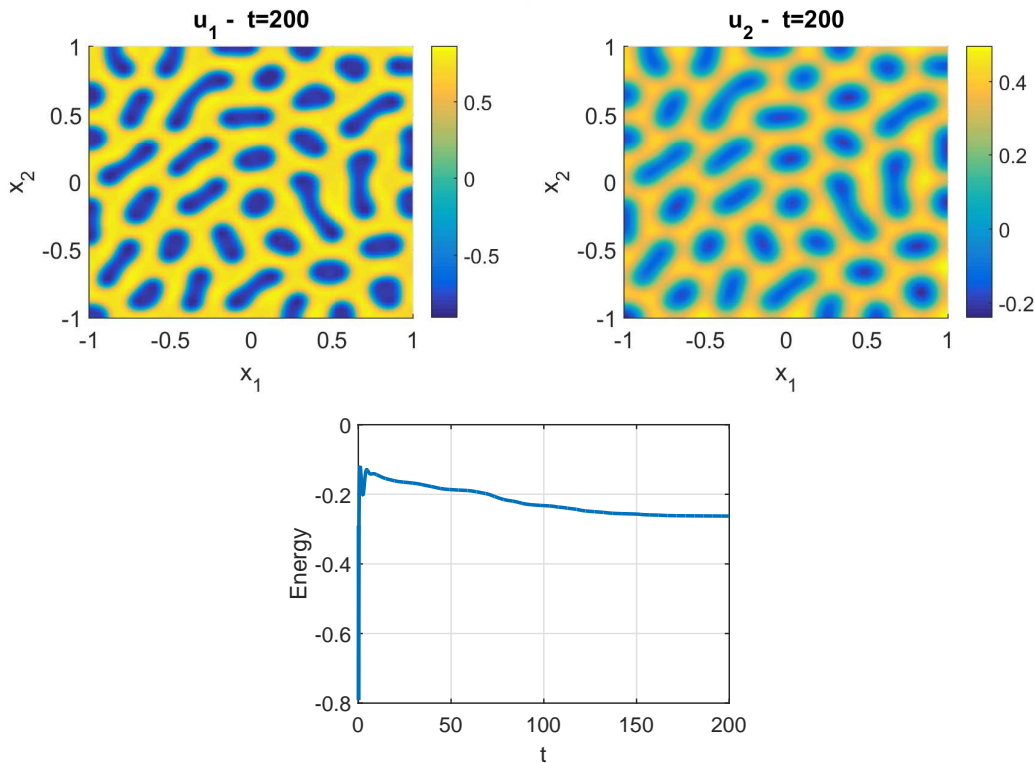


Figure 2.5: Spot patterns solutions for both state component u_1, u_2 (top), evolution of the discrete energy (bottom).

2.5.2.2 Labyrinth-like Patterns

For $\kappa = 0$, the pattern formation of the diffusive FHNE (2.14) will be labyrinth-like patterns as it shown in Figure 2.6.

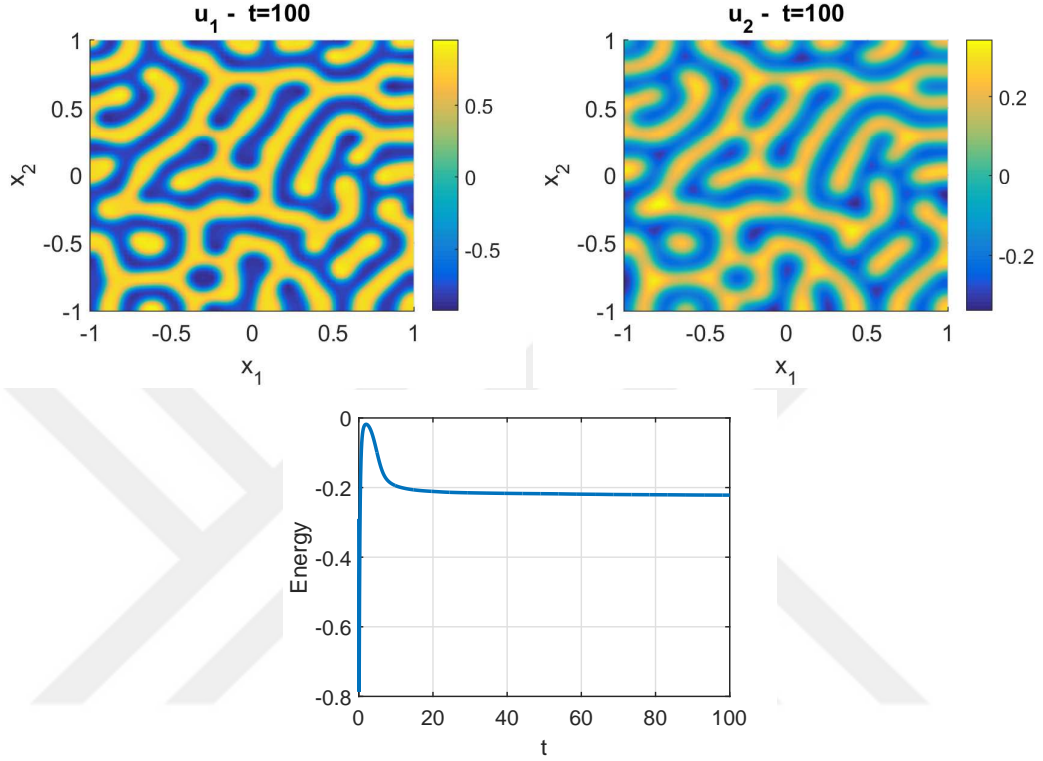


Figure 2.6: Labyrinth-like patterns solutions for both state component u_1 , u_2 (top), evolution of the discrete energy (bottom).

2.5.3 Three Component Diffusive FHNE in 1D

As described in subsection 2.1.3, we have one-pulse (two-front), two-pulse (four front), and multi-pulse (multi-front) solutions for triple 1D diffusive FHNE (2.15). The reference solutions can be found in [93, 84, 94].

2.5.3.1 One-pulse

We consider the equation (2.15) on the spatial domain $\Omega = [-1000, 1000]$ with the spatial and temporal step sizes $\Delta x = \Delta t = 0.5$. For one pulse and two pulse solutions we fix the following set of parameters

$$(\alpha, \beta, \epsilon, \gamma, d, \tau_1, \tau_2, \tau_3) = (3, 1, \frac{1}{100}, -\frac{1}{4}, 5, 1, \frac{100}{3}, 100)$$

satisfying the skew-gradient condition (2.21) and “*Turing instability*” conditions (2.11) under the homogeneous Neumann boundary conditions on both ends in one dimensional spatial domain.

We set the initial conditions for one-pulse solutions as

$$u(x, 0) = \begin{cases} 1 & \text{if } x \in [-50, 50] \\ -1 & \text{otherwise} \end{cases}, \quad v(x, 0) = s(x, 0) = 0.$$

The solution profiles up to final time $T = 200$ and at $t = 25, 100, 200$ together with the energy plots are shown in Figures 2.7- 2.8.

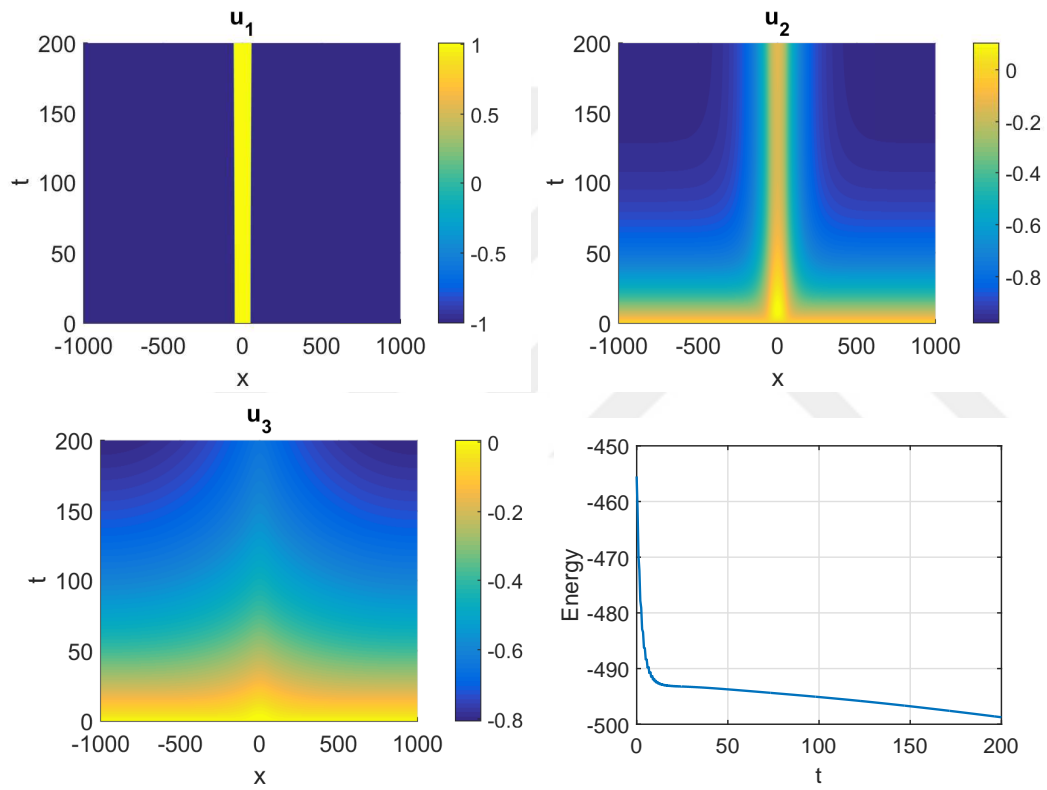


Figure 2.7: One-pulse solutions for three state components u_1, u_2, u_3 , evolution of the discrete energy.

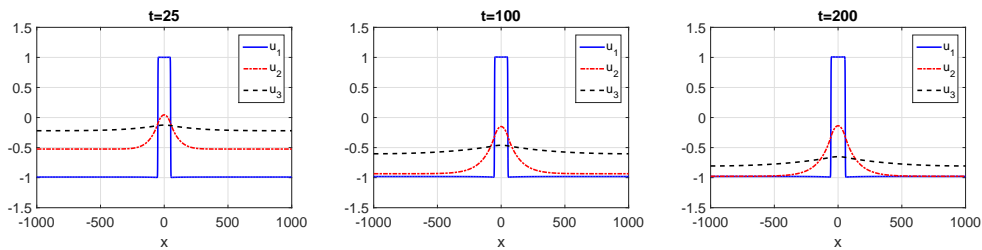


Figure 2.8: One-pulse solutions at different time instances.

2.5.3.2 Two-pulse

For two-pulse solutions, we set the initial conditions as

$$u(x, 0) = \begin{cases} 1 & \text{if } x \in [-350, -150] \cup [150, 350] \\ -1 & \text{otherwise} \end{cases}, \quad v(x, 0) = s(x, 0) = 0.$$

The solution profiles up to $T = 100$ and at $t = 20, 50, 100$ together with the energy plots, are shown in Figures 2.9- 2.10.

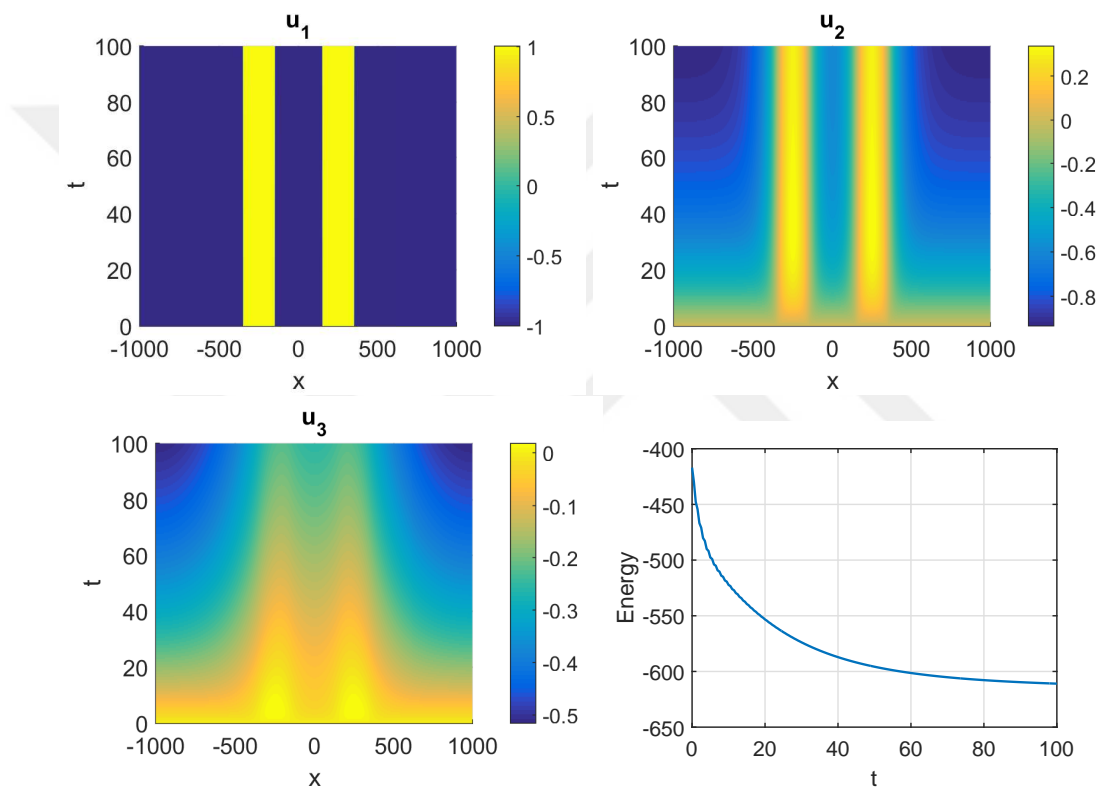


Figure 2.9: Two-pulse solutions for three state components u_1, u_2, u_3 , evolution of the discrete energy.

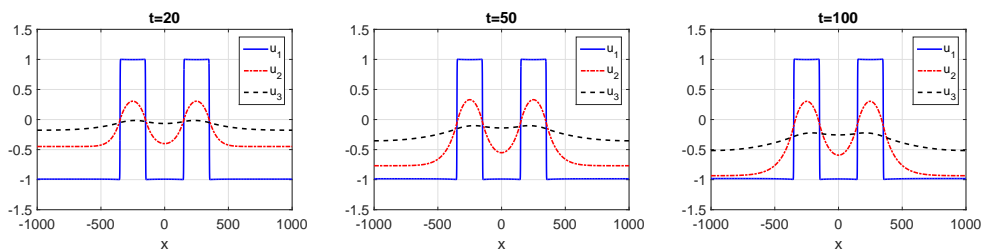


Figure 2.10: Two-pulse solutions at different time instances.

2.5.3.3 Multi-pulse

For multi-pulse solutions, the set of parameters is chosen

$$(\alpha, \beta, \epsilon, \gamma, d, \tau_1, \tau_2, \tau_3) = (100, 100, \frac{1}{100}, -\frac{1}{4}, 5, 1, 1, 1)$$

with the following initial conditions

$$u(x, 0) = \begin{cases} 0 & \text{if } x \in [50, 50] \\ -1 & \text{otherwise} \end{cases}, \quad v(x, 0) = s(x, 0) = 0.$$

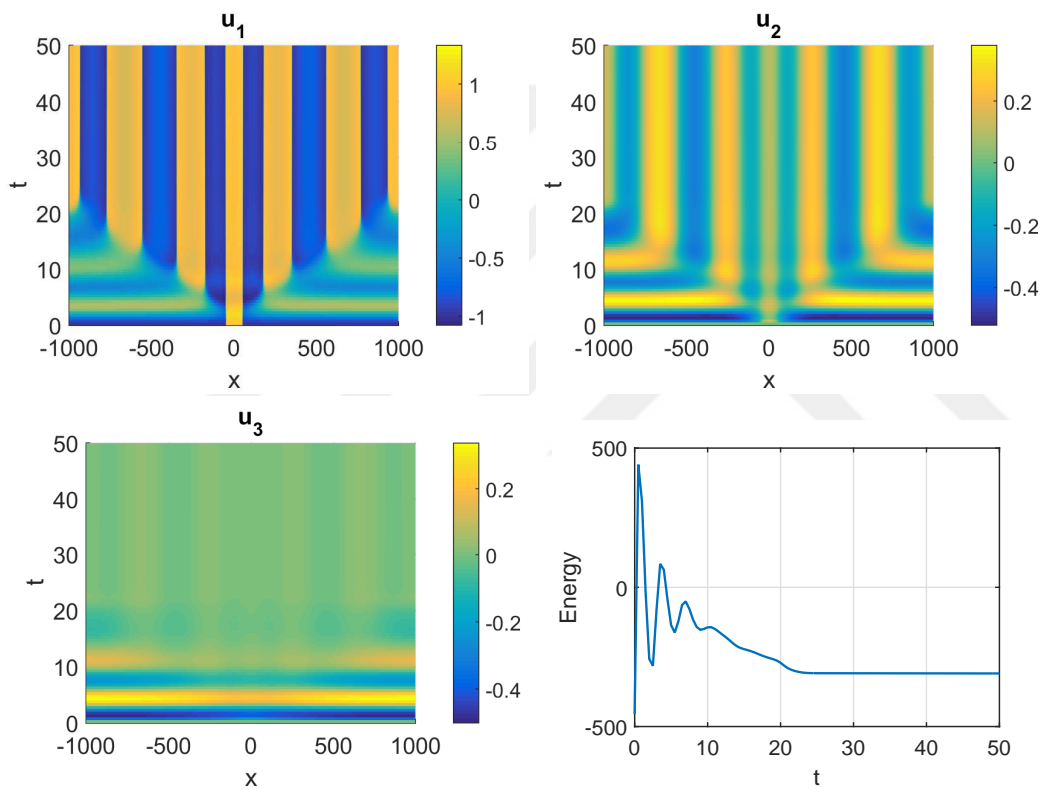


Figure 2.11: Multi-pulse solutions for three state components u_1 , u_2 , u_3 , evolution of the discrete energy.

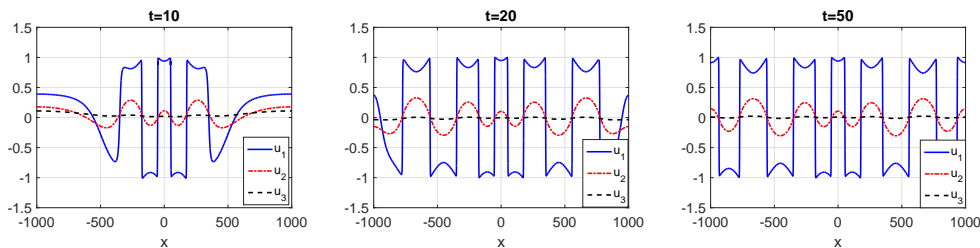


Figure 2.12: Multi-pulse solutions at different time instances.

The solution profiles up to $T = 50$ and at $t = 10, 20, 50$ together with the energy plots, are given in Figures 2.11- 2.12. After an oscillation, the discrete energy remains constant, which means that the solution reaches the steady state at $t \approx 22$.

2.5.4 Convective FHNE in 2D

As described in subsection 2.2.1, we consider the convective FHNE (2.23) on the space-time domain $\Omega = [0, L] \times [0, H]$ with the spatial mesh size $\Delta x_1 = \Delta x_2 = 0.5$ and the temporal mesh size $\Delta t = 0.05$. We choose the parameters as [35]

$$(\tau_1, \tau_2, d_2, c_1, c_2, c_3, \epsilon) = (1, 1, 1, 1, 9, 0.02, 5, 0.1)$$

satisfying the mono-stability condition (2.27).

We have obtained similar results, shown in Figures 2.13- 2.14, as obtained in [35].

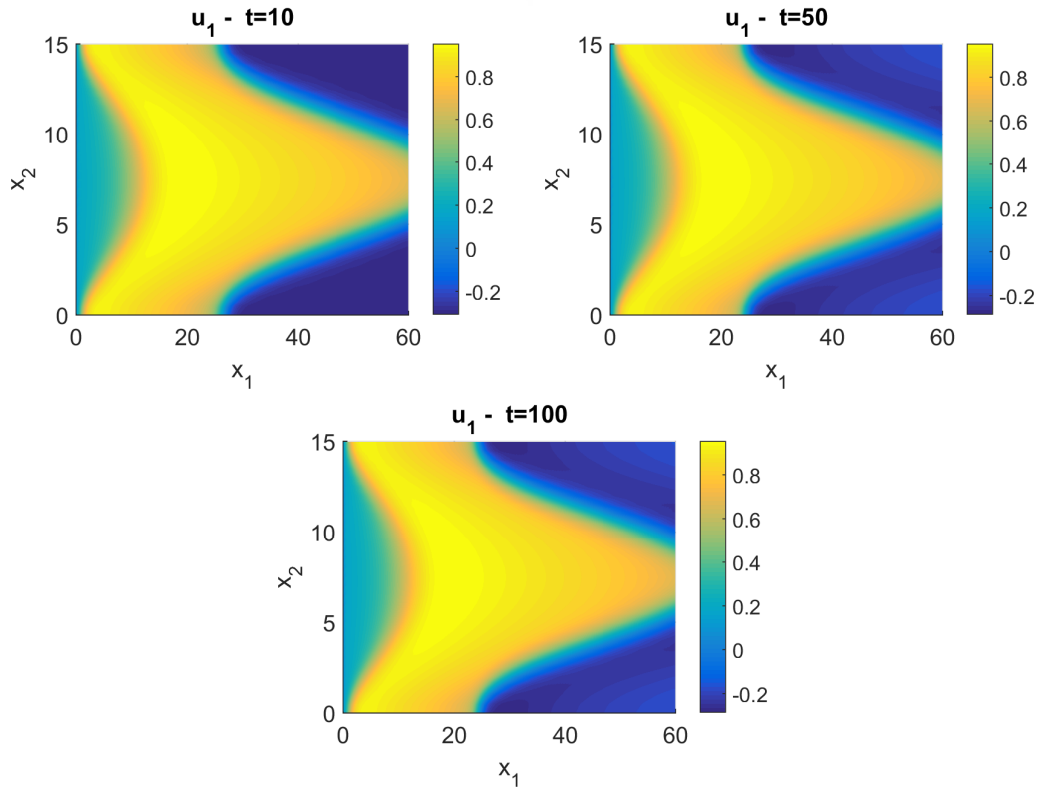


Figure 2.13: Wave-type pattern solutions for state u_1 at different time instances.

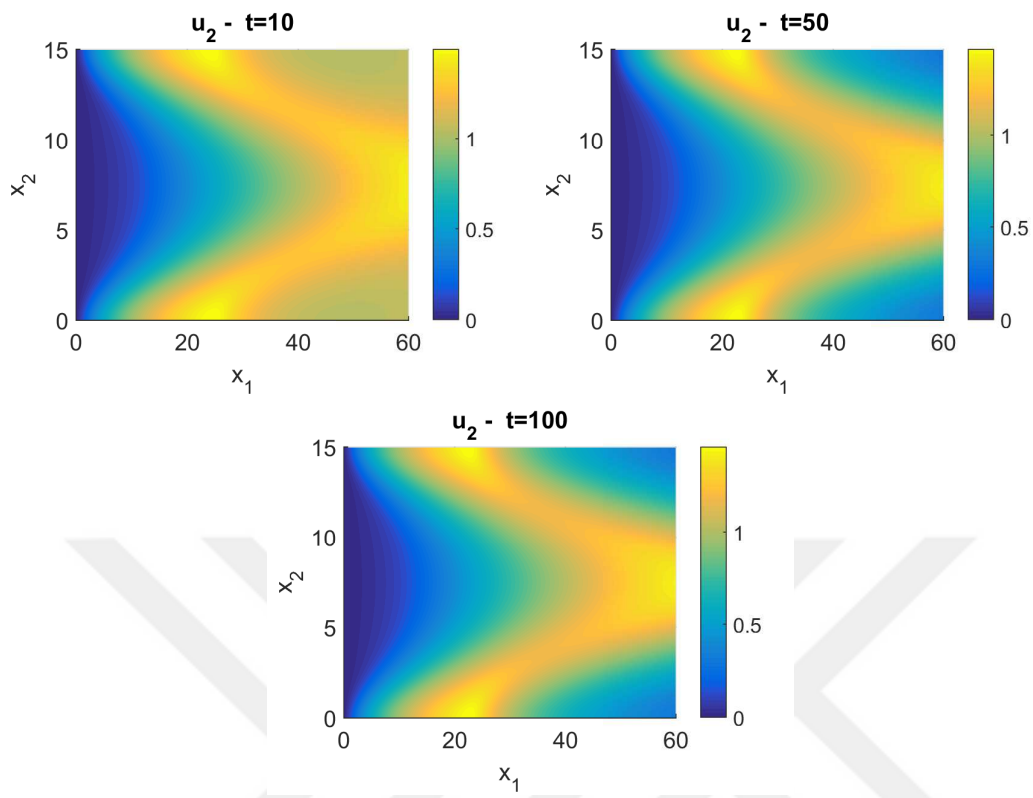


Figure 2.14: Wave-type pattern solutions for state u_2 at different time instances.

CHAPTER 3

REDUCED ORDER MODELLING

Discretization of PDEs by finite differences, finite volumes, spectral methods or finite element methods, leads to very large algebraic systems. The high-fidelity or full order approximate solutions require very long computational times. When the problem involves the repetitive solution of the PDEs for different parameters (parametrized PDEs), or PDE-constrained optimization, optimal control and design, parameter estimation, and inverse problems, then CPU time increases more. In the last two decades, reduced order modelling (ROM) has emerged as an efficient way to reduce computing time and produce reliable solutions. A common purpose of the model order reduction (MOR) techniques is to reduce the computational time by approximating the large-scale sized full order models (FOMs) with lower dimensional models essentially capturing their dynamics. There exists a variety of ROM methods: balanced truncation, proper orthogonal decomposition, moment matching methods, Krylov subspace based methods [6].

The proper orthogonal decomposition (POD) method, also known as Karhunen-Loève decomposition or principal component analysis, is one of the most popular and successful MOR technique for linear and nonlinear PDEs. The POD is used to project the dynamical system onto subspaces comprising of basis elements inheriting the main characteristics of the expected or unknown solution. In the literature, the POD method has been successfully used in many areas including coherent structures [46, 82], fluid dynamics [55, 57]. It is based on the snapshots, i.e. FOM solutions of the discretized PDE based method. The POD is optimal in terms of energy content [82]. Even though the POD is a heuristic method, i.e. there exists no a-priori error estimate, it is used widely in the literature. Despite the fact that the POD is extremely effective MOR technique for linear problems, for nonlinear problems the computational complexity of the evaluation of the nonlinear terms still depends on the dimension of FOM. The empirical interpolation method [28], discrete empirical interpolation method (DEIM) [13] are the mostly used techniques to overcome this issue. The DEIM was originally developed for nonlinear functions which depend componentwise on single variable, arising from the finite difference discretization of nonlinear PDEs. The nonlinear terms are computed at selected DEIM interpolation points from the POD basis of the nonlinear function in a greedy algorithm, then interpolation and projection are combined to derive approximation in the low dimensional POD-DEIM space. The cost and overall quality of the ROM is directly influenced by the POD-DEIM. For the finite element discretization, the nonlinear functions depend on the mesh and on the poly-

mial degree of the finite elements. Therefore the efficiency of the POD-DEIM can be degraded. Using the unassembled finite elements such that each DEIM point is related to one element, the number of nonlinear function calls during the online computation can be reduced, but the size of the nonlinear snapshots are enlarged, which increases the offline computational cost [5]. For the DG discretization, each component of the nonlinear functions depends only on the local mesh in contrast to the continuous finite element discretization where the nonlinear function depends on multiple components of the state vector. Therefore the number of POD-DEIM function evaluations for DG discretization is comparable with the finite elements discretization, which will be illustrated in the next sections.

The Dynamic Mode Decomposition (DMD) method [60] is emerged in the recent years as an alternative to the above mentioned projection based ROM techniques. The DMD was first introduced by Schmid [80] and Rowley [76] for the time evaluation of fluid flows; however, it is also related to the Koopman operator [52]. The Koopman operator is a linear infinite dimensional operator that represents and provides information about nonlinear, finite dimensional dynamics of a dynamical system. The DMD can be considered as an approximation of the Koopman operator. The DMD can use the snapshots obtained by numerical simulations, but at the same time it is a data driven MOR technique, using for example experimental data. The DMD is equation-free, i.e. there is no need to solve small sized reduced order systems as in case of the POD. We apply here the DMD to approximate the nonlinear terms as an alternative to the DEIM [3].

In Section 3.1, application of the POD to the diffusive FHNE is described for the DG discretization. The DEIM under DG discretization is given in Section 3.2. The POD-DMD approximation of the nonlinear terms are described in Section 3.3. The chapter is concluded with some numerical results for the diffusive (2.1) and convective FHNEs (2.23), comparing the ROM solutions for SIPG discretization with application of the POD, POD-DEIM and POD-DMD in terms of computational time and accuracy in Section 3.5.

3.1 Proper Orthogonal Decomposition

In the following, we consider the SIPG semi-discretized form

$$\begin{aligned}
\tau_1 \mathbf{M} \frac{d\vec{u}_1}{dt} &= \mathbf{A}_{u_1} \vec{u}_1 + \mathbf{f}(\vec{u}_1) - \mathbf{M} \vec{u}_2 + \mathbf{K}, \\
\tau_2 \mathbf{M} \frac{d\vec{u}_2}{dt} &= \mathbf{A}_{u_2} \vec{u}_2 - \mathbf{M} \vec{u}_2 + \mathbf{M} \vec{u}_1, \\
\vec{u}_1(\cdot, 0) &= \vec{u}_{10}, \\
\vec{u}_2(\cdot, 0) &= \vec{u}_{20},
\end{aligned} \tag{3.1}$$

of the 2D diffusive FHNE

$$\begin{aligned}
\tau_1 u_{1t} &= d_1 \Delta u_1 + f(u_1) - u_2 + \kappa, \\
\tau_2 u_{2t} &= d_2 \Delta u_2 - u_2 + u_1,
\end{aligned} \tag{3.2}$$

where $\vec{u}_{10}, \vec{u}_{20} \in \mathbb{R}^N$ denote the given initial data. $\mathbf{A}_{u_1}, \mathbf{A}_{u_2} \in \mathbb{R}^{N \times N}$ are the stiffness matrices, $\mathbf{M} \in \mathbb{R}^{N \times N}$ is the mass matrix, $\mathbf{K} \in \mathbb{R}^N$ is the constant vector related to the integral containing the parameter κ , and $\mathbf{f} \in \mathbb{R}^N$ is the vector of the bi-stable nonlinear terms of $\vec{u}_1 = (u_{11}, u_{22}, \dots, u_{NN})$.

We note that the nonlinear vector \mathbf{f} is uniformly Lipschitz continuous with respect to \vec{u}_1 with a Lipschitz constant $L_F = 1 > 0$, i.e., $\forall \vec{u}_{11}(t), \vec{u}_{12}(t) \in \mathbb{R}^N$ and $t \in [0, T]$,

$$\|\mathbf{f}(\vec{u}_{11}(t)) - \mathbf{f}(\vec{u}_{12}(t))\| \leq L_F \|\vec{u}_{11}(t) - \vec{u}_{12}(t)\|. \quad (3.3)$$

The Lipschitz continuity (3.3) ensures that (3.1) has a unique solution for any t values.

The diffusive FHNE (3.2) has spot and labyrinth-like patterns depending on the values of κ . Because the computation of those Turing patterns for equation (3.1) is time consuming, we develop a POD reduced order model [54] for the SIPG semi-discretized system (3.1) of small dimension $k \ll N$.

The semi-discrete solutions of (3.1) in ordered form using the basis functions are given as

$$w(t) = \underbrace{\sum_{j=1}^{n_e} \sum_{i=1}^{n_k} w_i^j(t) \phi_i^j(x)}_{\text{Usual}} \implies w(t) = \underbrace{\sum_{i=1}^N w_i(t) \phi_i(x)}_{\text{Ordered}} = \phi \mathbf{w}(t)$$

with $w_i \in \mathbb{R}$, $\phi_i \in \mathbb{R}^N$, $i = 1, 2, \dots, N$, and

$$\mathbf{w}(t) = (w_1^1(t), w_2^1(t), \dots, w_{n_k}^1(t), \dots, w_1^{n_e}(t), w_2^{n_e}(t), \dots, w_{n_k}^{n_e}(t))^T$$

for $\mathbf{w} \in \{\vec{u}_1, \vec{u}_2\}$, and

$$\phi = (\phi_1^1, \phi_2^1, \dots, \phi_{n_k}^1, \dots, \phi_1^{n_e}, \phi_2^{n_e}, \dots, \phi_{n_k}^{n_e})^T.$$

The reduced-order system of (3.1) is constructed using the Galerkin projection which approximates $u_1(t)$ and $u_2(t)$ of the full system (3.1) from the subspaces, spanned by sets of orthogonal basis $\{\psi_{u_1,i}\}_{i=1}^k$ and $\{\psi_{u_2,i}\}_{i=1}^k$ of dimension k in \mathbb{R}^N :

$$u_1(t) \approx \sum_{i=1}^k \tilde{u}_{1i}(t) \psi_{u_1,i}, \quad u_2(t) \approx \sum_{i=1}^k \tilde{u}_{2i}(t) \psi_{u_2,i},$$

where $\vec{\tilde{u}}_1 = (\tilde{u}_{11}(t), \dots, \tilde{u}_{1k}(t))^T$ and $\vec{\tilde{u}}_2 = (\tilde{u}_{21}(t), \dots, \tilde{u}_{2k}(t))^T$ are the solutions of the reduced system. On the other hand, because the basis functions are spanned by the DG basis functions $\{\phi_i\}_{i=1}^N$, the functions $\{\psi_{u_1,i}\}$ and $\{\psi_{u_2,i}\}$ are given by

$$\psi_{u_1,i} = \sum_{j=1}^N \Psi_{u_1,j,i} \phi_j(x), \quad \psi_{u_2,i} = \sum_{j=1}^N \Psi_{u_2,j,i} \phi_j(x),$$

where the i -th columns of the matrices $\Psi_{u_1} = [\Psi_{u_1,\cdot,1}, \dots, \Psi_{u_1,\cdot,k}] \in \mathbb{R}^{N \times k}$ and $\Psi_{u_2} = [\Psi_{u_2,\cdot,1}, \dots, \Psi_{u_2,\cdot,k}] \in \mathbb{R}^{N \times k}$ are the coefficient vectors of the i -th reduced basis functions $\psi_{u_1,i}$ and $\psi_{u_2,i}$, respectively.

In order to find the orthogonal (in L^2 -sense) basis functions $\{\psi_{u_1,i}\}$ and $\{\psi_{u_2,i}\}$, $i = 1, 2, \dots, k$, the matrices Ψ_{u_1} and Ψ_{u_2} are constructed from the coefficient vectors through an application of the POD method [54].

The snapshot matrices are defined as

$$U_1 = [\vec{u}_1^1, \dots, \vec{u}_1^J], \quad U_2 = [\vec{u}_2^1, \dots, \vec{u}_2^J],$$

where $U_1, U_2 \in \mathbb{R}^{N \times J}$, and each member of the snapshot matrices are the corresponding coefficient vectors of the discrete solutions $\{u_1^i\}_{i=1}^J$ and $\{u_2^i\}_{i=1}^J$, respectively, of the FOM (3.1) at the time trajectories t_i , $i = 0, 1, \dots, J$ with $u_1^i \approx u_1(t_i)$ and $u_2^i \approx u_2(t_i)$.

Definition 3.1. For $w \in \{u_1, u_2\}$, the L^2 -orthogonal POD basis functions $\{\psi_{w,i}\}$, $i = 1, 2, \dots, k$, are given by the solution of the minimization problem

$$\min_{\psi_{w,1}, \dots, \psi_{w,k}} \frac{1}{J} \sum_{j=1}^J \left\| w^j - \sum_{i=1}^k (w^j, \psi_{w,i})_{L^2(\Omega)} \psi_{w,i} \right\|_{L^2(\Omega)}^2 \quad (3.4)$$

$$\text{subject to } (\psi_{w,i}, \psi_{w,j})_{L^2(\Omega)} = \Psi_{w,i}^T M \Psi_{w,j} = \delta_{ij}, \quad 1 \leq i, j \leq k, \quad (3.5)$$

where δ_{ij} is the Kronecker delta, i.e., $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ otherwise.

Then, the coefficient vectors $\Psi_{u_1,i}$ and $\Psi_{u_2,i}$ of the POD basis functions $\psi_{u_1,i}$ and $\psi_{u_2,i}$, respectively, are given by the solutions of the following eigenvalue problem [54]

$$U_1 U_1^T M \Psi_{u_1,i} = \sigma_{u_1,i}^2 \Psi_{u_1,i}, \quad U_2 U_2^T M \Psi_{u_2,i} = \sigma_{u_2,i}^2 \Psi_{u_2,i}, \quad i = 1, 2, \dots, k. \quad (3.6)$$

Defining $\widehat{U}_1 = R U_1$ and $\widehat{U}_2 = R U_2$ with the lower triangular matrix R^T as the Cholesky factor of the mass matrix M , $M = R^T R$, we obtain the equivalent formulations

$$\widehat{U}_1 \widehat{U}_1^T \widehat{\Psi}_{u_1,i} = \sigma_{u_1,i}^2 \widehat{\Psi}_{u_1,i}, \quad \widehat{U}_2 \widehat{U}_2^T \widehat{\Psi}_{u_2,i} = \sigma_{u_2,i}^2 \widehat{\Psi}_{u_2,i}, \quad i = 1, 2, \dots, k \quad (3.7)$$

with $\widehat{\Psi}_{\cdot,i} = R \Psi_{\cdot,i}$. The solution $\widehat{\Psi}_{\cdot,i}$ of the above problems (3.7) can be computed using the first k left singular vectors $\zeta_{u_1,i}$ and $\zeta_{u_2,i}$ of the singular value decomposition (SVD) of the matrices \widehat{U}_1 and \widehat{U}_2 , respectively:

$$\widehat{U}_1 = \zeta_{u_1} \Sigma_{u_1} \beta_{u_1}^T, \quad \widehat{U}_2 = \zeta_{u_2} \Sigma_{u_2} \beta_{u_2}^T,$$

and the diagonal matrices Σ_{u_1} and Σ_{u_2} contain the singular values $\sigma_{u_1,i}$ and $\sigma_{u_2,i}$ on their diagonals, respectively. Then, since $\widehat{\Psi}_{\cdot,i} = R \Psi_{\cdot,i}$, the coefficient vectors $\Psi_{\cdot,i}$ of the POD basis functions are computed as

$$\Psi_{u_1,i} = R^{-1} \zeta_{u_1,i}, \quad \Psi_{u_2,i} = R^{-1} \zeta_{u_2,i}, \quad i = 1, 2, \dots, k.$$

Note also that for $i \neq j$

$$\delta_{ij} = (\widehat{\Psi}_{\cdot,i}, \widehat{\Psi}_{\cdot,j}) = \widehat{\Psi}_{\cdot,i}^T \widehat{\Psi}_{\cdot,j} = \Psi_{\cdot,i}^T M \Psi_{\cdot,j} = (\psi_{\cdot,i}, \psi_{\cdot,j})_{L^2(\Omega)},$$

which means that the coefficient vectors of the POD basis functions satisfy the M-orthogonality.

Meanwhile, letting $\phi = [\phi_1, \dots, \phi_N] \in \mathbb{R}^{N \times N}$ and using the corresponding linear combinations, we have for $w \in \{u_1, u_2\}$

$$\begin{aligned} w(t) &= \psi_w \tilde{\mathbf{w}}, \quad w(t) = \phi \mathbf{w}, \quad \psi_w = \phi \Psi_w \\ &\Rightarrow \quad \phi \mathbf{w} = \phi \Psi_w \tilde{\mathbf{w}} \\ &\Rightarrow \quad \mathbf{w} = \Psi_w \tilde{\mathbf{w}}. \end{aligned} \quad (3.8)$$

On the other hand, using the fact that the singular vectors $\widehat{\Psi}_{u_1, \cdot, i}$ of (3.7) are orthogonal and that $\Psi_{w, \cdot, i} = R^{-1} \widehat{\Psi}_{w, \cdot, i}$, we obtain

$$\begin{aligned} \tilde{\mathbf{w}} &= \Psi_w^{-1} \mathbf{w} = \widehat{\Psi}_w^{-1} R \mathbf{w} = \widehat{\Psi}_w^T R \mathbf{w} = \Psi_w^T R^T R \mathbf{w} = \Psi_w^T M \mathbf{w} \\ &\Rightarrow \quad \tilde{\mathbf{w}} = \Psi_w^T M \mathbf{w}. \end{aligned} \quad (3.9)$$

Now, in order to obtain the k -dimensional reduced systems, we substitute for $\mathbf{w} \in \{\vec{u}_1, \vec{u}_2\}$ the relation (3.8) in the full system (3.1) and we project the systems onto the reduced spaces spanned by $\{\psi_{u_1, 1}, \dots, \psi_{u_1, k}\}$ and $\{\psi_{u_2, 1}, \dots, \psi_{u_2, k}\}$, respectively, to get

$$\begin{aligned} \Psi_{u_1}^T M \Psi_{u_1} \vec{u}_1 t &= \Psi_{u_1}^T A_{u_1} \Psi_{u_1} \vec{u}_1 + \Psi_{u_1}^T F(\Psi_{u_1} \vec{u}_1) - \Psi_{u_1}^T M \Psi_{u_2} \vec{u}_2 + \Psi_{u_1}^T K, \\ \Psi_{u_2}^T M \Psi_{u_2} \vec{u}_2 t &= \Psi_{u_2}^T A_{u_2} \Psi_{u_2} \vec{u}_2 - \Psi_{u_2}^T M \Psi_{u_2} \vec{u}_2 + \Psi_{u_2}^T M \Psi_{u_1} \vec{u}_1. \end{aligned}$$

Or using the M-orthogonality of the POD basis coefficients, we obtain

$$\begin{aligned} \vec{u}_1 &= \tilde{A}_{u_1} \vec{u}_1 + \Psi_{u_1}^T F(\Psi_{u_1} \vec{u}_1) - \tilde{M}_{u_1} \vec{u}_2 + \Psi_{u_1}^T K, \\ \vec{u}_2 &= \tilde{A}_{u_2} \vec{u}_2 - \vec{u}_2 + \tilde{M}_{u_2} \vec{u}_1 \end{aligned} \quad (3.10)$$

with the reduced matrices

$$\tilde{A}_{u_1} = \Psi_{u_1}^T A_{u_1} \Psi_{u_1}, \quad \tilde{A}_{u_2} = \Psi_{u_2}^T A_{u_2} \Psi_{u_2}, \quad \tilde{M}_{u_1} = \Psi_{u_1}^T M \Psi_{u_2}, \quad \tilde{M}_{u_2} = \Psi_{u_2}^T M \Psi_{u_1}.$$

To find the solutions \vec{u}_1 and \vec{u}_2 of the reduced system, the system of k -dimensional ODEs (3.10) is solved with the same time integrator of the full system. Then, the approximate solutions \vec{u}_1 and \vec{u}_2 of the full system (3.1) can be calculated via the relation (3.8).

Here, the choice of the dimension of the ROM or POD modes k is important, since there is no a-priori estimates for the choice of k . The number of the POD basis functions are determined according to the following relative information content [82]:

$$\epsilon(k) = \frac{\sum_{i=1}^k \sigma_i^2}{\sum_{i=1}^K \sigma_i^2}, \quad (3.11)$$

which represents the energy captured by the first k POD modes over all K POD modes, and σ_i is the corresponding singular value of i -th mode. The choice of this indicator (3.11) is supported by the following fact:

Theorem 3.1. *The error in (3.4) is given with $\lambda_i = \sigma_i^2$,*

$$\frac{1}{J} \sum_{j=1}^J \left\| w^j - \sum_{i=1}^k (w^j, \psi_{w,i})_{L^2(\Omega)} \psi_{w,i} \right\|_{L^2(\Omega)}^2 = \sum_{i=k+1}^K \lambda_i. \quad (3.12)$$

For a detailed proof of this theorem, see, e.g., [96].

Algorithm 1 The POD Algorithm

Input: Snapshots $\{\vec{u}_1^1, \dots, \vec{u}_1^J\}$, $\{\vec{u}_2^1, \dots, \vec{u}_2^J\}$.

Output: POD basis functions Ψ_{u_1} , Ψ_{u_2} .

- 1: Construct the snapshot matrices

$$U_1 = [\vec{u}_1^1, \dots, \vec{u}_1^J], \quad U_2 = [\vec{u}_2^1, \dots, \vec{u}_2^J].$$

- 2: Compute the mean of the columns

$$\bar{u}_1 = \frac{1}{J} \sum_{i=1}^J \vec{u}_1^i, \quad \bar{u}_2 = \frac{1}{J} \sum_{i=1}^J \vec{u}_2^i.$$

- 3: Construct new snapshot matrices

$$U'_1 = [\vec{u}_1^1 - \bar{u}_1, \dots, \vec{u}_1^J - \bar{u}_1], \quad U'_2 = [\vec{u}_2^1 - \bar{u}_2, \dots, \vec{u}_2^J - \bar{u}_2].$$

- 4: Calculate the empirical correlation matrix

$$\hat{U}'_1 = \frac{1}{J} U'_1 U'^T_1, \quad \hat{U}'_2 = \frac{1}{J} U'_2 U'^T_2,$$

where U'^T_1 , U'^T_2 denote the transposes of the corresponding mean-subtracted snapshot matrices U'_1 , U'_2 .

- 5: Calculate the singular values of the matrices \hat{U}'_1 , \hat{U}'_2 :

$$\hat{U}'_1 \psi_{u_1} = \lambda_i \psi_{u_1}, \quad \hat{U}'_2 \psi_{u_2} = \lambda_i \psi_{u_2} \text{ for } i = 1, \dots, K,$$

where K denotes total number of eigenvalues.

- 6: Find the number of POD basis functions k according to the energy criteria (3.11) capturing at least 99.99% of the energy.
- 7: Choose the first k eigenvectors $\{\psi_{u_1,1}, \dots, \psi_{u_1,k}\}$ and $\{\psi_{u_2,1}, \dots, \psi_{u_2,k}\}$ yielding the following POD basis functions of order k

$$\Psi_{u_1} = \frac{1}{\sqrt{\lambda_i}} \hat{U}'_1 \psi_{u_1}, \quad \Psi_{u_2} = \frac{1}{\sqrt{\lambda_i}} \hat{U}'_2 \psi_{u_2} \text{ for } i = 1, \dots, k.$$

- 8: Project the system onto the reduced spaces.
-

Remark 3.1. Here we use a version of the POD constructed from mean-subtracted snapshots [18] according to the Algorithm 1:

$$U'_1 = [\vec{u}_1^1 - \bar{u}_1, \dots, \vec{u}_1^J - \bar{u}_1], \quad U'_2 = [\vec{u}_2^1 - \bar{u}_2, \dots, \vec{u}_2^J - \bar{u}_2]$$

$$\text{with } \bar{u}_1 = \frac{1}{J} \sum_{i=0}^J \vec{u}_1^i, \quad \bar{u}_2 = \frac{1}{J} \sum_{i=0}^J \vec{u}_2^i.$$

The reduced system is written as

$$\begin{aligned} \tilde{u}_1 &= \tilde{A}_{u_1}(\tilde{u}_1 + \bar{u}_1) + \Psi_{u_1}^T F(\Psi_{u_1} \tilde{u}_1 + \bar{u}_1) - \tilde{M}_{u_1}(\tilde{u}_2 + \bar{u}_2) + \Psi_{u_1}^T K, \\ \tilde{u}_2 &= \tilde{A}_{u_2}(\tilde{u}_2 + \bar{u}_2) - \tilde{u}_2 + \tilde{M}_{u_2}(\tilde{u}_1 + \bar{u}_1) \end{aligned} \quad (3.13)$$

with the reduced matrices

$$\tilde{A}_{u_1} = \Psi_{u_1}^T A_{u_1} \Psi_{u_1}, \quad \tilde{A}_{u_2} = \Psi_{u_2}^T A_{u_2} \Psi_{u_2}, \quad \tilde{M}_{u_1} = \Psi_{u_1}^T M \Psi_{u_2}, \quad \tilde{M}_{u_2} = \Psi_{u_2}^T M \Psi_{u_1}.$$

There is no a significant difference between the reduced form of our system (3.1) from the reduced form using classical POD (3.10), whereas both (3.10)-(3.13) satisfy the same boundary conditions [18].

3.2 Discrete Empirical Interpolation

The computation of the nonlinear term $F(\Psi_{u_1} \tilde{u}_1)$ of the reduced system still depends on the dimension of the full order model N :

$$N(\tilde{u}_1) = \Psi_{u_1}^T F(\Psi_{u_1} \tilde{u}_1).$$

Further, the reduced system (3.10) has to be solved by the Newton method

$$\frac{\partial}{\partial \tilde{u}_1} N(\tilde{u}_1) = \Psi_{u_1}^T J_F \Psi_{u_1},$$

which requires the computation of the Jacobian matrix $J_F \in \mathbb{R}^{N \times N}$

$$(J_F)_{i,j} = \frac{\partial}{\partial u_{1j}} F_i(\vec{u}_1), \quad i, j = 1, 2, \dots, N,$$

where u_{1j} 's are the components of the solution \vec{u}_1 of the full system (3.1). To reduce the computational complexity of the nonlinear term, we apply DEIM [28] to the nonlinear term $F(\Psi_{u_1} \tilde{u}_1)$, which approximates the nonlinear function $F(\Psi_{u_1} \tilde{u}_1)$ by projecting it onto a subspace of the space generated by the nonlinear functions and spanned by a basis of dimension $m \ll N$. Let

$$\mathcal{F} = [F_1, F_2, \dots, F_J] \in \mathbb{R}^{N \times J}$$

denotes the snapshot matrix of the nonlinear functions at each time instances t_1, \dots, t_J computed during the solution of the full system (3.1) with $F_i = F(\Psi_{u_1} \vec{u}(t_i))$, $i =$

1, 2, ..., J. Through the SVD of the matrix \mathcal{F} , we can find the $m \ll N$ orthogonal POD basis functions $\{W_i\}_{i=1}^m$ spanning the m -dimensional subspace of the space spanned by \mathcal{F} . Then, for $W = [W_1, \dots, W_m] \in \mathbb{R}^{N \times m}$, we use the ansatz

$$F(\Psi_{u_1} \tilde{u}_1) \approx Ws(t) \quad (3.14)$$

with the corresponding coefficient vector $s(t)$. Note that the system above is overdetermined. Thus, to compute the coefficient vector $s(t)$, we take m distinguished rows from the system $Ws(t)$ through the projection using a permutation matrix $P = [e_{p_1}, \dots, e_{p_m}] \in \mathbb{R}^{N \times m}$ with e_{p_i} is the i -th column of the identity matrix $I \in \mathbb{R}^{N \times N}$, and it is computed using the DEIM algorithm [28] in Algorithm 2 so that $P^T W$ is non-singular.

Algorithm 2 DEIM Algorithm

Input: POD basis functions $\{W_i\}_{i=1}^m$.

Output: Index vector $\mathbf{p} = [p_1, \dots, p_m]^T$

Permutation Matrix $P = [e_{p_1}, \dots, e_{p_m}]$.

- 1: $[|\rho|, \mathbf{p}_1] = \max\{|W_1|\}$.
 - 2: $W = [W_1], P = [e_{p_1}], \mathbf{p} = [\mathbf{p}_1]$.
 - 3: for $l = 2$ to m do
 - 4: Solve $(P^T W)s = P^T W_l$ for s .
 - 5: $r = W_l - Ws$.
 - 6: $[|\rho|, \mathbf{p}_l] = \max\{|r|\}$.
 - 7: $W \leftarrow [W \ W_l], P \leftarrow [P \ e_{p_l}], \mathbf{p} \leftarrow \begin{bmatrix} \mathbf{p} \\ \mathbf{p}_l \end{bmatrix}$.
 - 8: end for
-

The projection of (3.14) leads to the system

$$\begin{aligned} P^T F(\Psi_u \tilde{u}(t)) &= (P^T W)s(t) \\ \implies s(t) &= (P^T W)^{-1} P^T F(\Psi_u \tilde{u}(t)). \end{aligned} \quad (3.15)$$

Then, using (3.14) and (3.15), the nonlinear term can be approximated as

$$N(\tilde{\mathbf{u}}(t)) \approx \tilde{N}(\tilde{\mathbf{u}}(t)) = Q\tilde{F}, \quad (3.16)$$

where the matrix $Q = \Psi_u^T W (P^T W)^{-1} \in \mathbb{R}^{k \times m}$ can be precomputed only once, and $\tilde{F} = P^T F(\Psi_{u_1} \tilde{u}_1) \in \mathbb{R}^m$ is the m -dimensional nonlinear vector which can be computed in an efficient way.

For instance, if $m = 4$, $\varphi_2 = 6$ and $\varphi_3 = 2$, we obtain

$$P^T F = \begin{bmatrix} 0 & \dots & \dots & 1 & \dots & 0 \\ 0 & \dots & 1 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} F_1 \\ \vdots \\ F_N \end{bmatrix} = \begin{bmatrix} F_{\varphi_1} \\ F_6 \\ F_2 \\ F_{\varphi_4} \end{bmatrix},$$

which means that, instead of N entries, only $m \ll N$ entries are computed.

Now before giving a priori error bound for the DEIM approximation of the nonlinear vector $F(\Psi_{u_1} \tilde{u}_1)$, let us define the *interpolatory projector*.

Definition 3.2. Let $W = [W_1, \dots, W_m]$ and $P = [e_{\varphi_1}, \dots, e_{\varphi_m}]$ be $N \times m$ real matrices with linearly independent columns. Suppose $(P^T W)^{-1}$ exists, then

$$\mathbb{P} := W(P^T W)^{-1} P^T$$

is called an *interpolatory projector* onto the range of W .

Lemma 3.2. Let $W = [W_1, W_2, \dots, W_m]$ be a $N \times m$ real matrix whose columns are orthonormal DEIM basis vectors. The DEIM approximation $\mathbb{P}F$ satisfies

$$\|F - \mathbb{P}F\| \leq C\varepsilon^*, \quad (3.17)$$

where $C := \|(P^T W)^{-1}\|$ and $\varepsilon^* = \|(I - WW^T)F\|$.

Proof. Let $F^* := WW^T F$ be the optimal 2-norm approximation to F in $\text{Ran}(W)$ and $\|y\| = \min\{\|F - G\| : G \in \text{Ran}(W)\}$. So we can write F as

$$F = (F - F^*) + F^* = y + F^*. \quad (3.18)$$

In addition, the DEIM approximation can be written as

$$\mathbb{P}F = \mathbb{P}(y + F^*) = \mathbb{P}y + F^* \quad (3.19)$$

since F^* is already in the range of W , i.e. $\mathbb{P}F^* = F^*$.

Subtracting (3.19) from (3.18) and using sub-unitary of W and P lead

$$\begin{aligned} F - \mathbb{P}F &= y - \mathbb{P}y = (I - \mathbb{P})y \implies \|F - \mathbb{P}F\| \leq \|I - \mathbb{P}\| \|y\| \\ &= \|\mathbb{P}\| \|y\| \\ &= \|W(P^T W)^{-1} P^T\| \|y\| \\ &= \|(P^T W)^{-1}\| \|y\|, \end{aligned} \quad (3.20)$$

which gives the result. \square

In the case of the entries $\{F_i\}_{i=1}^N$, they are just integrals of the form

$$F_i(\vec{u}) = \int_{E_{t_i}} f(\vec{u}) \phi_i dx, \quad i = 1, 2, \dots, N, \quad (3.21)$$

where E_{t_i} is the triangular element on which the i -th basis function ϕ_i is defined. For instance, if linear basis functions are used ($n_k = 3$), then $\{\phi_i\}_{i=1}^3$ are defined on E_1 , whereas, $\{\phi_i\}_{i=13}^{15}$ are defined on E_5 . The DG requires only computation of the integrals on a single triangular element, which is not the case in continuous finite elements where all the interior degrees of freedoms are shared by usually 6 triangular elements (see Figure 3.1). The unassembled finite element approach is used in [5] so that each DEIM point is related to one element. This reduces the online computational cost but increases the number of snapshots and therefore the cost of the offline computation.

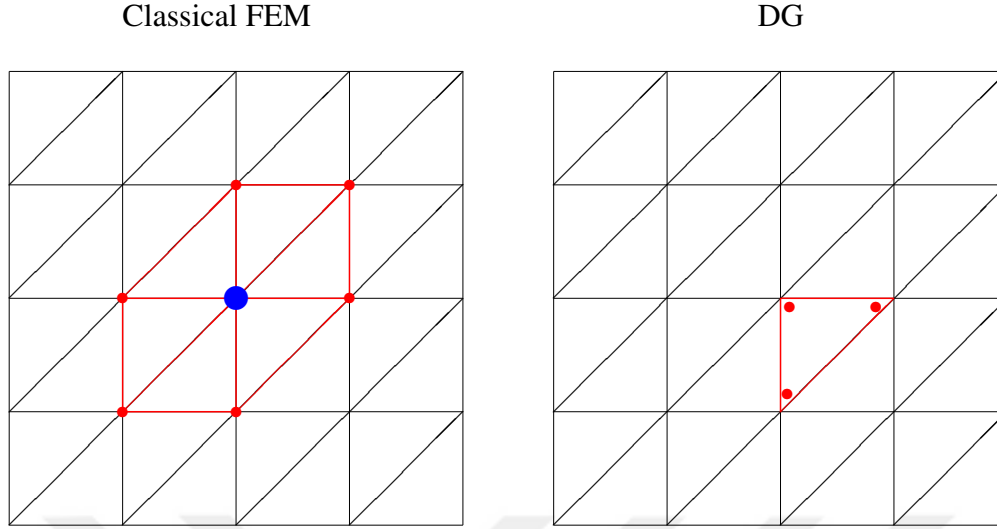


Figure 3.1: Connectivity of degrees of freedom for linear basis

Due to its' local nature, the DG discretization is automatically in the unassembled form and it does not require computation of additional snapshots (see Figure 3.1).

The integrals in (3.21) are computed by Gaussian quadrature of order 4, which are computed exactly for the cubic nonlinear terms. When n_q denotes the number of quadrature points in the reference triangle, then, through an affine map from the reference triangle to the physical triangles, once the values of basis functions at $n_e \times n_q$ quadrature points are evaluated in the very-beginning, the entries F_i 's can be computed in an efficient way since the triangular elements E_{t_i} 's on which the integral related to the entries are unique.

For the solution of the nonlinear system by Newton's method, the entries of the Jacobian $J_F \in \mathbb{R}^{N \times N}$ of the nonlinear terms are of the form $i, j = 1, 2, \dots, N$

$$(J_F)_{i,j} = \frac{\partial}{\partial u_{1j}} F_i(\vec{u}_1) = \int_{E_{t_i}} \frac{\partial f(\vec{u}_1)}{\partial u_{1j}} \phi_i dx = \int_{E_{t_i}} f'(\vec{u}_1) \phi_j \phi_i dx.$$

Because the DG basis functions are defined only a single element and they vanish outside that element, the integral terms of the Jacobian matrix $(J_F)_{i,j}$ vanish on the triangular elements E_{t_i} where the basis function ϕ_j is not defined. Unlike the continuous finite elements where the Jacobian matrix contains overlapping blocks, Jacobian matrix in DG appears in block diagonal form, and has the form

$$J_F = \begin{bmatrix} A_{11} & 0 & \dots & 0 \\ 0 & A_{22} & 0 & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & A_{n_e, n_e} \end{bmatrix}$$

whose of blocks $A_{ii} \in \mathbb{R}^{n_k \times n_k}$, $i = 1, 2, \dots, n_e$, corresponds to a single triangular element E_i (remember that $N := n_k \times n_e$). As a consequent, every component of

an arbitrary block A_{ii} of the Jacobian matrix J_F only depends on the n_k components $u_{\mu+1}, u_{\mu+2}, \dots, u_{\mu+n_k}$ for $\mu = (i-1)n_k$. The Jacobian matrix arising from the DEIM has the form:

$$\frac{\partial}{\partial \tilde{u}_1} N(\tilde{u}_1) = \Psi_{u_1}^T J_F \Psi_{u_1}, \quad \frac{\partial}{\partial \tilde{u}_1} \tilde{N}(\tilde{u}_1) = Q(P^T J_F) \Psi_{u_1}.$$

We note that $(P^T J_F) \in \mathbb{R}^{m \times N}$ is the matrix whose i -th row is the \wp_i -th row of the Jacobian J_F , $i = 1, 2, \dots, m$, and in each row of the Jacobian there are only n_k non-zero terms because of the local structure of the DG. For instance, again, if $m = 4$, $\wp_2 = 6$, $\wp_3 = 2$ and linear DG basis functions are used ($n_k = 3$), we get

$$P^T J_F = \begin{bmatrix} \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \frac{\partial F_6}{\partial u_4} & \frac{\partial F_6}{\partial u_5} & \frac{\partial F_6}{\partial u_6} & 0 & \dots & 0 \\ \frac{\partial F_2}{\partial u_1} & \frac{\partial F_2}{\partial u_2} & \frac{\partial F_2}{\partial u_3} & 0 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}.$$

Hence, only $m \times n_k$ entries are needed to compute for the term $P^T J_F$, whereas without DEIM, it requires the computation of $n_e \times n_k^2$ entries for the term J_F .

3.3 Dynamic Mode Decomposition

Dynamic mode decomposition (DMD) was first proposed in fluid dynamics by Schmid [80] spatio-temporal decomposition of the complex fluid flow. Afterward the DMD was shown as an approximation of the infinite dimensional Koopman operator [52], see, e.g., [63, 76]. It is an equation-free, data-driven and snapshot based reduced order method. Without explicit knowledge of the dynamical system, the DMD algorithm determines eigenvalues, eigenmodes, and spatial structures for each mode and combines each mode with its Fourier series in time. Although the Koopman operator is an infinite dimensional linear operator, the DMD approximates this operator by a finite dimensional linear operators, say \mathcal{A} [31, 80].

If we consider the following dynamical system

$$\bar{u}^{n+1} = g(\bar{u}^n)$$

on a manifold \mathcal{M} , then the Koopman operator \mathcal{A} maps any observable function $f : \mathcal{M} \rightarrow \mathcal{C}$ such that

$$\mathcal{A}f(\bar{u}^n) = f(g(\bar{u}^n)).$$

For DMD, this observable function is chosen as linear: $f(\bar{u}^n) = \bar{u}^n$. Then, the DMD algorithm determines eigen-elements of the Koopman operator directly from the snapshots.

Afterwards, the expansion of the function f can be written as

$$f(\bar{u}^n) = \sum_{j=1}^{\infty} \lambda_j^n \alpha_j(w_0) v_j,$$

where $\lambda_j = e^{\sigma_j + iw_j}$ denotes the complex-valued Ritz eigenvalues of the modal decomposition $\mathcal{A}v_j = \lambda_j v_j$ with the corresponding growth rate σ_j and the frequency w_j , $v_j \in \mathbb{C}$ denotes the dynamic modes as Koopman functions, and $\alpha_j \in \mathbb{C}$ denotes the amplitudes of the Koopman modes [18].

We consider the following two sets of shifted snapshot matrices for

$$\begin{aligned} U_1(J-1) &= [\vec{u}_1^1, \dots, \vec{u}_1^{J-1}], & U_1(J) &= [\vec{u}_1^2, \dots, \vec{u}_1^J], \\ U_2(J-1) &= [\vec{u}_2^1, \dots, \vec{u}_2^{J-1}], & U_2(J) &= [\vec{u}_2^2, \dots, \vec{u}_2^J], \end{aligned} \quad (3.22)$$

where $U_1(J-1), U_1(J), U_2(J-1), U_2(J) \in \mathbb{R}^{N \times J-1}$, and each member of the snapshot matrices are the corresponding coefficient vectors of the discrete solutions $\{u_1^i\}_{i=1}^J$ and $\{u_2^i\}_{i=1}^J$, respectively, of the FOM (3.1) at $J-1$ equally spaced time instances.

There are two main assumptions for the construction of the DMD algorithms:

- There exists two linear operator \mathcal{A}_1 and \mathcal{A}_2 such that

$$\begin{aligned} \vec{u}_1^{t+1} &= \mathcal{A}_1 \vec{u}_1^t, \\ \vec{u}_2^{t+1} &= \mathcal{A}_2 \vec{u}_2^t, \end{aligned} \quad (3.23)$$

for $t \in \{1, \dots, J-1\}$.

- The final snapshots can be written as a linear combination of the previous snapshots as

$$\begin{aligned} \vec{u}_1^J &= c_1 \vec{u}_1^1 + \dots + c_{m-1} \vec{u}_1^{J-1} + \vec{r}_1, \\ \vec{u}_2^J &= d_1 \vec{u}_2^1 + \dots + d_{m-1} \vec{u}_2^{J-1} + \vec{r}_2, \end{aligned} \quad (3.24)$$

where \vec{r}_1, \vec{r}_2 denote residual vectors.

The goal in the DMD algorithm is to determine the eigen-elements of $\mathcal{A}_1, \mathcal{A}_2$. Since the dimension of the problem is very large, the matrices $\mathcal{A}_1, \mathcal{A}_2$ are generally ill-conditioned. Instead, the DMD provides a lower rank approximations in terms of the POD-projected matrix $\tilde{\mathcal{A}}_1, \tilde{\mathcal{A}}_2$.

From the first assumption, the collected snapshot sets correspond to the $(J-1)$ -th Krylov subspace generated by the Koopman operators $\mathcal{A}_1, \mathcal{A}_2$ from \vec{u}_1^1 and \vec{u}_2^1 , respectively. So, using the first assumption (3.23) leads the following relation between the snapshot matrices

$$\begin{aligned} U_1(J-1) &= \mathcal{A}_1 U_1(J), \\ U_2(J-1) &= \mathcal{A}_2 U_2(J), \end{aligned} \quad (3.25)$$

and using the second assumption (3.24) leads

$$\begin{aligned} \mathcal{A}_1 U_1(J-1) &= U_1(J-1) \tilde{\mathcal{A}}_1 + \vec{r}_1^T e_{J-1}^T, \\ \mathcal{A}_2 U_2(J-1) &= U_2(J-1) \tilde{\mathcal{A}}_2 + \vec{r}_2^T e_{J-1}^T, \end{aligned} \quad (3.26)$$

where e^T denotes i -th Euclidean unitary vector of length $J - 1$ and $\tilde{\mathcal{A}}_1, \tilde{\mathcal{A}}_2$ are the companion matrices defined as

$$\tilde{\mathcal{A}}_1 = \begin{bmatrix} 0 & \cdots & 0 & c_1 \\ 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & c_{m-1} \end{bmatrix}, \quad \tilde{\mathcal{A}}_2 = \begin{bmatrix} 0 & \cdots & 0 & d_1 \\ 1 & \cdots & 0 & d_2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & d_{m-1} \end{bmatrix}.$$

Therefore, it follows from (3.26) that the matrices $\mathcal{A}_1, \mathcal{A}_2$ are similar to the companion matrices $\tilde{\mathcal{A}}_1, \tilde{\mathcal{A}}_2$, respectively. Also note that, (3.26) implies the matrices $\tilde{\mathcal{A}}_1, \tilde{\mathcal{A}}_2$ are low dimensional representations of $\mathcal{A}_1, \mathcal{A}_2$ in the subspaces spanned by the POD modes of snapshot matrices $U_1(J - 1), U_2(J - 1)$. It follows that eigen-elements of $\mathcal{A}_1, \mathcal{A}_2$ can be approximated by the low dimensional companion matrices $\tilde{\mathcal{A}}_1, \tilde{\mathcal{A}}_2$, respectively.

The matrices $\tilde{\mathcal{A}}_1, \tilde{\mathcal{A}}_2$ can be determined by the following minimization problems with respect to the Frobenious norm:

$$\min \left\| U_1(J) - U_1(J - 1)\tilde{\mathcal{A}}_1 \right\|_F^2, \quad \min \left\| U_2(J) - U_2(J - 1)\tilde{\mathcal{A}}_2 \right\|_F^2. \quad (3.27)$$

Frobenious norm of a matrix A is defined as

$$\|A\|_F^2 = \text{trace}(A^*A) = \text{trace}(AA^*),$$

where A^* denotes a complex conjugate transpose of A .

Solutions to these minimization problems (3.27) are given in [90] by

$$\tilde{\mathcal{A}}_1 = U_1(J - 1)^\dagger U_1(J), \quad \tilde{\mathcal{A}}_2 = U_2(J - 1)^\dagger U_2(J), \quad (3.28)$$

where \dagger denotes the Moore-Penrose pseudoinverse.

In the literature, there are different DMD algorithms [31, 90]. Here, we consider the exact DMD algorithm by Tu et al. in [90]:

Algorithm 3 Exact DMD Algorithm

Input: Snapshots $U_1(J-1)$, $U_1(J)$, $U_2(J-1)$, $U_2(J)$.

Output: DMD modes $\Psi_{u_1}^{\text{DMD}}$, $\Psi_{u_2}^{\text{DMD}}$.

- 1: Compute SVDs of $U_1(J-1)$, $U_2(J-1)$:

$$U_1(J-1) = U_1 \Sigma_1 V_1^*, \quad U_2(J-1) = U_2 \Sigma_2 V_2^*.$$

- 2: Substitute into $\tilde{\mathcal{A}}_1$, $\tilde{\mathcal{A}}_2$:

$$\tilde{\mathcal{A}}_1 = U_1^* U_1(J) V_1 \Sigma_1^{-1}, \quad \tilde{\mathcal{A}}_2 = U_2^* U_2(J) V_2 \Sigma_2^{-1}.$$

- 3: Find eigen-elements of $\tilde{\mathcal{A}}_1$, $\tilde{\mathcal{A}}_2$:

$$\tilde{\mathcal{A}}_1 W_1 = \lambda_1 W_1, \quad \tilde{\mathcal{A}}_2 W_2 = \lambda_2 W_2.$$

- 4: Get DMD modes $\Psi_{u_1}^{\text{DMD}}$, $\Psi_{u_2}^{\text{DMD}}$:

$$\Psi_{u_1}^{\text{DMD}} = U_1(J) V_1 \Sigma_1^{-1} W_1, \quad \Psi_{u_2}^{\text{DMD}} = U_2(J) V_2 \Sigma_2^{-1} W_2.$$

3.3.1 DMD Modes with Optimal Amplitudes

In spite of the fact that POD modes are ranked according to their energy contribution, for DMD there is not such criteria for the selection of the DMD modes. Depending on the context of the problem, different criteria are developed, for instance based on their frequency/growth rates or amplitudes. However all these have some disadvantages. For example, the criteria based on frequency/growth rates relies on a-priori physical knowledge, and there can be some modes with high amplitudes but fast damped when the criteria based on amplitudes are used [87]. Moreover, since the DMD modes are not orthogonal, it makes this situation more complicated. Therefore, several algorithms are developed for the optimal selection of the DMD modes, like gradient-based algorithm, optimal amplitudes weighted its temporal coefficient [87], construction of the mode shapes forming the low rank basis [99], or spatial structures using combinatorial search [31].

Here, we use the optimal modes selection according to the sparsity-promoting framework [48]. We consider the following reduced dynamical system:

$$\begin{aligned} \vec{u}_1^{t+1} &= \tilde{\mathcal{A}}_1 \vec{u}_1^t, \\ \vec{u}_2^{t+1} &= \tilde{\mathcal{A}}_2 \vec{u}_2^t, \end{aligned} \tag{3.29}$$

for $t \in \{1, \dots, J-1\}$ and reduced matrices

$$\tilde{\mathcal{A}}_1 = U_1^* U_1(J) V_1 \Sigma_1^{-1}, \quad \tilde{\mathcal{A}}_2 = U_2^* U_2(J) V_2 \Sigma_2^{-1},$$

obtained from Algorithm 3.

The reduced solution of (3.29) can be written as

$$\begin{aligned}\vec{u}_1 &= \sum_{i=1}^r = \alpha_1^i \lambda_i(u_1)^t \phi_{u_1}^i, \\ \vec{u}_2 &= \sum_{i=1}^r = \alpha_2^i \lambda_i(u_2)^t \phi_{u_2}^i,\end{aligned}\tag{3.30}$$

where

$$\Phi_{u_1} = [\phi_{u_1}^1, \dots, \phi_{u_1}^m], \quad \Phi_{u_2} = [\phi_{u_2}^1, \dots, \phi_{u_2}^m]$$

are the matrices of the DMD modes, $\lambda_i(u_1)$, $\lambda_i(u_2)$ are the corresponding complex eigenvalues of $\phi_{u_1}^i$, $\phi_{u_2}^i$, respectively, and $t \in \{0, \dots, m-1\}$. The unknown amplitudes $\alpha_1 = [\alpha_1^1, \dots, \alpha_1^r]$, $\alpha_2 = [\alpha_2^1, \dots, \alpha_2^r]$ are found by solving the following minimization problems

$$\begin{aligned}\min & \|U_1(J-1) - \Phi_{u_1} D(\alpha_1) V_{and}^1\|_F^2, \\ \min & \|U_2(J-1) - \Phi_{u_2} D(\alpha_2) V_{and}^2\|_F^2,\end{aligned}\tag{3.31}$$

where

$$D(\alpha_1) = \text{diag}(\alpha_1^1, \dots, \alpha_1^r), \quad D(\alpha_2) = \text{diag}(\alpha_2^1, \dots, \alpha_2^r),$$

and

$$V_{and}^1 = (\lambda_1(u_1), \dots, \lambda_r(u_1)), \quad V_{and}^2 = (\lambda_1(u_2), \dots, \lambda_r(u_2))$$

are Vandermonde matrices.

Finally the approximate DMD solutions are given as

$$\begin{aligned}\vec{u}_1 &= \sum_{i=1}^r = \alpha_1^i(0) \phi_{u_1}^i \exp(w_1^i), \\ \vec{u}_2 &= \sum_{i=1}^r = \alpha_2^i(0) \phi_{u_2}^i \exp(w_2^i)\end{aligned}\tag{3.32}$$

with $w_1^i = \log(\lambda_i(u_1))/\Delta t$, $w_2^i = \log(\lambda_i(u_2))/\Delta t$.

For a detailed information, we refer [48] to the reader.

3.3.2 POD with DMD

In this part, we apply DMD for nonlinear reaction part of the diffusive FHNE (3.2) as an alternative method to DEIM, described in Section 3.2. So, we try to approximate the nonlinear term $f(\vec{u}_1)$ of the dynamical system (3.1) by means of the DMD approximation [3]. We construct the following snapshot matrix for nonlinear vector of the SIPG semi-discretized FHNE (3.1):

$$\mathcal{F} = [F_1, F_2, \dots, F_J] \in \mathbb{R}^{N \times J}.$$

We apply the DMD Algorithm 3 to nonlinear snapshots \mathcal{F}

$$\vec{f}(\vec{u}_1) = \sum_{i=1}^r \alpha_1^i(0) \phi_{u_1}^i \exp(w_1^i), \quad (3.33)$$

where $\phi_{u_1}^i$ denotes DMD basis functions of rank r for the nonlinear vector $\mathbf{f}(\vec{u}_1)$. In a closed form, the nonlinearity can be approximated as

$$\vec{f}(\vec{u}_1) \approx \Psi^{DMD} \text{diag}(\exp(w^{DMD}t))b, \quad (3.34)$$

where $b = (\Psi^{DMD})^\dagger(\vec{u}_1)_1$ denotes initial amplitude of each mode, determined by projecting back to the initial data. Hence, the nonlinear term is approximated by its DMD representation, which implies that no further evaluation of the nonlinear term is needed. When we plug this approximation (3.34) into the POD system (3.10), we obtain the following ROM for POD-DMD

$$\begin{aligned} \tilde{u}_1 &= \tilde{A}_{u_1} \tilde{u}_1 + \Psi_{u_1}^T \Psi^{DMD} \text{diag}(\exp(w^{DMD}t))b - \tilde{M}_{u_1} \tilde{u}_2 + \Psi_{u_1}^T K, \\ \tilde{u}_2 &= \tilde{A}_{u_2} \tilde{u}_2 - \tilde{u}_2 + \tilde{M}_{u_2} \tilde{u}_1, \\ N(\tilde{u}_1) &= \underbrace{\Psi_{u_1}^T \Psi^{DMD}}_{\in \mathbb{R}^{k \times r}} \underbrace{\text{diag}(\exp(w^{DMD}t))b}_{\in \mathbb{R}^r}. \end{aligned} \quad (3.35)$$

The most powerful feature that distinguishes the POD-DMD from the POD-DEIM model is that the ROM obtained from POD-DMD (3.35) is linear which makes POD-DMD significantly faster than the other ROM methods. However, the DMD modes obtained from DMD Algorithm 3 are not orthogonal.

The POD-DMD algorithm is given in [3]

Algorithm 4 POD-DMD Algorithm

Input: k (number of POD Basis), r (number of DMD basis).

Output: \vec{u}_1, \vec{u}_2 .

- 1: Construct the snapshot matrices U_1, U_2 for both state variables \vec{u}_1, \vec{u}_2 .
 - 2: Compute POD basis functions $\{\psi_{u_1,i}, \psi_{u_2,i}\}_{i=1}^k$.
 - 3: Construct the snapshot matrix \mathcal{F} for the nonlinear term.
 - 4: Compute DMD modes Ψ^{DMD} according to the Algorithm 3.
 - 5: Calculate the DMD approximation for the nonlinear term (3.34).
 - 6: Solve the ROM (3.35).
-

3.4 Randomized Singular Value Decomposition

Accurate and fast computation of the randomized (rSVD) is described [62, 4].

Let $A \in \mathbb{R}^{m \times n}$ be a given arbitrary matrix. We construct a random low-dimensional basis $Q^{m \times k}$ for the column space of A satisfying the following relation:

$$A \approx QB,$$

where $B \in \mathbb{R}^{k \times n}$. We use here a random projection matrix, a Gaussian random matrix $\Omega^{m \times k}$, whose entries is formed using Gaussian distribution. Then we construct the sample matrix Y for the column space of A as

$$Y = A\Omega.$$

The orthonormal basis for the column space of A , are computed by the QR -decomposition of Y as $Y = QR$

$$A \approx QQ^*A.$$

As a result, B is defined as the projection of the matrix A onto the low-dimensional subspace Q

$$B = Q^*A \in \mathbb{R}^{k \times n}.$$

Restricting the column space of A to small matrix B , we apply SVD to the relatively small matrix B to obtain right singular vectors and singular elements of A .

The rSVD algorithm is given as

Algorithm 5 rSVD Algorithm

Input: A matrix $A^{m \times n}$, target rank k , over-sampling parameter p .

Output: Singular vectors U .

- 1: Construct Gaussian random matrix Ω of dimension $m \times k$ as $\Omega = \text{randn}(m, k + p)$.
 - 2: Form the corresponding sample matrix: $Y = A\Omega$.
 - 3: Compute QR -decomposition of A : $A = QR$.
 - 4: Set the matrix B : $B = Q^*A$.
 - 5: Compute SVD of B : $B = \tilde{U}\Sigma V^*$.
 - 6: Obtain right singular vectors U : $U = Q\tilde{U}$.
-

The over-sampling parameter p is used to reduce the approximation error of the rSVD [43]. Instead of using k samples, extra p samples can be added ($k = k + p$) which improve the quality of the basis. In general, a small value for p , for instance $p = 5$, is sufficient.

3.5 Numerical Results

In this section, we present some numerical results supporting the ROM techniques described in the previous Sections 3.1-3.3: POD, POD-DEIM, DMD as an equation-free, and POD-DMD for both diffusive and convective FHNEs. As a space discretization, we use DG method, as a time discretization we use the average vector field method for the diffusive FHNE, implicit Euler for the convective FHNE.

Since there is no a-priori error estimates for the ROMs, we compare the ROM solutions with the FOM solutions in terms of the Frobenious error. We compare the performance of the ROM techniques in terms of CPU times, and speed-up of the ROM solutions with respect to the FOM solutions.

3.5.1 Diffusive FHNE

We consider the diffusive FHNE (3.2) with labyrinth-like patterns on $\Omega = [-1, 1]^2$ with zero flux boundary conditions and with the initial conditions as uniformly distributed random numbers between -1 and 1 . Space mesh size and time step size are taken as $\Delta x_1 = \Delta x_2 = 1/16$ and $\Delta t = 1/10$, respectively. According to the energy plot, we see that the steady state is reached at approximately $t = 100$. So, we can take $T = 100$ as a final time. We set the parameters $d_1 = 0.00028$, $d_2 = 0.005$ and $\kappa = 0$.

We present numerical results for FOM and ROMs computed by only POD, POD-DEIM, and POD-DMD. We use $k = 20$ POD basis functions, $m = 30$ DEIM basis functions, and $r = 70$ DMD basis functions. The snapshot matrices is formed by the solutions at 1001 time instances resulting in a matrix in $\mathbb{R}^{24576 \times 1001}$ which has the rank 150 and 144 for \vec{u}_1 and \vec{u}_2 , respectively. After the POD basis computation, we obtain a steep descent in the singular values of the system related to \vec{u}_1 , \vec{u}_2 , and the nonlinear term \vec{f} , shown in Figure 3.3. The average number of Newton iterations is 1 and tolerance number for the Newton iteration is $1e - 10$ for the computation of the FOMs and ROMs except POD-DMD on each time step.

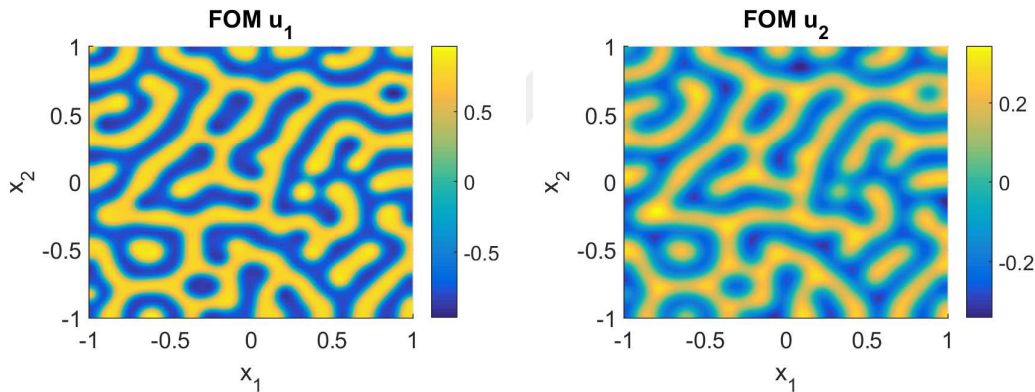


Figure 3.2: FOM solutions for the component u_1 (left), u_2 (right).

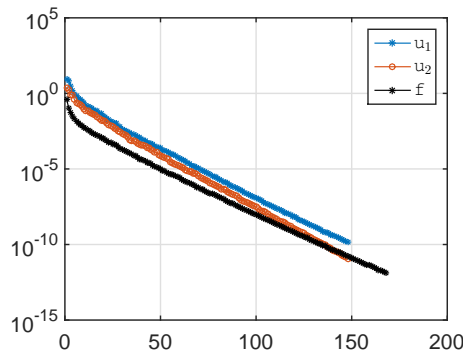


Figure 3.3: Decay of the singular values for u_1 , u_2 and the non-linearity f .

The computed labyrinth-like Turing pattern solutions of the FOM are given in Figure 3.2, the same with the Figure 2.6. The corresponding solution profiles of the ROMs with errors for the first state component in Figure 3.4 and for the second state component in Figure 3.5 are consistent with the solution profiles of the FOMs. As expected, POD error is smaller than the other ROM errors. Moreover, POD-DMD approximates to FOM solutions better than POD-DEIM.

The CPU time to compute the POD, DEIM and DMD basis functions, which are included in offline stage, are 5.30, 2.36, 337.70 seconds, respectively.

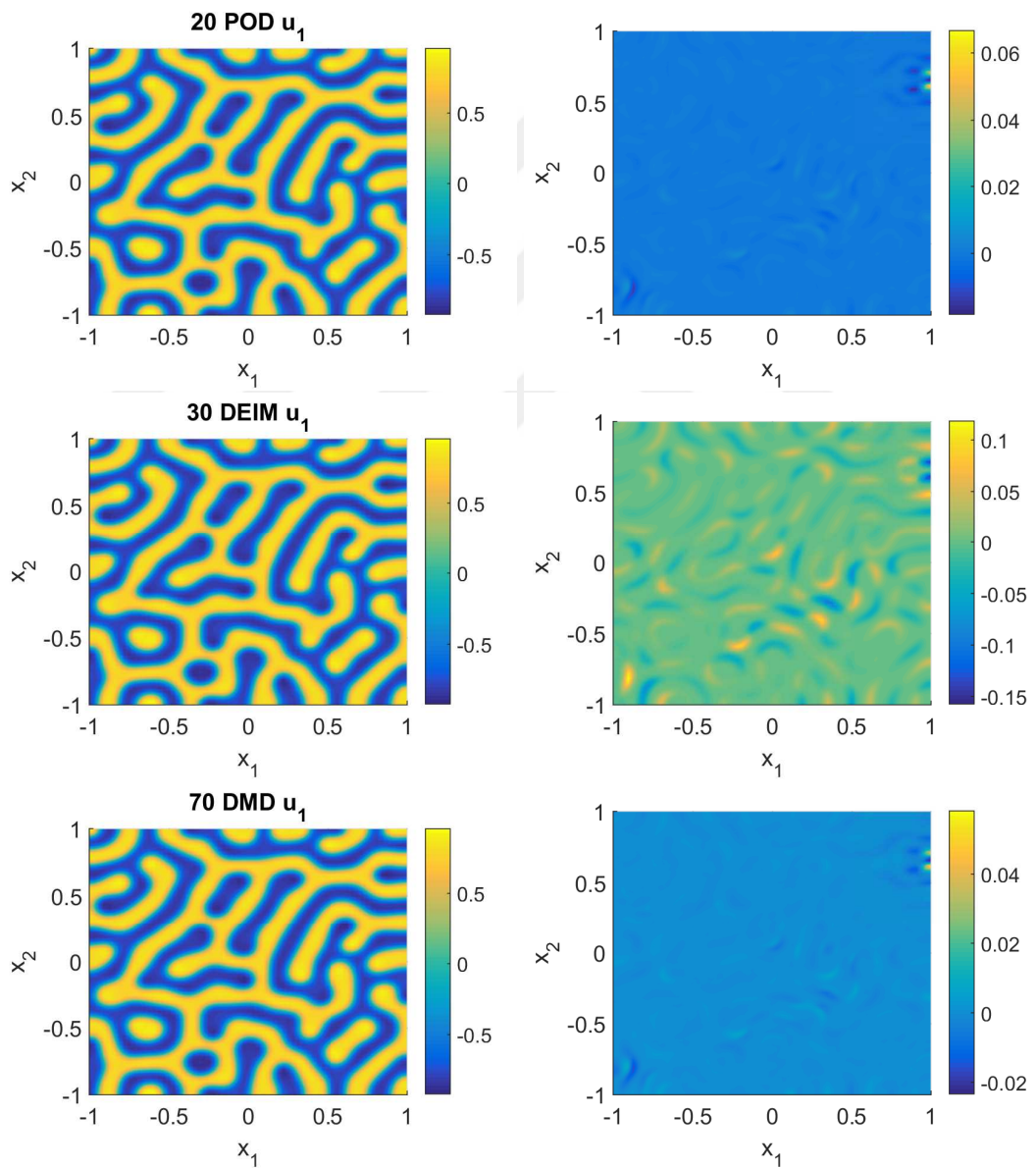


Figure 3.4: From top to bottom POD, DEIM, DMD solutions (left) and errors (right) for the component u_1 .

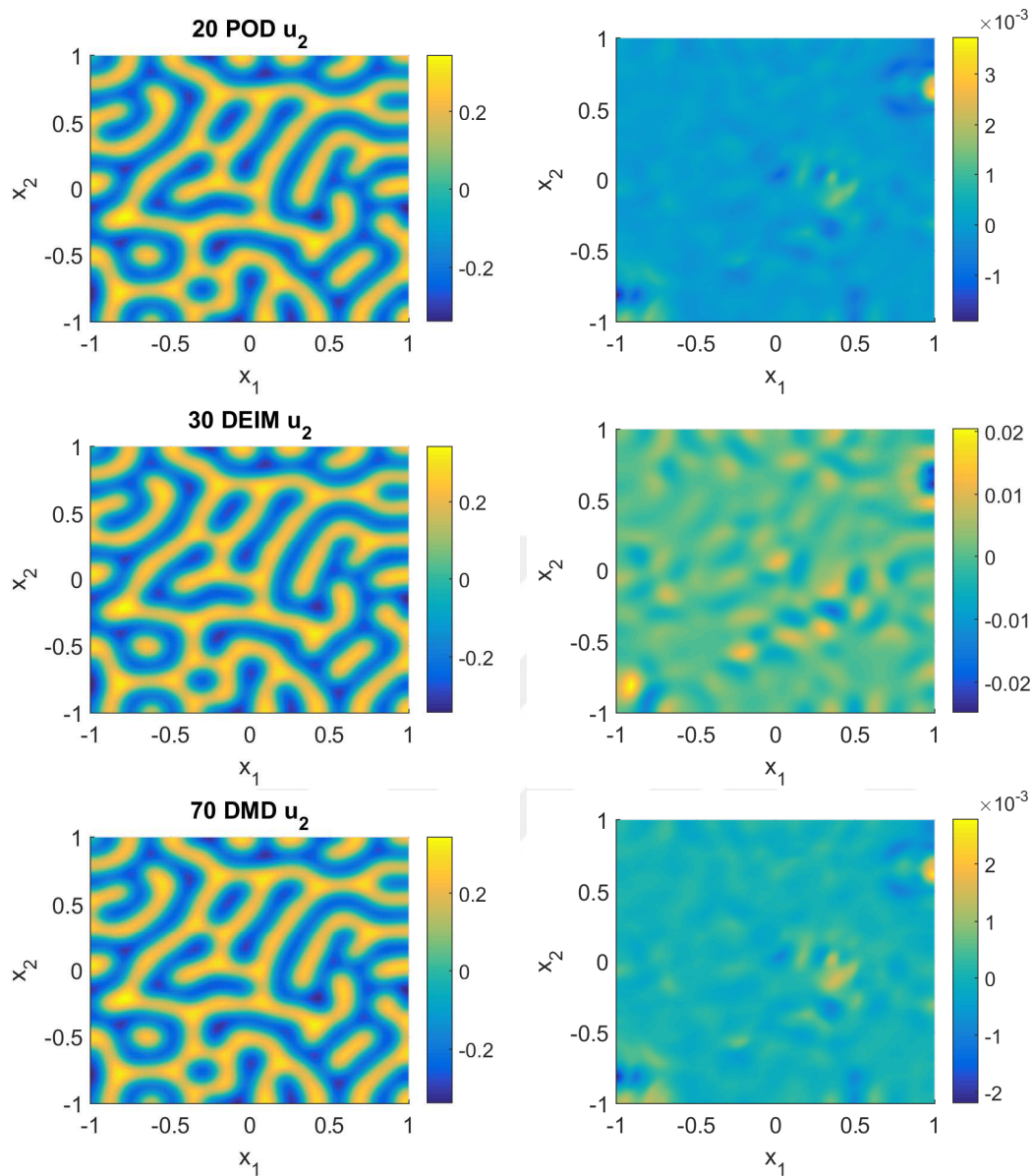


Figure 3.5: From top to bottom POD, DEIM, DMD solutions (left) and errors (right) for the component u_2 .

The corresponding energy plots for FOM and ROMs are shown in Figure 3.6. Besides, the energy errors are also shown in Figure 3.7. As it is seen, all ROM methods preserve the energy; however, the POD method captures the energy better than the other ROM methods as expected.

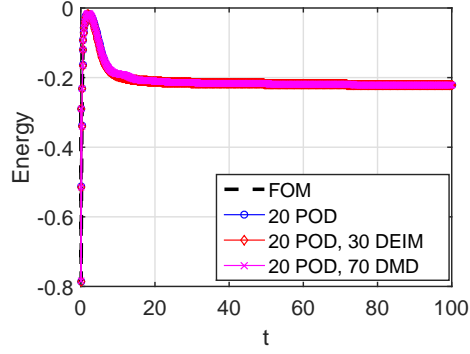


Figure 3.6: Energy plots.

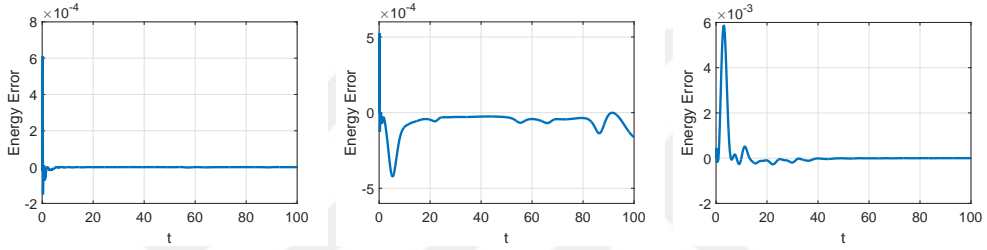


Figure 3.7: Energy errors.

In Table 3.1, we show the computed Frobenious errors between FOM and ROM solutions. POD errors are less than the other methods, which is acceptable since POD-DEIM and POD-DMD methods approximate the nonlinearity in the POD reduced form. Besides, for all ROM methods the errors for the first state variable is greater than the other component.

The speed-up's are calculated according to the following equations:

$$S_{POD} = \frac{\text{CPU time for FOM}}{\text{CPU time for POD}}, \quad S_{DEIM} = \frac{\text{CPU time for FOM}}{\text{CPU time for POD-DEIM}},$$

$$S_{DMD} = \frac{\text{CPU time for FOM}}{\text{CPU time for POD-DMD}},$$

where S_{POD} , S_{DEIM} , S_{DMD} are the speed-up factors of POD- POD-DEIM, and POD-DMD, respectively.

Table 3.1: FOM-ROM errors & speed-up with 20 POD, 30 DEIM, and 70 DMD basis functions.

	L_2 -Error u_1	L_2 -Error u_2	Energy Error	Speed-up
POD	1.81e-02	8.95e-03	2.60e-03	1.73
POD-DEIM	1.72e-01	4.49e-02	8.23e-03	14.11
POD-DMD	3.78e-01	1.00e-01	8.40e-03	173.10

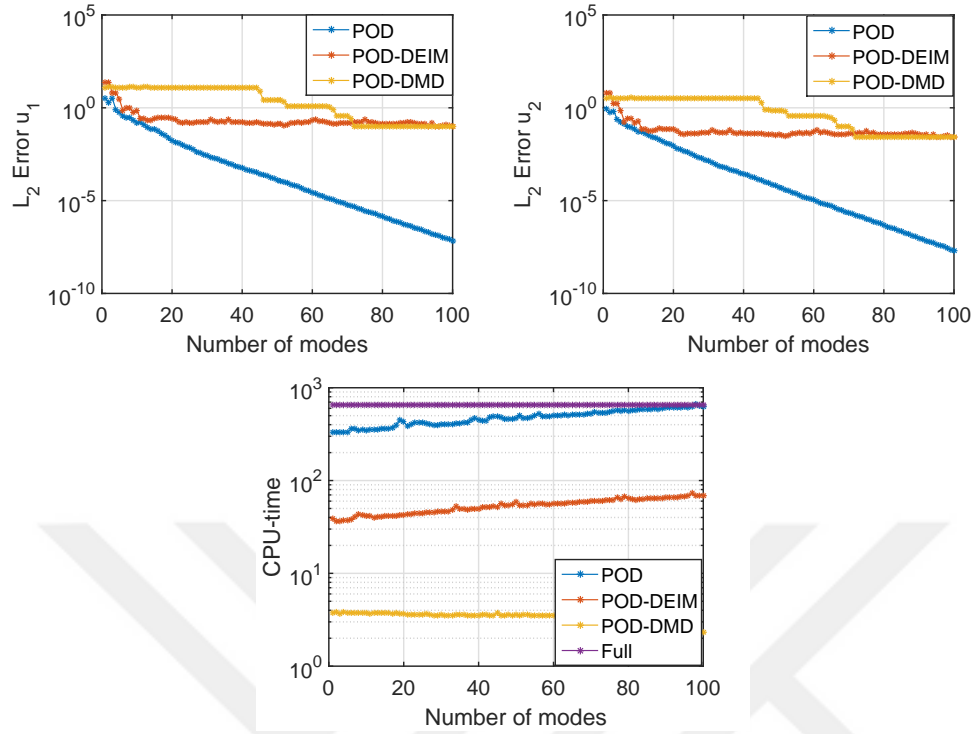


Figure 3.8: (Top) FOM-ROM $L^2(Q)$ errors for the states u_1 (left), u_2 (right) with increasing number of POD for POD solution, and increasing number of DEIM and DMD basis functions with 20 fixed POD basis functions for POD-DEIM and POD-DMD solutions; (bottom) corresponding CPU times.

Choosing the number of basis functions for POD, DEIM, and DMD is crucial for numerical simulations. In general, in order to improve the numerical results, or decrease the error between FOM and ROM solution, we have to increase the number of modes. However, what should be the relation between the number of POD, DEIM, and DMD basis functions is an important subject. Although there is no a-priori error estimates for ROMs, we focus on heuristic energy criteria for POD (3.11). So, when the basis functions contain at least %99.99 of the energy of the full system, we use that number as a number of POD basis functions. Although increasing the number of POD modes improves the result for POD solution, for POD-DEIM and POD-DMD it is not the case all the time. Hence, in order to find the relation between POD and the other basis functions we check two cases. The first one is after we fix the number of POD basis functions according to the energy criteria (3.11), we increase the number of DEIM modes and DMD modes for POD-DEIM and POD-DMD solutions, respectively to show the Frobenius error as it seen in Figure 3.8. On the other hand, we show the Frobenius errors with respect to the same increasing number of POD, DEIM and DMD basis functions in Figure 3.9.

Increasing the number of POD basis functions leads to more accurate solutions for both cases. The error for POD is monotonically decreasing. However, this situation is not that much good for POD-DEIM and POD-DMD solutions. In order to obtain

more accurate solutions, the number of DEIM and DMD modes must be more than POD modes. So, we have used different number of modes to obtain POD-DEIM and POD-DMD solutions in Figures 3.4-3.7. The errors in all ROM methods are decreased, when the number of basis functions are increased.

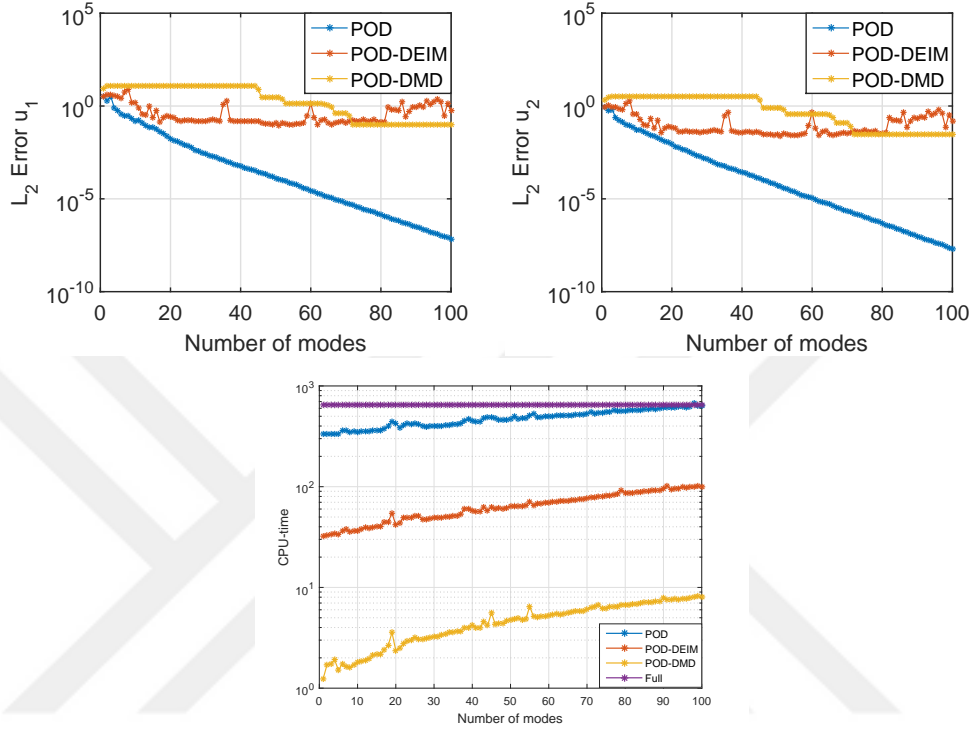


Figure 3.9: (Top) FOM-ROM $L^2(Q)$ errors for the states u_1 (left), u_2 (right) with the same increasing number of POD, DEIM and DMD basis functions; (bottom) corresponding CPU times.

On the other hand, we have to consider also CPU times to show the efficiency of ROM solutions. As it seen in Figures 3.8-3.9, all ROM methods we use requires less time than FOM except for the method that we use many modes in the POD method. As expected, POD-DMD provides a significant reduction since DMD approximates the nonlinear equation by a linear one. So, it is solved directly without using Newton's method.

3.5.2 Parametrized Model Order Reduction for Diffusive FHNE

In this section, we give the numerical results for the parametrized diffusive FHNE [51]. We consider the following diffusive parametrized FHNE [61]:

$$\begin{aligned} u_{1t} &= d_1 \Delta u_1 + f(u_1; \mu) - \alpha(u_2 - u_1), \\ u_{2t} &= d_2 \Delta u_2 - \beta(u_2 - u_1), \end{aligned} \tag{3.36}$$

where the parametrized bistable nonlinear term is defined as $f(u_1; \mu) = (u - \mu) - \beta(u^2 - 1)$ on $[-10, 10]^2 \times [0, 1000]$ with uniformly distributed initial conditions between -1 and 1 . The other parameters $d_1 = 0.04$, $d_2 = 1$, $\alpha = 0.3$, $\beta = 1$ are fixed as in [61]. We use linear DG polynomials ($n_k = 3$), and as the discrete mesh, we form the partition of $[-10, 10]^2$, by 5 times uniform refinement with 2048 triangular elements leading to 6144 DoFs. Different from the usual time discretization technique, here we use implicit Euler method. Snapshots are taken in the time interval $[0, 1000]$ with the time step $\Delta t = 0.5$. For POD/POD-DEIM basis construction, we use the parameter samples $\mu \in \{-0.04, -0.02, 0, 0.02, 0.04\}$, $n_s = 5$. The reduced systems are solved for the set $\{-0.03, -0.01, 0.01, 0.03\}$ of parameter values of μ , which are not contained in the set of sample parameters. The average number of Newton iterations is 1 and the Newton tolerance is $1e - 10$ for the computation of the FOMs and ROMs on each time step.

Let \mathcal{U} , \mathcal{V} , \mathcal{F} denote the singular values for the solution snapshots \vec{u}_1 , \vec{u}_2 and nonlinear snapshots. Then, the decay of those singular values is given in Figure 3.10, left, and the CPU times of the FOMs and ROMs for the parameter value $\mu = 0.03$ are shown in Figure 3.10, right.

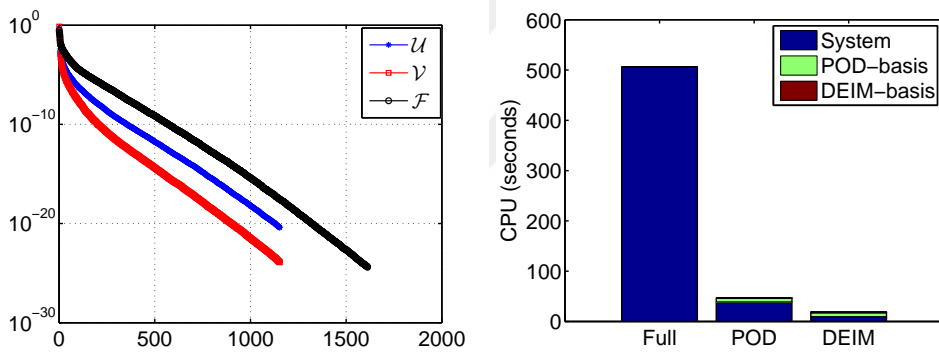


Figure 3.10: (Left) Decay of the singular values of solution snapshots \mathcal{U} , \mathcal{V} and of the nonlinear snapshots \mathcal{F} ; (right) CPU times for the computation of FOMs, POD and POD-EIM ROMs for the parameter value $\mu = 0.03$.

Table 3.2: The computation times (in sec), speed-up factors S_{POD} and S_{DEIM} , and the DEIM projection error bounds $\|(P^T W)^{-1}\|_2$.

μ	FOM	POD	POD-DEIM	S_{POD}	S_{DEIM}	$\ (P^T W)^{-1}\ _2$
-0.03	527.3	34.5	7.5	15.31	70.21	28
-0.01	501.9	33.4	13.2	15.05	38.08	33
0.01	522.3	32.9	11.9	15.88	43.67	41
0.03	505.9	38.6	9.0	13.10	56.43	33

In Table 3.2 we give the CPU times for FOMs, POD and POD-DEIM ROMs together with the speed-up factors S_{POD} and S_{DEIM} , which demonstrate the efficiency of the DEIM. We note that in the POD-DEIM algorithm, the nonlinearity is discretized at six points of the mesh by linear continuous FEM, whereas at three points of the mesh

by linear DG method (see Figure 3.1). Therefore, during the online computation, the DG requires less more work than the continuous FEM. In Figure 3.11, the patterns of FOMs, POD and POD-DEIM reduced solutions are shown at the final time $T = 1000$. The ROM patterns in Figure 3.11 computed with POD are very close to those of the FOMs as in [61]. But the patterns computed with POD-DEIM are less accurate than those with the POD computed ones for some parameter values in Figure 3.11. The DEIM does not improve the accuracy of the POD reduced model, but enormously reduces the computational complexity [5]. The error bounds $\|(P^T W)^{-1}\|_2$ of moderate size for the DEIM approximations are also given in Table 3.2.

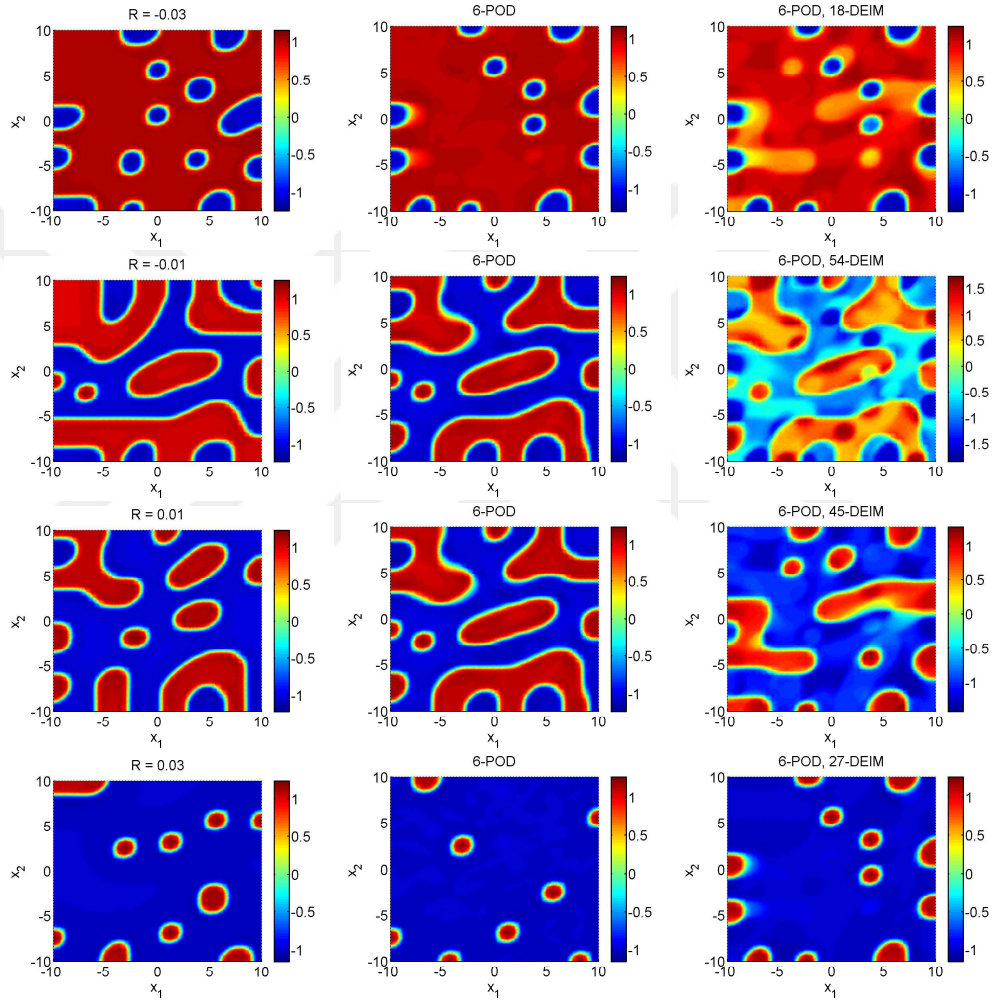


Figure 3.11: Pattern solutions for u_1 at the final time $T = 1000$ with FOM (left), POD (middle) and POD-DEIM (right) for the parameter values $\mu \in \{-0.03, -0.01, 0.01, 0.03\}$ from top to bottom.

3.5.3 Convective FHNE

In this part, we consider the convective FHNE on $\Omega = [0, 60] \times [0, 15]$ with zero flux except the left boundary, which is Dirichlet 0.1. As initial conditions we set

$u_1 = 0.1$ inside a narrow triangle $0 \leq x \leq 0.5$, $0 \leq y \leq 15$ and $u_1 = 0$ outside this rectangle; $u_2 = 0$ everywhere. The spatial and temporal mesh size are chosen as $\Delta x_1 = \Delta x_2 = 0.5$ and $\Delta t = 0.05$, respectively. The parameters are chosen as $(\tau_1, \tau_2, d_2, c_1, c_2, c_3, \epsilon) = (1, 1, 1, 1, 9, 0.02, 5, 0.1)$.

We show numerical results for FOM and ROMs computed by POD, POD-DEIM, DMD (as an equation-free), and POD-DMD. We use $k = 15$ POD basis functions, $m = 15$ DEIM basis functions, and $r = 21$ DMD basis functions. To form the snapshot matrix we use 101 time instances resulting in a matrix in $\mathbb{R}^{21600 \times 101}$ which has the rank 60 and 53 for \vec{u}_1 and \vec{u}_2 , respectively. POD basis computation leads a steep descent in the singular values of the system related to \vec{u}_1 , \vec{u}_2 and the nonlinear term \vec{f} , shown in Figure 3.13. The average number of Newton iterations are 2, 3, 1 and tolerance number for the Newton iteration is $1e - 10$ for FOM, POD, and POD-DEIM, respectively.

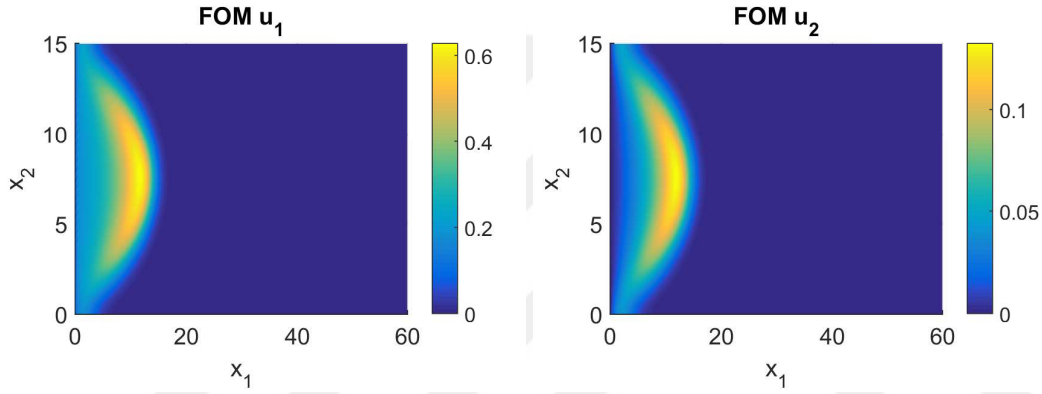


Figure 3.12: FOM solutions for the component u_1 (left), u_2 (right).

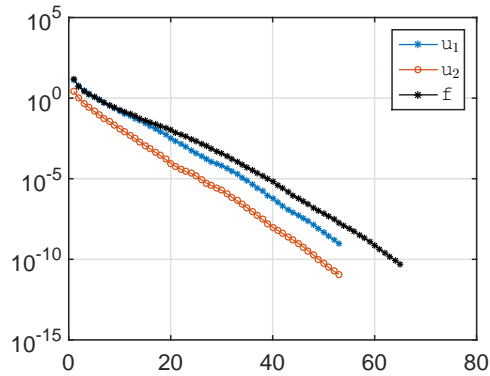


Figure 3.13: FOM solutions for the component u_1 (left), u_2 (right).

The computed wave type solutions, given in Figure 3.12, for the FOM is consisted with the results in [35]. The corresponding solution profiles with ROMs are given for the first state component in Figure 3.14, for the second component in Figure 3.15. Also, their errors are given in the same Figures.

The CPU time to compute POD, DEIM, and DMD basis functions are 1.12, 0.91, 9.51 seconds, respectively.

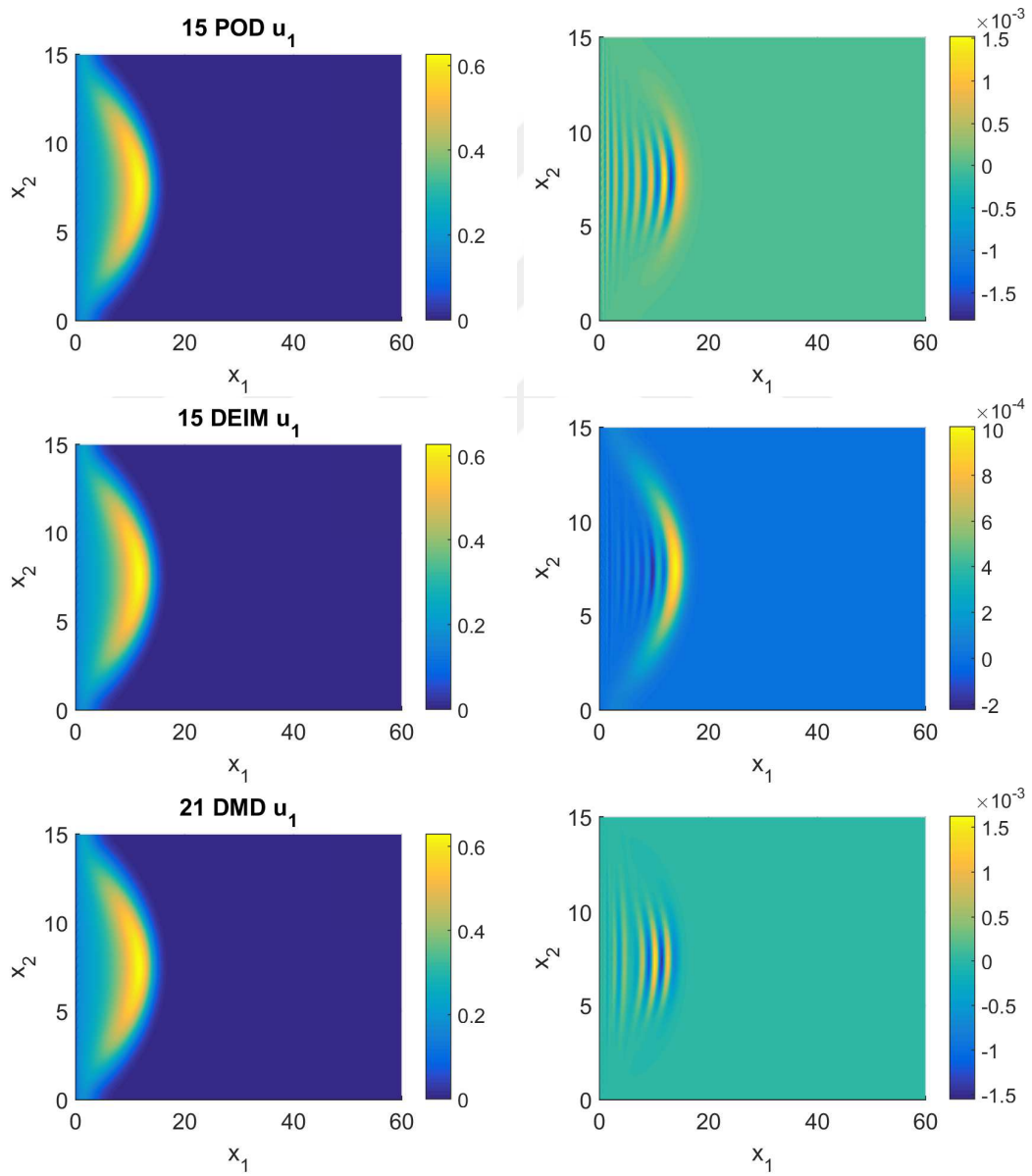


Figure 3.14: From top to bottom POD, DEIM, DMD solutions (left) and errors (right) for the component u_1 .

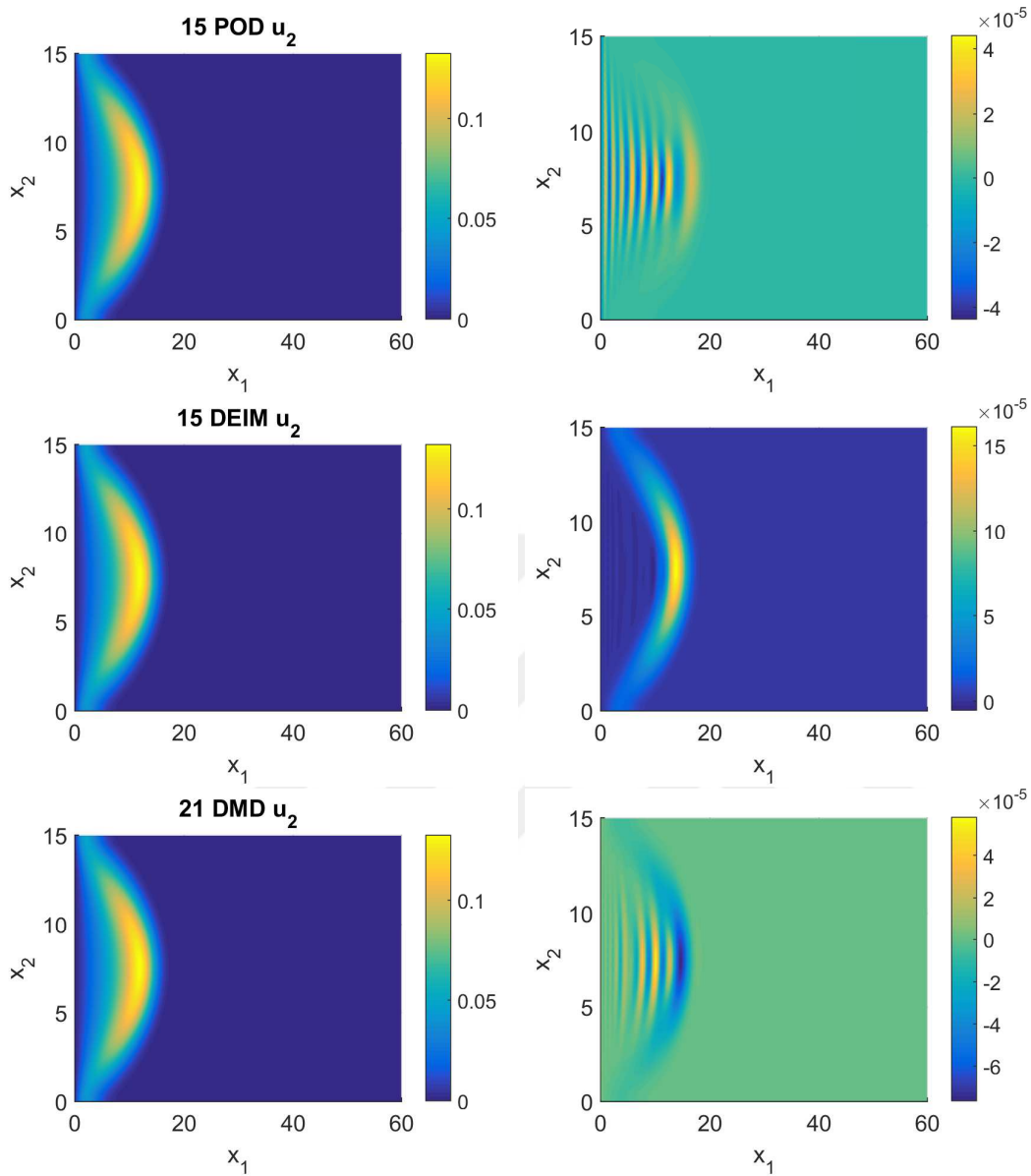


Figure 3.15: From top to bottom POD, DEIM, DMD solutions (left) and errors (right) for the component u_2 .

Table 3.3: FOM-ROM errors & speed-up with 15 POD, 15 DEIM, and 21 DMD basis functions.

	L_2 -Error u_1	L_2 -Error u_2	Speed-up
POD	2.48e-03	9.48e-05	1.49
POD-DEIM	3.13e-03	2.96e-04	51.77
POD-DMD	3.66e-03	2.18e-04	805.60

In Table 3.3, the computed Frobenius errors between FOM and ROM solutions are shown. As expected, the POD errors are less than the other methods, which is ac-

ceptable. Also the other errors are close to FOM errors. Here, the main contribution is having really high speed-up factor for POD-DMD solutions with almost the same errors with POD-DEIM. This makes POD-DMD method more preferable. In a similar to diffusive FHNE results, the errors for the first state variable is greater than the other component. It is again due to the fact that the non-linear term occurs in the first equation.

In order to choose the number of basis functions for ROM methods, we try two cases mentioned in the diffusive FHNE results. If we increase all of them at once, or keep one constant and increase the others, we obtain similar results. There is not much difference between these two cases, as it is seen in Figures 3.16- 3.17. So we prefer to use the second one. First, we check the energy criteria (3.11) in order to choose the number of POD basis functions. When the basis functions contain at least %99.99 of the energy of the full system, we cut from there, and use that number as a number of POD basis functions. Then, the number of the DEIM and the DMD basis functions is increased until the errors are less than a certain and acceptable tolerance.

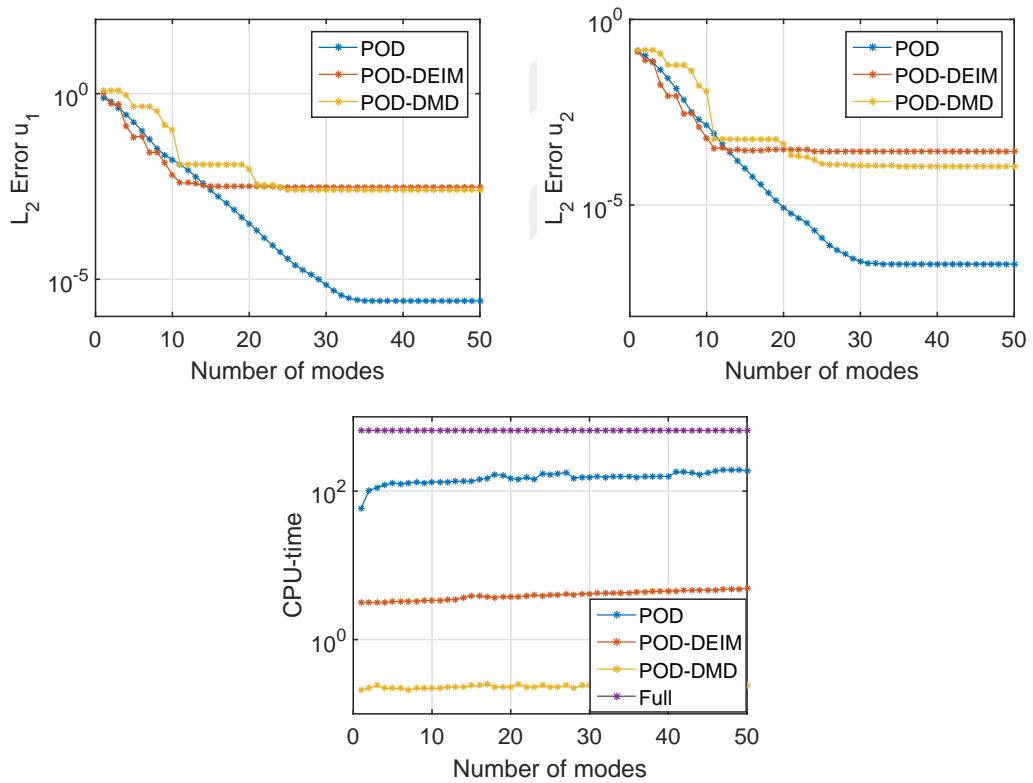


Figure 3.16: (Top) FOM-ROM $L^2(Q)$ errors for the states u_1 (left), u_2 (right) with increasing number of POD for POD solution, and increasing number of DEIM and DMD basis functions with 15 fixed POD basis functions for POD-DEIM and POD-DMD solutions; (bottom) corresponding CPU times.

As expected, increasing the number of POD basis functions leads to more accurate solutions for both cases. The error for the POD is solutions monotonically decreasing.

However, for the POD-DEIM and the POD-DMD it is not much obvious. The rate of decline of their errors is slower.

On the other hand, we have to consider also CPU times to show the efficiency of ROM solutions. As it is seen in Figures 3.16-3.17, all ROM methods we use requires less time than FOM except for the first second case in which when we use many modes in the POD method. Increasing the number of bases does not much affect the time. There is not a big difference between CPU times. As expected again, the POD-DMD provides a significant reduction.

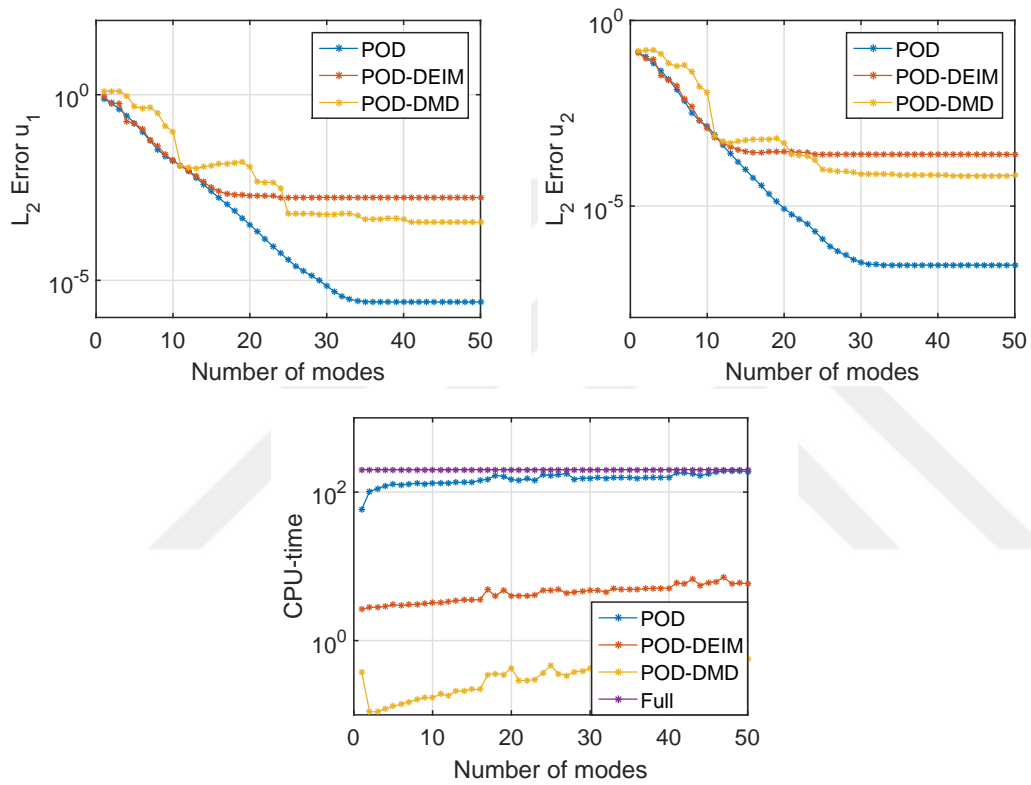


Figure 3.17: (Top) FOM-ROM $L^2(Q)$ errors for the states u_1 (left), u_2 (right) with the same increasing number of POD, DEIM and DMD basis functions; (bottom) corresponding CPU times.

CHAPTER 4

OPTIMAL CONTROL AND SUB-OPTIMAL CONTROL USING REDUCED ORDER MODELLING

In the current years, control of PDEs with wave type solutions received more attention: the control for spiral waves of Schlögl or Nagumo equation [21], spiral and travelling waves of classical FHNE [23, 24, 77], spiral waves of lambda-omega systems [19], and pattern formation [85].

In this Chapter, we investigate a numerical approximation of optimal control of the convective FHNE. As described in Chapter 3, the uncontrolled solutions of the convective FHNE are in the form of travelling waves [35]. Our aim is to control those waves by sparse and non-sparse control in the whole space-time domain and at the terminal time. The sparse control is more efficient and realistic, but the cost function becomes non-smooth. When the controls are localized in some part of the regions, the sparse control provides solutions without any a priori knowledge of these sub-areas. Sparse optimal control was investigated in [83] for linear elliptic equations, and for semi-linear elliptic and parabolic equations in [22, 24, 97].

The DG methods were successfully implemented in a variety of the optimal control problems (OCPs), for instance, for the steady state OCPs [102, 103], the time dependent linear convection-diffusion-reaction OCPs [1, 2], and the semi-linear steady state OCPs [104]. There are two approaches for solving OCPs with PDE-constrained. The first one is the *discretize-then-optimize* approach, where the objective function is discretized and the discrete Lagrangian is built, and then the optimality conditions are derived in the discrete setting. The second one is the *optimize-then-discretize* approach, where the Lagrangian is built for the infinite dimensional problem and then continuous form of the first order optimality conditions are discretized. There is no preferred approach between them, here we follow the *optimize-then-discretize* approach. The first order optimality conditions are derived and then they are discretized by using the DG method. We employ the projected nonlinear conjugate gradient (CG) method [41] for solving the nonlinear discrete optimization problem.

The OCP problem governed by the convective FHNE is non-convex because of the semi-linearity of the activator equation. Therefore, the fulfillment of the first order necessary conditions does not guarantee the optimality. Also the second order sufficient optimality conditions (SSCs) have to be checked. In the recent years, the fulfillment of the SSCs for infinite dimensional and finite dimensional semi-linear PDE-constrained

optimal control problems has been investigated extensively [25]. The proof of the SSCs for the infinite dimensional problems is not possible because of the necessity of optimal solution, except for few examples whose analytical solutions exist. Therefore the finite dimensional approximations of the infinite dimensional problem are considered. Provided that the local minima satisfies the SSCs, one can check the SSCs numerically by finding a bound for the distance between the local minima and discrete solution [75]. For this purpose, the associated coercivity constant of the reduced Hessian operator is estimated numerically by computing its smallest eigenvalue. Similar techniques were applied to measure how far the control obtained by a reduced order optimization model is away from a local full order control solution [49]. Moreover, the Tikhonov regularization parameter in the cost function expresses the cost of the control, and it increases the numerical stability of the optimal solution. Recently the second order optimality conditions have been investigated for semi-linear parabolic control problems with the objective function, not including the Tikhonov regularization term [25]. We test the discrete optimization problem for vanishing Tikhonov regularization parameter as done in [24]. The numerical results of the control of two dimensional waves confirm the convergence of the optimal solutions for vanishing Tikhonov regularization parameter as it was demonstrated for the one dimensional wave solutions of the classical FHNE in [24, 77].

The numerical solution of OCP governed by PDEs are time consuming due to the repetitive solutions of PDEs during the optimization. To reduce the computational time and complexity, the sub-optimal control problems with reduced solutions are considered [15]. The reduced order OCPs are solved usually applying the Galerkin projection based on the proper orthogonal decomposition (POD) using the snapshots of the discretized state equation. Increasing the number of the POD basis functions leads to more accurate reduced order or sub-optimal controls, which can be considered as a compensate for the lack of an a priori analysis of the POD. ROMs using the POD-Galerkin projection for OCPs with linear PDE-constrained are investigated in [45, 56, 89] and with semi-linear PDE-constrained in [49, 86, 53]. To deal efficiently with the nonlinear term in the activator equation of the convective FHNE, we use the DEIM [28] and the DMD [3, 4] as described in Chapter 3.

To our best knowledge, reduced order optimal controls using POD-DMD approach have not discussed in literature yet. When dealing with the MOR techniques, having a balance between the cost of calculation and the accuracy of the reduced solutions is indispensable. The fastest of these three methods is obviously POD-DMD, because the reduced system in the discrete form becomes linear and the OCP problem will be convex. Again the POD is the most accurate, but it is the slowest. The success of MOR techniques depends on the type of the problem. The reduced or sub-optimal control of semi-linear PDEs [21] and convection dominated problems with travelling wave solutions require greater number of POD basis functions [66], which is also confirmed by numerical results in Section 4.5.

The outline of this chapter is as follows: We give basic definition of function spaces in Section 4.1. In Section 4.2, the FOM of sparse optimal control problem with pointwise box constraints governed by convective FHNE is described. We first prove the existence and uniqueness of a convective FHNE, called as state equation, by transforming

into the one with monotone nonlinearity. Then, we introduce the first and second order optimality conditions. The fully discrete form of the optimality system is derived in Section 4.3. In Section 4.4, we propose reduced order models as, POD, POD-DEIM and POD-DMD for the OCP. We compare these three approaches to solve our sparse optimal control problem and to give POD-DEIM and POD-DMD algorithms for OCPs. Finally, we present numerical examples to confirm the accuracy and efficiency of the MOR techniques in Section 4.5.

4.1 Preliminaries

In this section, we introduce the function spaces which we use in this chapter.

Definition 4.1. (Lebesgue integrable functions) Let Ω be a Lebesgue measurable and bounded subset of \mathbb{R}^n , $n \geq 1$. The Lebesgue spaces L^p are defined

$$L^p(\Omega) := \{u : \Omega \rightarrow \mathbb{R} \mid \|u\|_{L^p(\Omega)} < \infty\}$$

with the endowed norm

$$\|u\|_{L^p(\Omega)} = \left(\int_{\Omega} |u(x)|^p dx \right)^{\frac{1}{p}},$$

for $1 \leq p < \infty$.

For $p = \infty$,

$$L^\infty(\Omega) := \{u : \Omega \rightarrow \mathbb{R} \mid \|u\|_{L^\infty(\Omega)} < \infty\}$$

with the norm

$$\|u\|_{L^\infty(\Omega)} := \operatorname{ess\,sup}_{x \in \Omega} |u(x)|.$$

For $1 \leq p \leq \infty$, locally Lebesgue spaces are defined as

$$L^p_{loc}(\Omega) := \{u : \Omega \rightarrow \mathbb{R} \mid u \in L^p(\Omega_0) \text{ for any compact subset } \Omega_0 \subset \Omega\}.$$

The space $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ is a Banach space for $1 \leq p \leq \infty$. In the case $p = 2$, the $L^2(\Omega)$ space is also a Hilbert space and denotes the set of square integrable functions with the inner product

$$(u, v)_{L^2(\Omega)} = \int_{\Omega} u(x)v(x)dx.$$

Definition 4.2. (Weak derivative) Let $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ be a multi-index with order $|\alpha| = \alpha_1 + \dots + \alpha_n$. For $u \in L^p_{loc}(\Omega)$,

$$D^\alpha u := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} u$$

is called α -th weak partial derivative of u .

Definition 4.3. (Sobolev spaces) Let Ω be a bounded subset of \mathbb{R}^n , $n \geq 1$. Then the Sobolev spaces $W^{k,p}(\Omega)$ are defined by

$$W^{k,p}(\Omega) := \{u \in L^p(\Omega) \mid D^\alpha u \in L^p(\Omega) \text{ for } 0 \leq |\alpha| \leq k\}$$

with the endowed norm

$$\|u\|_{W^{k,p}(\Omega)} := \left(\sum_{0 \leq |\alpha| \leq k} \int_{\Omega} |D^\alpha u(x)|^p dx \right)^{\frac{1}{p}},$$

for $p \geq 1$.

The space $(W^{k,p}(\Omega), \|\cdot\|_{W^{k,p}(\Omega)})$ is a Banach space. In the case $p = 2$, the Sobolev spaces are defined as $W^{k,2}(\Omega) := H^k(\Omega)$, and constitutes Hilbert spaces. The set of Hilbert spaces $H^k(\Omega)$ is given as

$$H^k(\Omega) := \{u \in L^2(\Omega) \mid D^\alpha u \in L^2(\Omega) \text{ for } 0 \leq |\alpha| \leq k\},$$

associated with the following inner product

$$(u, v)_{H^k(\Omega)} = \sum_{0 \leq |\alpha| \leq k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}.$$

In the case $k = 1$, the set of the Hilbert space $H^1(\Omega)$ is written as

$$H^1(\Omega) := \{u \in L^2(\Omega) \mid \frac{\partial}{\partial x_i} u \in L^2(\Omega), i = 1, \dots, n\}$$

with the norm

$$\|u\|_{H^1(\Omega)} := \left(\int_{\Omega} |u|^2 + |\nabla u|^2 dx \right)^{\frac{1}{2}}.$$

Furthermore, the set of measurable functions spaces defined from $(0, T)$ to the normed space X are written as

$$L^q(0, T; X) = \{u : [0, T] \rightarrow X \mid \|u(t)\|_{L^q(0,T;X)} < \infty\}$$

with the norm

$$\|u(t)\|_{L^q(0,T;X)} := \left(\int_0^T \|u(t)\|_X^q dt \right)^{\frac{1}{q}},$$

for $1 \leq q \leq \infty$. For $p = \infty$,

$$\|u(t)\|_{L^\infty(0,T;X)} := \text{ess sup}_{t \in [0,T]} \|u(t)\|_X.$$

We can define also the following Hilbert space

$$W(0, T) := \{u \in L^2(0, T; V); \mid u' \in L^2(0, T; V^*)\},$$

associated with the norm

$$\|u\|_{W(0,T)} = \left(\int_0^T (\|u(t)\|_V^2 + \|u_t(t)\|_{V^*}^2) dt \right)^{\frac{1}{2}},$$

where $V = H^1(\Omega)$ and V^* is the dual space of V .

4.2 Full Order Optimal Control

We consider the following OCP

$$\min_{u \in \mathcal{U}_{ad}} J(u) := I(u) + \mu j(u), \quad (4.1)$$

where

$$\begin{aligned} I(u) &= \frac{\omega_Q}{2} \left(\int_0^T \int_{\Omega} (u_1(x, t) - u_{1Q}(x, t))^2 dx dt + \int_0^T \int_{\Omega} (u_2(x, t) - u_{2Q}(x, t))^2 dx dt \right) \\ &\quad + \frac{\omega_T}{2} \left(\int_{\Omega} (u_1(x, T) - u_{1T}(x))^2 dx + \int_{\Omega} (u_2(x, T) - u_{2T}(x))^2 dx \right) \\ &\quad + \frac{\lambda}{2} \int_0^T \int_{\Omega} (u(x, t))^2 dx dt, \\ j(u) &= \int_0^T \int_{\Omega} |u(x, t)| dx dt, \end{aligned}$$

subject to the convective FHNE in Q_T

$$\begin{aligned} u_{1t}(x, t) - d_1 \Delta u_1(x, t) + \mathbf{V} \cdot \nabla u_1(x, t) + f(u_1(x, t)) + u_2(x, t) &= u(x, t), \\ u_{2t}(x, t) - d_2 \Delta u_2(x, t) + \mathbf{V} \cdot \nabla u_2(x, t) + \epsilon(u_2(x, t) - c_3 u_1(x, t)) &= 0 \end{aligned} \quad (4.2)$$

with the Neumann and Dirichlet boundary conditions on Σ_T^N

$$\begin{aligned} \partial_n u_1(x, t) &= 0, \quad \partial_n u_2(x, t) = 0, \\ u_1(x, t) &= u_{1D}(x, t), \quad u_2(x, t) = u_{2D}(x, t) \end{aligned}$$

and with the initial conditions

$$u_1(x, 0) = u_{10}(x), \quad u_2(x, 0) = u_{20}(x) \quad \text{in } \Omega,$$

and subject to the pointwise box constraints

$$u \in \mathcal{U}_{ad} := \{u \in L^\infty(Q) : u_l \leq u(x, t) \leq u_r \text{ for a.e } (x, t) \in Q_T\} \quad (4.3)$$

with the real numbers $u_l \leq u_r$.

The function $f(u_1)$ denotes the cubic polynomial nonlinearity

$$f(u_1) = c_1 u_1 (u_1 - c_2) (u_1 - 1) \quad (4.4)$$

with the non-negative real numbers c_i , $i = 1, 2$, which is monostable for $0 < c_1 < 20.803$ and $c_2 = 0.02$.

The functions $u_{1T}, u_{2T} \in L^\infty(\Omega)$ and $u_{1Q}, u_{2Q} \in L^\infty(Q_T)$ denote the given targets or desired states for the state variables u_1 and u_2 in the final domain or whole domain

control, respectively. We have given constants ω_Q, ω_T . The parameter $\lambda \geq 0$ stands for the Tikhonov regularization parameter, and $\mu \geq 0$ for the sparse parameter.

Here, we make the following assumption for the solutions u_1, u_2

$$0 =: u_{10} \leq u_1 \leq u_{11}, \quad 0 =: u_{20} \leq u_2 \leq u_{21} \quad \text{a.e. } Q_T, \quad (4.5)$$

for some constant value bounds. This presumption is acceptable for physical authenticity.

The aim of the OCP is to ensure that the state variables u_1, u_2 are as close as possible to the desired states u_{1Q}, u_{2Q} for the whole domain with chosen parameters $\omega_Q = 1, \omega_T = 0$, or to u_{1T}, u_{2T} for the final time control with $\omega_Q = 0, \omega_T = 1$. The activator component u_1 of the state equation (4.2) is to be controlled by u , while the inhibitor u_2 has only some auxiliary character regarding the control with respect to the established classical FHNE [23, 24, 77] since just controlling the activator appears to be plausible [95].

From the biological point of view, a mathematical modelling of an excitable medium requires at least two equations with activator and inhibitor components. In blood coagulation process, although the activator component variable u_1 describes the concentrations of thrombin the excitation itself, the inhibitor component variable describes the inhibition of this excitation and recovery of the medium variables, u_2 activated factor XI [34]. The complex process of coagulation comprises of cascadic enzymatic reactions and feedback loops, separated into three phases, initiation (restricted at the vascular harm site); propagation, or spatial expansion of coagulation wave into the vessel; end of the biochemical responses and cluster expansion stoppage [59]. These ultimately catalyze fibrinogen conversion into fibrin, which polymerizes to form a clot. These reactions allow autocatalytic thrombin generation far from the damage site. The most important property of blood coagulation process is the formation of autowaves with the velocity independent of the initial conditions [35, 34, 59].

4.2.1 Well-posedness of the State Equation

In this Section, we prove that the state equation (4.2) has a unique solution for each $u \in \mathcal{U}_{ad}$. Because of the non-monotonicity of the state equation which results from nonlinearity in the activator component, the existence-uniqueness theory for the solution of the state equation (4.2) is a bit delicate [21, 23]. The well-posedness of the state equation can be proved in the literature using different methods like the Schauder fixed point theorem applied to the classical FHNE [47], the Leray-Schauder fixed-point theorem for the weak solution of semi-linear parabolic equations, Brouwer's fixed point theorem for non-monotone quasi-linear elliptic discrete problem [39], and the Faedo–Galerkin method for an anti-periodic quasi-linear hemivariational inequality [71]. After transforming the non-monotone nonlinearity in the activator equation to a monotone one [21], we derive upper and lower solutions [14, 70] to prove the well-posedness of the state equation (4.2).

We transform the state equation (4.2) by substituting $u_1 = e^{\eta t} v_1$. Since the derivative

of the cubic nonlinearity $f(u_1)$ is a quadratic, then it is bounded from below by a constant, say c , then η is chosen as $|c|$ [21]. Then, we obtain the following system:

$$\begin{aligned} v_{1t} - d_1 \Delta v_1 + \mathbf{V} \cdot \nabla v_1 + e^{-\eta t} f(e^{\eta t} v_1) + \eta v_1 &= e^{-\eta t} (u - u_2), \\ u_{2t} - d_2 \Delta u_2 + \mathbf{V} \cdot \nabla u_2 + \epsilon (u_2 - c_3 e^{\eta t} v_1) &= 0 \end{aligned} \quad (4.6)$$

with the boundary conditions

$$\begin{aligned} \partial_n v_1 &= 0, \quad \partial_n u_2 = 0, \\ e^{\eta t} v_1 &= u_{1D}, \quad u_2 = u_{2D}, \end{aligned}$$

and the initial conditions

$$v_1(x, 0) = u_{10}(x), \quad u_2(x, 0) = u_{20}(x).$$

Here, the nonlinear term

$$\tilde{f} : v_1 \mapsto e^{-\eta t} g(e^{\eta t} v_1) + \eta v_1$$

is a monotone non-decreasing function with respect to v_1 for all $(x, t) \in Q_T$ [88, Sec. 4.3]:

- (i) For every fixed $v_1 \in \mathbb{R}$ is Lebesgue measurable in Q_T .
- (ii) For almost all $(x, t) \in Q_T$, \tilde{f}_1 is twice continuously differentiable with respect to v_1 and locally Lipschitz continuous of order 2 with respect to v_1 , i.e., there exists $L(\rho) = 6c_1 e^{2\eta t} > 0$ such that

$$\left| \frac{\partial f}{\partial v_1 \partial v_1}(x, t, v_1) - \frac{\partial f}{\partial v_1 \partial v_1}(x, t, v_2) \right| \leq L(\rho) |v_{11} - v_{12}|$$

holds with for all $v_{11}, v_{12} \in \mathbb{R}$ with $|v_{1i}| \leq \rho, i = 1, 2$.

The nonlinearity is uniformly bounded and monotone increasing:

- (iii) There exists a constant $C = c_1 c_2 + \eta + 2e^{\eta t} (c_1 c_2 + c_1) > 0$ such that

$$|\tilde{f}_1(x, t, 0)| + |\tilde{f}_{1v}(x, t, 0)| + |\tilde{f}_{1vv}(x, t, 0)| \leq C.$$

- (iv) It holds $0 \leq \tilde{f}_{1v_1}(x, t, v_1)$ for almost all $(x, t) \in Q_T$, all $v_1 \in \mathbb{R}$.

Definition 4.4. A pair of functions $(v_1, u_2) \in (W(0, T) \cap L^\infty(Q_T))^2$ is called weak solution of the system (4.6) if

$$\begin{aligned} \int_0^T (v_{1t}, \varphi)_{V^*, V} dt + \iint_{Q_T} (d_1 \nabla v_1 \cdot \nabla \varphi + \mathbf{V} \cdot \nabla v_1 \varphi + \tilde{f} \varphi - e^{\eta t} (u - u_2) \varphi) dx dt &= 0, \\ \int_0^T (u_{2t}, \varphi)_{V^*, V} dt + \iint_{Q_T} (d_2 \nabla u_2 \cdot \nabla \varphi + \mathbf{V} \cdot \nabla u_2 \varphi + \epsilon (u_2 - c_3 e^{\eta t} v_1) \varphi) dx dt &= 0 \end{aligned} \quad (4.7)$$

with $V = H^1(\Omega)$ and its dual space V^* and

$$v_1(\cdot, 0) = u_{10}, \quad u_2(\cdot, 0) = u_{20}$$

are satisfied for all $\varphi \in L^2(0, T; H^1(\Omega))$. Here ∇ denotes the gradient with respect to x .

Definition 4.5. The pair functions $(\tilde{v}_1, \tilde{u}_2)$ and (\hat{v}_1, \hat{u}_2) in $(W(0, T) \cap L^\infty(Q_T))^2$ are said to be ordered upper and lower solutions of (4.6), respectively, if they satisfy

$$(\hat{v}_1, \hat{u}_2) \leq (\tilde{v}_1, \tilde{u}_2)$$

and

$$\begin{aligned} \hat{v}_{1t} - d_1 \Delta \hat{v}_1 + \mathbf{V} \cdot \nabla \hat{v}_1 + \tilde{f}(x, t, \hat{v}_1) - e^{-\eta t}(u - \tilde{u}_2) &\leq 0 \\ &\leq \tilde{v}_{1t} - d_2 \Delta \tilde{v}_1 + \mathbf{V} \cdot \nabla \tilde{v}_1 + \tilde{f}(x, t, \tilde{v}_1) - e^{-\eta t}(u - \hat{u}_2), \\ \hat{u}_{2t} - d_1 \Delta \hat{u}_2 + \mathbf{V} \cdot \nabla \hat{u}_2 + \epsilon(u_2 - c_3 e^{\eta t} \tilde{v}_1) &\leq 0 \\ &\leq \tilde{u}_{2t} - d_2 \Delta \tilde{u}_2 + \mathbf{V} \cdot \nabla \tilde{u}_2 + \epsilon(u_2 - c_3 e^{\eta t} \hat{v}_1), \\ \partial_n \hat{v}_1 &\leq 0 &\leq \partial_n \tilde{v}_1, \\ \partial_n \hat{u}_2 &\leq 0 &\leq \partial_n \tilde{u}_2, \\ \hat{v}_1 &\leq e^{-\eta t} u_{1D} &\leq \tilde{v}_1, \\ \hat{u}_2 &\leq u_{2D} &\leq \tilde{u}_2, \\ \hat{v}_1(x, 0) &\leq u_{10}(x) &\leq \tilde{v}_1(x, 0), \\ \hat{u}_2(x, 0) &\leq u_{20}(x) &\leq \tilde{u}_2(x, 0). \end{aligned}$$

Taking

$$\tilde{v}(x, t) = \tilde{z}(x, t) = M, \quad \hat{v}(x, t) = \hat{z}(x, t) = 0$$

for some $M > 0$ in (4.6) leads the following system:

$$\begin{aligned} v_{1t} - d_1 \Delta v_1 + \mathbf{V} \cdot \nabla v_1 + \tilde{f}(e^{\eta t} v_1) - e^{-\eta t}(u - M) &= e^{-\eta t}(M - u_2), \\ u_{2t} - d_2 \Delta u_2 + \mathbf{V} \cdot \nabla u_2 + \epsilon(u_2 - c_3 e^{\eta t} M) &= \epsilon c_3 e^{\eta t}(M - v_1), \\ \partial_n v_1 &= 0, \quad \partial_n u_2 = 0, \\ v_1 &= e^{-\eta t} u_{1D}, \quad u_2 = u_{2D}, \\ v_1(x, 0) &= u_{10}(x), \quad u_2(x, 0) = u_{20}(x). \end{aligned}$$

Here, we have

$$\begin{aligned} \frac{\partial(e^{-\eta t}(M - u_2))}{\partial v_1}, \quad \frac{\partial(\epsilon c_3 e^{\eta t}(M - v_1))}{\partial u_2} &\geq 0, \\ \frac{\partial(e^{-\eta t}(M - u_2))}{\partial u_2}, \quad \frac{\partial(\epsilon c_3 e^{\eta t}(M - v_1))}{\partial v_1} &\leq 0 \end{aligned}$$

for all $v_1, u_2 \in [0, M]$.

The existence and uniqueness of the system (4.6) for each control variable u is given by the following theorem:

Theorem 4.1. *Suppose that the initial conditions u_{10} and u_{20} are non-negative functions, and (4.5) holds. Then, the system (4.6) admits a unique solution $(v_1, u_2) \in (W(0, T) \cap C(Q_T))^2$ for each control $u \in \mathcal{U}_{ad}$.*

Proof. We adopt the iteration technique introduced in [70] and construct sequences $\{(\bar{v}_1^k, \bar{u}_2^k)\}_{k=0}^\infty, \{(\underline{v}_1^k, \underline{u}_2^k)\}_{k=0}^\infty$ with initial elements

$$\begin{aligned}\bar{v}_1^0 &= \tilde{v}_1 = M, & \bar{u}_2^0 &= \tilde{u}_2 = M, \\ \underline{v}_1^0 &= \hat{v}_1 = 0, & \underline{u}_2^0 &= \hat{u}_2 = 0.\end{aligned}$$

$$\begin{aligned}\bar{v}_{1t}^{k+1} - d_1 \Delta \bar{v}_1^{k+1} + \mathbf{V} \cdot \nabla \bar{v}_1^{k+1} + \tilde{f}(e^{\eta t} \bar{v}_1^{k+1}) - e^{-\eta t}(u - M) &= e^{-\eta t}(M - \underline{u}_2^k), \\ \bar{u}_{2t}^{k+1} - d_2 \Delta \bar{u}_2^{k+1} + \mathbf{V} \cdot \nabla \bar{u}_2^{k+1} + \epsilon(\bar{u}_2^{k+1} - c_3 e^{\eta t} M) &= \epsilon c_3 e^{\eta t}(M - \underline{v}_1^k), \\ \partial_n \bar{v}_1^{k+1} &= 0, & \partial_n \bar{u}_2^{k+1} &= 0, \\ \bar{v}_1^{k+1} &= e^{-\eta t} u_{1D}, & \bar{u}_2^{k+1} &= u_{2D}, \\ \bar{v}_1^{k+1}(x, 0) &= u_{10}(x), & \bar{u}_2^{k+1}(x, 0) &= u_{20}(x)\end{aligned}$$

and

$$\begin{aligned}\underline{v}_{1t}^{k+1} - d_1 \Delta \underline{v}_1^{k+1} + \mathbf{V} \cdot \nabla \underline{v}_1^{k+1} + \tilde{f}(e^{\eta t} \underline{v}_1^{k+1}) - e^{-\eta t}(u - M) &= e^{-\eta t}(M - \bar{u}_2^k), \\ \underline{u}_{2t}^{k+1} - d_2 \Delta \underline{u}_2^{k+1} + \mathbf{V} \cdot \nabla \underline{u}_2^{k+1} + \epsilon(\underline{u}_2^{k+1} - c_3 e^{\eta t} M) &= \epsilon c_3 e^{\eta t}(M - \bar{v}_1^k), \\ \partial_n \underline{v}_1^{k+1} &= 0, & \partial_n \underline{u}_2^{k+1} &= 0, \\ \underline{v}_1^{k+1} &= e^{-\eta t} u_{1D}, & \underline{u}_2^{k+1} &= u_{2D}, \\ \underline{v}_1^{k+1}(x, 0) &= u_{10}(x), & \underline{u}_2^{k+1}(x, 0) &= u_{20}(x),\end{aligned}$$

respectively. The constructed sequence $\{(\bar{v}_1^k, \bar{u}_2^k)\}_{k=0}^\infty$ is monotone non-increasing and upper solution for all k . Conversely $\{(\underline{v}_1^k, \underline{u}_2^k)\}_{k=0}^\infty$ is monotone non-decreasing and lower solution for all k . Further, we have

$$\underline{u}^k(x, t) \leq \bar{u}^k(x, t) \quad \text{and} \quad \underline{v}^k(x, t) \leq \bar{v}^k(x, t)$$

for all $k \in \mathbb{N}$ and $(x, t) \in Q_T$.

By induction, we can verify the monotonicity of the sequence $\{\bar{v}_1^k\}_{k=0}^\infty$: For $k = 0$,

$$\begin{aligned}\bar{v}_{1t}^1 - d_1 \Delta \bar{v}_1^1 + \mathbf{V} \cdot \nabla \bar{v}_1^1 + \tilde{g}(e^{\eta t} \bar{v}_1^1) - e^{-\eta t}(u - M) &= e^{-\eta t}(M - \underline{u}_1^0), \\ \bar{u}_{2t}^1 - d_2 \Delta \bar{u}_2^1 + \mathbf{V} \cdot \nabla \bar{u}_2^1 + \epsilon(\bar{u}_2^1 - c_3 e^{\eta t} M) &= \epsilon c_3 e^{\eta t}(M - \underline{v}_1^0), \\ \partial_n \bar{v}_1^1 &= 0, & \partial_n \bar{u}_2^1 &= 0, \\ \bar{v}_1^1 &= e^{-\eta t} u_{1D}, & \bar{u}_2^1 &= u_{2D}, \\ \bar{v}_1^1(x, 0) &= u_{10}(x), & \bar{u}_2^1(x, 0) &= u_{20}(x).\end{aligned}$$

The property that \bar{v}_1^0 is an upper solution

$$\begin{aligned}\bar{v}_{1t}^0 - d_1 \Delta \bar{v}_1^0 + \mathbf{V} \cdot \nabla \bar{v}_1^0 + \tilde{f}(e^{\eta t} \bar{v}_1^0) - e^{-\eta t}(u - M) &\geq e^{-\eta t}(M - \underline{u}_2^0), \\ \bar{u}_{2t}^0 - d_2 \Delta \bar{u}_2^0 + \mathbf{V} \cdot \nabla \bar{u}_2^0 + \epsilon(\bar{u}_2^0 - c_3 e^{\eta t} M) &\geq \epsilon c_3 e^{\eta t}(M - \underline{v}_1^0), \\ \partial_n \bar{v}_1^0 &\geq 0, & \partial_n \bar{u}_2^0 &\geq 0, \\ \bar{v}_1^0 &\geq e^{-\eta t} u_{1D}, & \bar{u}_2^0 &\geq u_{2D}, \\ \bar{v}_1^0(x, 0) &\geq u_{10}(x), & \bar{u}_2^0(x, 0) &\geq u_{20}(x).\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}
\bar{v}_{1t}^0 - \bar{v}_{1t}^1 - d_1 \Delta(\bar{v}_1^0 - \bar{v}_1^1) + \mathbf{V} \cdot \nabla(\bar{v}_1^0 - \bar{v}_1^1) + \tilde{f}(e^{nt} \bar{v}_1^0) - \tilde{f}(e^{nt} \bar{v}_1^1) &\geq 0, \\
\bar{u}_{2t}^0 - \bar{u}_{2t}^1 - d_2 \Delta(\bar{u}_2^0 - \bar{u}_2^1) + \mathbf{V} \cdot \nabla(\bar{u}_2^0 - \bar{u}_2^1) + \epsilon(\bar{u}_2^0 - \bar{u}_2^1) &\geq 0, \\
\partial_n(\bar{v}_1^0 - \bar{v}_1^1) &\geq 0, \quad \partial_n(\bar{u}_2^0 - \bar{u}_2^1) \geq 0, \\
(\bar{v}_1^0 - \bar{v}_1^1) &\geq 0, \quad (\bar{u}_2^0 - \bar{u}_2^1) \geq 0, \\
(\bar{v}_1^0 - \bar{v}_1^1)(x, 0) &\geq 0, \quad (\bar{u}_2^0 - \bar{u}_2^1)(x, 0) \geq 0.
\end{aligned}$$

So it follows from the comparison principle for nonlinear parabolic equations $\bar{v}_1^0 - \bar{v}_1^1, \bar{u}_2^0 - \bar{u}_2^1 \geq 0$. Now if we assume that $\bar{v}_1^{k-1} - \bar{v}_1^k, \bar{u}_2^{k-1} - \bar{u}_2^k \geq 0$, one can easily show that $\bar{v}_1^k - \bar{v}_1^{k+1}, \bar{u}_2^k - \bar{u}_2^{k+1} \geq 0$. Analogously, the monotonicity of $(\underline{v}_1^k, \underline{u}_2^k)$ can be proved.

Now, we show the convergence of the sequence $\{\bar{v}_1^k, \bar{u}_2^k\}$ to a solution of (4.6). The sequence $\{\bar{v}_1^k\}$ is monotone non-increasing and bounded from below by $\hat{u} = 0$. Hence by Lebesgue dominated convergence theorem [14, 70], it converges to v_1 in space $L^p(Q)$, $p < \infty$. On the other hand, the sequence $\{\bar{u}_2^k\}$ is monotone non-decreasing and bounded from above by $\tilde{u}_2 = M$. It converges to u_2 in a similar way.

Finally, we prove the uniqueness of the solution of (4.6). Suppose that $(v_1^1, u_2^1), (v_1^2, u_2^2)$ are solutions of (4.7) and set $v_1 := v_1^1 - v_1^2, u_2 := u_2^1 - u_2^2$. Then v_1 and u_2 satisfy the initial conditions obviously. Moreover, the following equations

$$\begin{aligned}
\int_0^T (v_{1t}, \varphi)_{V^*, V} dt + \iint_{Q_T} (d_1 \nabla v_1 \cdot \nabla \varphi + \mathbf{V} \cdot \nabla v_1 \varphi + \tilde{f}(x, t, v_1^1) \varphi \\
- \tilde{f}(x, t, v_1^2) \varphi - e^{nt}(u_2 - u_2^2) \varphi) dx dt = 0, \\
\int_0^T (u_{2t}, \varphi)_{V^*, V} dt + \iint_{Q_T} (d_2 \nabla u_2 \cdot \nabla \varphi + \mathbf{V} \cdot \nabla u_2 \varphi \\
+ \epsilon(u_2 - c_3 e^{nt} v_1) \varphi) dx dt = 0
\end{aligned} \tag{4.8}$$

hold for all $\varphi \in W(0, T)$. Then, following the steps by taking $\varphi = v_1$ in the first equation of (4.8) and $\varphi = u_2$ in the second equation of (4.8), we obtain the desired result $v_1 = 0$ and $u_2 = 0$. \square

The existence of an optimal control u for the optimal control problem (4.1) is given by the following theorem.

Theorem 4.2. *The optimal control problem (4.1) has at least one optimal solution u with associated optimal state u_1 .*

Proof. Here we simply outline the key ideas of the proofs in [88, Sec. 5.3, Thm. 7.4]. Since \mathcal{U}_{ad} is non-empty and bounded in $L^\infty(Q_T)$, it is bounded in any space $L^p(Q_T)$ and it follows from the existence and uniqueness of the state variables that they are also bounded. It follows that, the cost functional is bounded below, and so it has an infimum point. Therefore, because of the boundedness of this sequence, we can find a weakly convergent minimizing sequence. Then, the compact embedding results give

us the strong convergence of the state in a weaker norm. Consequently, there exists a feasible limit point, and using the continuity argument one can show the convergence of the objective function. \square

4.2.2 First Order Optimality Conditions

In this Section, we derive first order optimality conditions for the minimization problem (4.1)

$$\min_{u \in \mathcal{U}_{ad}} J(u) := f(u_1, u_2, u) = I(u) + \mu j(u), \quad (4.9)$$

which is non-convex due to the nonlinearity of $f(u_1)$ (4.4). Therefore, different local minima might occur.

We introduce the Lagrangian for deriving the optimality conditions by

$$\mathcal{L}(u_1, u_2, u, p_1, p_2) = J(u) + (p_1, R_1(u_1, u_2, u)) + (p_2, R_2(u_1, u_2)),$$

where R_1 and R_2 are the first two equations in (4.2), and the adjoint variables p_1 and p_2 are the Lagrange multipliers [88, Chapter 1]. Hence, the first order optimality conditions are obtained as

$$\begin{aligned} \mathcal{L}_{u_1}(u_1, u_2, u; p_1, p_2)u_1 &= 0, \\ \mathcal{L}_{u_2}(u_1, u_2, u; p_1, p_2)u_2 &= 0, \\ \mathcal{L}_{p_1}(u_1, u_2, u; p_1, p_2)p_1 &= 0, \\ \mathcal{L}_{p_2}(u_1, u_2, u; p_1, p_2)p_2 &= 0, \\ \mathcal{L}_u(u_1, u_2, u; p_1, p_2)(v - u) &\geq 0, \quad \forall u \in \mathcal{U}_{ad}, \end{aligned} \quad (4.10)$$

where $\mathcal{L}_{u_i}, \mathcal{L}_{p_i}$ denote Fréchet derivatives with respect to u_i and p_i . The first two equations of (4.10) are the adjoint equations and the last one is the variational inequality. These equations (4.10) with state equations (4.2) constitute the first order optimality conditions.

The cost functional $J(u)$ in (4.9) comprises of two terms with different smoothness. While the first part $I(u)$ is smooth, the second part $j(u) : L^1(Q) \rightarrow \mathbb{R}$ is non-smooth. Although it is not differentiable, it has sub-derivative and its directional derivative is given by

$$j'(u, v - u) = \max_{\sigma \in \partial j(u)} \langle \lambda, v - u \rangle \quad (4.11)$$

with

$$\partial j(u) = \left\{ \sigma \in L^\infty(Q_T) : j(v) \geq j(u) + \int_0^T \int_\Omega \sigma(v - u) dx dt \quad \forall v \in L^\infty(Q_T) \right\},$$

where

$$\sigma(x, t) \in \begin{cases} \{1\}, & \text{if } v(x, t) > 0, \\ [-1, 1], & \text{if } v(x, t) = 0, \\ \{-1\}, & \text{if } v(x, t) < 0. \end{cases}$$

Now, we consider the following weak form of the state equations (4.2)

$$\begin{aligned}
(u_{1t}, w)_\Omega + d_1 a(u_1, w) + b(u_1, w) + (f(u_1), w)_\Omega + (u_2, w)_\Omega &= (u, w)_\Omega, \\
(u_1(\cdot, 0), w)_\Omega &= (u_{10}, w)_\Omega, \\
(u_{2t}, w)_\Omega + d_2 a(u_2, w) + b(u_2, w) + \epsilon(u_2, w)_\Omega - \epsilon c_3(u_1, w)_\Omega &= 0, \\
(u_2(\cdot, 0), v)_\Omega &= (u_{20}, v)_\Omega,
\end{aligned} \tag{4.12}$$

where $(\cdot, \cdot)_\Omega$ denotes the L_2 inner product in Ω , and from now, we set $(\cdot, \cdot) := (\cdot, \cdot)_\Omega$ for easy notation. Here, bilinear forms for diffusive and convective part is separated. The diffusive terms $a(u_1, w) = (\nabla u_1, \nabla w)$, $a(u_2, w) = (\nabla u_2, \nabla w)$ are symmetric and convective terms are $b(u_1, w) = (\mathbf{V} \cdot \nabla u_1, w)$, $b(u_2, w) = (\mathbf{V} \cdot \nabla u_2, w)$ non-symmetric form. We also use $w \in H_1^0(\Omega)$ as the test function.

Using Lagrangian framework (4.10), we obtain the following weak form for the adjoint equations

$$\begin{aligned}
-(p_{1t}, w) + d_1 a(p_1, w) - b(p_1, w) + (f_{u_1} p_1, w) - \epsilon c_3(p_2, w) &= 0, \\
(p_1(\cdot, T), w) &= (u_1 - u_{1T}, w), \\
-(p_{2t}, w) + d_2 a(p_2, w) - b(p_2, w) + \epsilon(p_2, w) + (p_1, w) &= 0, \\
(p_2(\cdot, T), w) &= (u_2 - u_{2T}, w)
\end{aligned} \tag{4.13}$$

with the mixed boundary conditions

$$d_1 \partial_n p_1(x, t) + (\mathbf{V} \cdot \mathbf{n}) p_1(x, t) = 0, \quad d_2 \partial_n p_2(x, t) + (\mathbf{V} \cdot \mathbf{n}) p_2(x, t) = 0, \tag{4.14}$$

the Dirichlet boundary conditions

$$p_1(x, t) = 0, \quad p_2(x, t) = 0,$$

and the final time conditions

$$p_1(x, T) = u_1(x, T) - u_{1T}(x), \quad p_2(x, T) = u_2(x, T) - u_{2T}(x).$$

As it is observed, the bilinear forms of the convection terms $b(p_1, w)$, $b(p_2, w)$ in (4.13) and $b(u_1, w)$, $b(u_2, w)$ in (4.12) have opposite directions.

In addition to adjoint equations, we obtain the following variational inequality

$$I'(u)(v - u) + \mu j'(u, v - u) \geq 0, \quad \forall v \in \mathcal{U}_{ad}, \tag{4.15}$$

or equivalently,

$$\int_0^T \int_\Omega (p_1(x, t) + \lambda u(x, t) + \mu \sigma(x, t)) (v(x, t) - u(x, t)) \, dx \, dt \geq 0, \quad \forall v \in \mathcal{U}_{ad},$$

where p_1 is the first adjoint variable and solution of the (4.13).

For positive Tikhonov regularization λ and sparse μ parameters, we have the following projection formulas [24, 77, 22, 23, 88] resulting from the variational inequality (4.15)

$$u(x, t) = \mathbb{P}_{[u_a, u_b]} \left\{ -\frac{1}{\lambda} (p_1(x, t) + \mu\sigma(x, t)) \right\}, \quad \text{for a.a. } (x, t) \in Q_T, \quad (4.16)$$

$$\lambda(x, t) = \mathbb{P}_{[-1, 1]} \left\{ -\frac{1}{\mu} p(x, t) \right\}, \quad \text{for almost all } (x, t) \in Q_T \quad (4.17)$$

with the following projection operator

$$\mathbb{P}_{[a, b]}(f) = \max\{a, \min\{f, b\}\}.$$

Moreover, we have the following relation

$$u(x, t) = 0 \Leftrightarrow \begin{cases} |p(x, t)| \leq \mu, & \text{if } u_a < 0, \\ p(x, t) \geq -\mu, & \text{if } u_a = 0, \end{cases} \quad (4.18)$$

for almost all $(x, t) \in Q_T$.

4.2.3 Second Order Optimality Conditions

The cubic nonlinearity (4.4) makes the minimization problem (4.1) non-convex. Hence, the fulfillment of the first order optimality conditions is not sufficient for proving global optimality. We need to check the SSCs related to certain critical cones. These critical cones must be chosen as small as possible. In our OCP (4.1) the presence of the Tikhonov regularization parameter λ is crucial for the selection of the critical cones. The standard second-order optimality conditions are not valid for vanishing Tikhonov parameter $\lambda = 0$. For sparse optimal control governed by classical FHNE, new critical cones are presented [24] for the second order optimality conditions. In addition, second order conditions is established for a simple equation, but with more general nonlinear functions in [26]. Therefore, the second order conditions with suitable chosen critical cones are needed in order to prove the stability of locally optimal solutions with respect to $\lambda \rightarrow 0$. Here, we use the following critical cones introduced in [24]

$$C_u = \left\{ v \in L^2(Q_T) : v \text{ satisfies the sign condition and } I'(u)v + \mu j'(u) = 0 \right\}$$

with satisfying the sign condition

$$\begin{cases} v(x, t) \geq 0, & \text{if } u(x, t) = u_a, \\ v(x, t) \leq 0, & \text{if } u(x, t) = u_b. \end{cases}$$

Note that, the set of critical cones C_u is a convex and closed cone in $L^2(Q_T)$. Besides, a local minima for our OCP (4.1) satisfies the following inequalities [24, Theorem 3.3]

$$I''(u)v^2 \geq 0 \quad \forall v \in C_u \setminus \{0\}, \quad \text{equivalently } I''(u)v^2 \geq \delta \|v\|_{L^2(Q_T)}^2 \quad \forall v \in C_u, \quad \delta > 0. \quad (4.19)$$

Then, under the assumption $I''(u)v^2 \geq 0 \quad \forall v \in C_u \setminus \{0\}$, there exists $\delta > 0$ and $r_0 > 0$ such that

$$J(u) + \frac{\delta}{2} \|v - u\|_{L^2(Q_T)}^2 \leq J(v) \quad \forall u \in U_{ad} \cap B_{r_0}(u), \quad (4.20)$$

where $B_{r_0}(u)$ is the $L^2(Q_T)$ ball centered at u with radius r_0 . This shows the existence of a local minima. For a detail information, one can see [24, Theorem 3.4].

Now we compute the SSCs (4.19) to measure the distance a numerical solution u_h of a discretization (4.1) and a local minima u . However, the fulfillment of the SSCs is complicated since it requires the solution of the infinite dimensional problem. Nonetheless, there exist some numerical studies on the SSCs [49, 64].

Here, we compute the smallest eigenvalue of the reduced Hessian and determine the constant δ [64] even though it is not reliable in estimating the constant δ for the infinite dimensional optimal control problem.

Theorem 4.3. *Suppose that a local minima of (4.1) u satisfies the SSCs (4.19) and (4.20). Then the discrete solution u_h satisfying $u_h \in B_{r_0}(u)$ implies*

$$\|u - u_h\|_{L^2(Q_T)} \leq \frac{1}{\delta'} \|\zeta\|_{L^2(Q_T)} \quad (4.21)$$

with defined perturbation function ζ

$$\zeta(x) := \begin{cases} -\min\{0, \lambda u_h + p_{1h} + \mu \mu_h\}, & \text{if } u_h = u_l, \\ -(\lambda u_h + p_{1h} + \mu \lambda_h), & \text{if } u_l < u_h < u_r, \\ -\max\{0, \lambda u_h + p_{1h} + \mu \mu_h\}, & \text{if } u_h = u_r, \end{cases}$$

where p_{1h} , μ_h are corresponding adjoint and sparse solutions to u_h , and $0 < \delta' < \delta$.

Proof. Let (u_h, p_{1h}, μ_h) be discrete solutions, which need not to be optimal, for the OCP (4.1). If u_h is optimal solution, then

$$p_{1h} + \lambda u_h + \mu \mu_h = 0$$

holds for almost all points $x \in \Omega$ with $u_l \leq u_h \leq u_r$. If not, then

$$p_{1h} + \lambda u_h + \mu \mu_h + \zeta = 0,$$

where ζ is a perturbation function adapted from [7]. So, when u_h is not optimal solution of (4.1), it is optimal solution for the following perturbed minimization problem

$$\min_{u \in U_{ad}} J(u) + (\zeta, u)_{L^2(Q_T)}.$$

Then, both the local minima u and discrete solution u_h of (4.1) have to satisfy variational inequality for optimality system as

$$\begin{aligned} (J'(u_h) + \zeta, u - u_h) &\geq 0, \\ (J'(u) + \zeta, u_h - u) &\geq 0. \end{aligned}$$

Adding these inequalities leads

$$\left(J'(u_h) - J'(u), u - u_h \right) + (\zeta, u - u_h) \geq 0.$$

Using mean value theorem implies

$$\left(J'(u_h) - J'(u), u - u_h \right) = -J''(\hat{u})(u - u_h)^2,$$

for some $\hat{u} \in \{v \in U_{ad} : v = u + t(u_h - u), t \in (0, 1)\}$. Inserting this to above inequality and using Cauchy-Schwarz inequality lead the following desired result

$$\delta' \|u_h - u\|_{L^2(Q_T)}^2 \leq \|\zeta\|_{L^2(Q_T)} \|u_h - u\|_{L^2(Q_T)}.$$

□

Now, we can select $\delta' := \delta/2$ as in [49], which produces the following estimate

$$\|u - u_h\|_{L^2(Q_T)} \leq \frac{2}{\delta} \|\zeta\|_{L^2(Q_T)}, \quad (4.22)$$

where u_h belongs to the neighborhood of u .

4.3 Space-Time Discretization of the Optimality System

In this section, we give the space-time discrete forms of the state and adjoint equations by the SIPG in space and the backward Euler discretization in time.

4.3.1 Space Discretization of the State and Adjoint Equations

Before giving SIPG semi-discretized system of state (4.12) and adjoint equations (4.13), we define the discontinuous discrete spaces as follows:

$$W_h = \{w \in L^2(\Omega) \mid w|_K \in \mathbb{P}^1(K), \quad \forall K \in \mathcal{T}_h\}, \quad (4.23)$$

where $\mathbb{P}^1(K)$ is the set of piecewise linear polynomials defined on an element K . Then, SIPG discretized forms of the state and adjoint equations read as: $\forall w \in W_h$ and a.e. $t \in (0, T]$

$$\begin{aligned} \left(\frac{du_{1h}}{dt}, w \right) + d_1 a_h(u_{1h}, w) + b_h(u_{1h}, w) + (f(u_{1h}), w) + (u_{2h}, w) &= \ell_{h,u_1}(w) + (u_h, w), \\ (u_{1h}(\cdot, 0), w) &= (u_{10}, w), \\ \left(\frac{du_{2h}}{dt}, w \right) + d_2 a_h(u_{2h}, w) + b_h(u_{2h}, w) + \epsilon(u_{2h}, w) - \epsilon c_3(u_{1h}, w) &= \ell_{h,u_2}(w), \\ (u_{2h}(\cdot, 0), w) &= (u_{20}, w), \end{aligned} \quad (4.24)$$

$$\begin{aligned}
-\left(\frac{dp_{1h}}{dt}, w\right) + d_1 a_{h,p_1}(p_{1h}, w) - b(p_{1h}, w) + (f_{u_{1h}} p_{1h}, w) - \epsilon c_3(p_{2h}, w) &= 0, \\
(p_{1h}(\cdot, T), w) &= (u_1 - u_{1T}, w), \\
-\left(\frac{dp_{2h}}{dt}, w\right) + d_2 a_{h,p_2}(p_{2h}, w) - b(p_{2h}, w) + \epsilon(p_{2h}, w) + (p_{1h}, w) &= 0, \\
(p_{2h}(\cdot, T), w) &= (u_2 - u_{2T}, w),
\end{aligned} \tag{4.25}$$

where the bilinear terms are defined for $i = 1, 2$ as

$$\begin{aligned}
a_h(u_{ih}, w) &= \sum_{K \in \mathcal{T}_h} \int_K \nabla u_{ih} \cdot \nabla w \, dx \\
&\quad - \sum_{E \in \mathcal{E}_h^0 \cup \mathcal{E}_h^D} \int_E \left(\{\nabla u_{ih}\} \cdot [w] + \{\nabla w\} \cdot [u_{ih}] \right) ds \\
&\quad + \sum_{E \in \mathcal{E}_h^0 \cup \mathcal{E}_h^D} \frac{\gamma}{h_E} \int_E [u_{ih}] \cdot [w] \, ds, \\
b_h(u_{ih}, w) &= \sum_{K \in \mathcal{T}_h} \int_K \mathbf{V} \cdot \nabla u_{ih} w \, dx \\
&\quad + \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \setminus \partial \Omega} \mathbf{V} \cdot \mathbf{n} (u_{ih}^e - u_{ih}) w \, ds - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \mathbf{V} \cdot \mathbf{n} u_{ih} w \, ds, \\
\ell_{h,i}(w) &= \sum_{E \in \mathcal{E}_h^D} \int_E \eta_D \left(\frac{\gamma d_i}{h_E} \mathbf{n} \cdot [w] - \{d_i \nabla w\} \right) ds \\
&\quad - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \mathbf{V} \cdot \mathbf{n} u_{iD} w \, ds, \\
f(u_{ih}, w) &= \sum_{K \in \mathcal{T}_h} \int_K f(u_{ih}) w \, dx, \quad f_{u_{ih}}(u_{ih}, w) = \sum_{K \in \mathcal{T}_h} \int_K f_{u_{ih}}(u_{ih}) u_{ih} w \, dx,
\end{aligned} \tag{4.26}$$

where γ is the penalty parameter. The bilinear forms for the adjoint equations $a_{h,p_{1h}}$ and $a_{h,p_{2h}}$ in (4.25) are similar to the ones for states a_h in (4.24), but contain extra terms, the contribution of the mixed boundary conditions as

$$\sum_{E \in \mathcal{E}_h^N} \int_E (\mathbf{V} \cdot \mathbf{n}) p_{1h} w \, ds \quad \text{and} \quad \sum_{E \in \mathcal{E}_h^N} \int_E (\mathbf{V} \cdot \mathbf{n}) p_{2h} w \, ds.$$

The solutions of (4.24) are given by

$$u_{1h}(t) = \sum_{i=1}^{n_e} \sum_{j=1}^{n_k} u_{1j}^i \phi_j^i, \quad u_{2h}(t) = \sum_{i=1}^{n_e} \sum_{j=1}^{n_k} u_{2j}^i \phi_j^i, \quad \text{and} \quad u_h(t) = \sum_{i=1}^{n_e} \sum_{j=1}^{n_k} u_j^i \phi_j^i, \tag{4.27}$$

with

$$\begin{aligned}\vec{u}_1 &= (u_{11}^1, \dots, u_{1n_k}^1, \dots, u_{11}^{n_e}, \dots, u_{1n_k}^{n_e}) \in \mathbb{R}^N, \\ \vec{u}_2 &= (u_{21}^1, \dots, u_{2n_k}^1, \dots, u_{21}^{n_e}, \dots, u_{2n_k}^{n_e}) \in \mathbb{R}^N, \\ \vec{u} &= (u_1^1, \dots, u_{n_k}^1, \dots, u_1^{n_e}, \dots, u_{n_k}^{n_e}) \in \mathbb{R}^N,\end{aligned}$$

where $N = n_e \times n_k$.

Then, we obtain the SIPG semi-discretized system of state equations (4.24) in form of ODEs as following

$$\begin{aligned}\mathbf{M} \frac{d\vec{u}_1}{dt} + d_1 \mathbf{S} \vec{u}_1 + \mathbf{B} \vec{u}_1 + \mathbf{F}(\vec{u}_1) + \mathbf{M} \vec{u}_2 &= \ell_{u_1} + \mathbf{M} \vec{u}, \\ \mathbf{M} \frac{d\vec{u}_2}{dt} + d_2 \mathbf{S} \vec{u}_2 + \mathbf{B} \vec{u}_2 + \epsilon \mathbf{M} \vec{u}_2 - \epsilon c_3 \mathbf{M} \vec{u}_1 &= \ell_{u_2},\end{aligned}\tag{4.28}$$

where \mathbf{M} is the mass matrix, \mathbf{S} and \mathbf{B} are symmetric and non-symmetric parts of the stiffness matrices corresponding to the diffusive and convective terms, $\mathbf{F}(\vec{u}_1)$ is the nonlinear vector, and ℓ_{u_1} and ℓ_{u_2} are the vectors corresponding to the linear forms ℓ_{h,u_1} and ℓ_{h,u_2} , respectively. In a similar setting, we obtain the SIPG semi-discretized system of the adjoint equations (4.25) as following

$$\begin{aligned}-\mathbf{M} \frac{d\vec{p}_1}{dt} + d_1 \mathbf{S} \vec{p}_1 - \mathbf{B} \vec{p}_1 + \mathbf{F}_{u_1} \vec{p}_1 - \epsilon c_3 \mathbf{M} \vec{p}_2 &= 0, \\ -\mathbf{M} \frac{d\vec{p}_2}{dt} + d_2 \mathbf{S} \vec{p}_2 - \mathbf{B} \vec{p}_2 + \epsilon \mathbf{M} \vec{p}_2 + \mathbf{M} \vec{p}_1 &= 0,\end{aligned}\tag{4.29}$$

where the matrix \mathbf{F}_{u_1} is related to the form $f_{u_1}(p_{1h}, w)$ which is defined in (4.3.1).

4.3.2 Fully Discrete Form of State and Adjoint Equations

We split the time intervals $[0, T]$ into J equally-length $\Delta t = t_n - t_{n-1}$ for $n = 1, 2, \dots, J$ subintervals as $0 = t_0 < t_1 < \dots < t_J = T$. Then, an application of the backward Euler method to (4.28) and (4.29) leads the following fully discrete system

$$\begin{aligned}\frac{1}{\Delta t} \mathbf{M} (\vec{u}_1^{n+1} - \vec{u}_1^n) + d_1 \mathbf{S} \vec{u}_1^{n+1} + \mathbf{B} \vec{u}_1^{n+1} + \mathbf{F}(\vec{u}_1^{n+1}) + \mathbf{M} \vec{u}_2^{n+1} &= \ell_{u_1}^{n+1} + \mathbf{M} \vec{u}_1^{n+1}, \\ \frac{1}{\Delta t} \mathbf{M} (\vec{u}_2^{n+1} - \vec{u}_2^n) + d_2 \mathbf{S} \vec{u}_2^{n+1} + \mathbf{B} \vec{u}_2^{n+1} + \epsilon \mathbf{M} \vec{u}_2^{n+1} - \epsilon c_3 \mathbf{M} \vec{u}_1^{n+1} &= \ell_{u_2}^{n+1}, \\ (\mathbf{M} \vec{u}_1^0)_i &= (u_{10}, \phi_i), \quad (\mathbf{M} \vec{u}_2^0)_i = (u_{20}, \phi_i), \quad i = 1, 2, \dots, N,\end{aligned}\tag{4.30}$$

for $n = 1, 2, \dots, J-1$, and

$$\begin{aligned}\frac{1}{\Delta t} \mathbf{M} (\vec{p}_1^n - \vec{p}_1^{n+1}) + d_1 \mathbf{S} \vec{p}_1^n - \mathbf{B} \vec{p}_1^n + \mathbf{F}_{u_1}(\vec{u}_1^{n+1}) \vec{p}_1^n - \epsilon c_3 \mathbf{M} \vec{p}_2^n &= 0, \\ \frac{1}{\Delta t} \mathbf{M} (\vec{p}_2^n - \vec{p}_2^{n+1}) + d_2 \mathbf{S} \vec{p}_2^n - \mathbf{B} \vec{p}_2^n + \epsilon \mathbf{M} \vec{p}_2^n + \mathbf{M} \vec{p}_1^n &= 0, \\ (\mathbf{M} \vec{p}_1^J)_i &= (\mathbf{M} \vec{u}_{1J})_i - (u_{1T}, \phi_i), \quad (\mathbf{M} \vec{p}_2^J)_i = (\mathbf{M} \vec{u}_{2J})_i - (u_{2T}, \phi_i) \quad i = 1, 2, \dots, N,\end{aligned}\tag{4.31}$$

for $n = J - 1, \dots, 2, 1$.

Although the full discrete state equation (4.30) is solved forward in time, the full discrete adjoint equation (4.31) is solved backward in time.

4.4 Reduced Order Optimal Control

We apply three MOR techniques: POD, POD-DEIM, POD-DMD, introduced in Chapter 3 to find the sub-optimal solutions of the full order OCP problem. The POD was successfully applied for distributed and boundary optimal control for classical FHNE without using nonlinear reduction techniques in [21, 78].

4.4.1 POD Approach for the State and Adjoint System

The reduced system of the state equations is given by

$$\begin{aligned} \frac{d}{dt} \tilde{u}_1 + d_1 \tilde{\mathbf{S}}_{u_1} \tilde{u}_1 + \tilde{\mathbf{B}}_{u_1} \tilde{u}_1 + \Psi_{u_1}^T \mathbf{F}(\Psi_{u_1} \tilde{u}_1) + \tilde{\mathbf{M}}_{u_1, u_2} \tilde{u}_2 &= \Psi_{u_1}^T \ell_{u_1} + \Psi_{u_1}^T \mathbf{M} u, \\ \frac{d}{dt} \tilde{u}_2 + d_2 \tilde{\mathbf{S}}_{u_2} \tilde{u}_2 + \tilde{\mathbf{B}}_{u_2} \tilde{u}_2 + \epsilon \tilde{\mathbf{M}}_{u_2} \tilde{u}_2 - \epsilon c_3 \tilde{\mathbf{M}}_{u_2, u_1} \tilde{u}_1 &= \Psi_{u_2}^T \ell_{u_2}, \\ \tilde{u}_1(0) &= \Psi_{u_1}^T \mathbf{M} \vec{u}_{10}, \quad \tilde{u}_2(0) = \Psi_{u_2}^T \mathbf{M} \vec{u}_{20}, \end{aligned} \quad (4.32)$$

where

$$\begin{aligned} \tilde{\mathbf{S}}_{u_1} &= \Psi_{u_1}^T \mathbf{S} \Psi_{u_1}, & \tilde{\mathbf{S}}_{u_2} &= \Psi_{u_2}^T \mathbf{S} \Psi_{u_2}, & \tilde{\mathbf{B}}_{u_1} &= \Psi_{u_1}^T \mathbf{B} \Psi_{u_1}, & \tilde{\mathbf{B}}_{u_2} &= \Psi_{u_2}^T \mathbf{B} \Psi_{u_2}, \\ \tilde{\mathbf{M}}_{u_2} &= \Psi_{u_2}^T \mathbf{M} \Psi_{u_2}, & \tilde{\mathbf{M}}_{u_2, u_1} &= \Psi_{u_2}^T \mathbf{M} \Psi_{u_1}, & \tilde{\mathbf{M}}_{u_1, u_2} &= \Psi_{u_1}^T \mathbf{M} \Psi_{u_2}. \end{aligned}$$

The reduced system for the adjoint equation (4.31) can be constructed either by constructing its own basis or we can use the same POD basis for the state equations. Different basis for adjoint and state equations does not improve the accuracy of the sub-optimal controls [49], which increases only the computational cost. Therefore, we use the same POD basis functions for the state and adjoint equations. The reduced system for the adjoint equation (4.31) is given by

$$\begin{aligned} -\frac{d}{dt} \tilde{p}_1 + d_1 \tilde{\mathbf{S}}_{u_1} \tilde{p}_1 - \tilde{\mathbf{B}}_{u_1} \tilde{p}_1 + \tilde{\mathbf{M}}_{u_1} \mathbf{F}_{u_1}(\vec{u}_1) \tilde{p}_1 - \epsilon c_3 \tilde{\mathbf{M}}_{u_1, u_2} \tilde{p}_2 &= 0, \\ -\frac{d}{dt} \tilde{p}_2 + d_2 \tilde{\mathbf{S}}_{u_2} \tilde{p}_2 - \tilde{\mathbf{B}}_{u_2} \tilde{p}_2 + \epsilon \tilde{\mathbf{M}}_{u_2} \tilde{p}_2 + \tilde{\mathbf{M}}_{u_2, u_1} \tilde{p}_1 &= 0, \\ \tilde{p}_1(T) &= \Psi_{u_1}^T \mathbf{M} \vec{p}_{1T}, \quad \tilde{p}_2(0) = \Psi_{u_2}^T \mathbf{M} \vec{p}_{2T}, \end{aligned} \quad (4.33)$$

in addition with the reduced matrix $\tilde{\mathbf{M}}_{u_1} = \Psi_{u_1}^T \mathbf{M} \Psi_{u_1}$. As it is seen, the reduced system of ODE's for adjoint equation (4.33) are solved backward in time.

Although we reduce the dimension of the state and adjoint equations (4.30)-(4.31), the dimension of the nonlinear term still depends on the dimension N of the FOM. We approximate the reduced nonlinear terms using the DEIM and the DMD in the next subsections.

4.4.2 Discrete Empirical Interpolation

The ROM for the state equation is given by

$$\begin{aligned} \frac{d}{dt} \tilde{u}_1 + d_1 \tilde{\mathbf{S}}_{u_1} \tilde{u}_1 + \tilde{\mathbf{B}}_{u_1} \tilde{u}_1 + \tilde{\mathbf{F}} + \tilde{\mathbf{M}}_{u_1, u_2} \tilde{u}_2 &= \Psi_{u_1}^T \ell_{u_1} + \Psi_{u_1}^T \mathbf{M}u, \\ \frac{d}{dt} \tilde{u}_2 + d_2 \tilde{\mathbf{S}}_{u_2} \tilde{u}_2 + \tilde{\mathbf{B}}_{u_2} \tilde{u}_2 + \epsilon \tilde{\mathbf{M}}_{u_2} \tilde{u}_2 - \epsilon c_3 \tilde{\mathbf{M}}_{u_2, u_1} \tilde{u}_1 &= \Psi_{u_2}^T \ell_{u_2} \end{aligned} \quad (4.34)$$

with reduced dimension m of nonlinearity $\tilde{\mathbf{F}} = \Psi_u^T W (P^T W)^{-1} P^T F(\Psi_{u_1} \tilde{u}_1)$.

The algorithm for POD-DEIM for the OCP (4.1) is outlined below.

Algorithm 6 POD-DEIM Optimization Algorithm for Reduced Order OCP

Input: u^0 , To1_1 , To1_2 (Stopping criteria tolerances for the optimization algorithm), k (Number of POD basis functions), m (Number of DEIM basis functions)

Output: \vec{u} , \vec{u}_1 , \vec{u}_2

- 1: Solve the full order optimal control problem and save the solutions.
- 2: Collect snapshot data for state variables

$$(\vec{u}_{1i}, \vec{u}_{2i}) = (\vec{u}_1(t_i), \vec{u}_2(t_i)),$$

for $t_i = i\Delta t$, $i = 0, \dots, J$.

- 3: Construct the snapshot matrices U_1, U_2 for state variables $\{\vec{u}_1, \vec{u}_2\}$.
 - 4: Compute POD basis functions $\{\psi_{u_1, i}, \psi_{u_2, i}\}_{i=1}^k$ using snapshot matrices.
 - 5: Find the number of POD basis functions capturing 99.99% of the snapshot energy (3.11).
 - 6: **if** the state equation is nonlinear **then**
 - 7: compute a new basis of order m for the nonlinear term using DEIM.
 - 8: **end if**
 - 9: Solve the reduced order optimal control problem using projected nonlinear conjugate gradient (CG) algorithm [40, 41].
-

4.4.3 Dynamic Mode Decomposition

The ROM state equation (4.28) is given by

$$\begin{aligned} \frac{d}{dt} \tilde{u}_1 + d_1 \tilde{\mathbf{S}}_{u_1} \tilde{u}_1 + \tilde{\mathbf{B}}_{u_1} \tilde{u}_1 + \Psi_{u_1}^T \mathbf{F}^{\text{DMD}}(\vec{u}_1) + \tilde{\mathbf{M}}_{u_1, u_2} \tilde{u}_2 &= \Psi_{u_1}^T \ell_{u_1} + \Psi_{u_1}^T \mathbf{M}u, \\ \frac{d}{dt} \tilde{u}_2 + d_2 \tilde{\mathbf{S}}_{u_2} \tilde{u}_2 + \tilde{\mathbf{B}}_{u_2} \tilde{u}_2 + \epsilon \tilde{\mathbf{M}}_{u_2} \tilde{u}_2 - \epsilon c_3 \tilde{\mathbf{M}}_{u_2, u_1} \tilde{u}_1 &= \Psi_{u_2}^T \ell_{u_2} \end{aligned} \quad (4.35)$$

with reduced dimension r of nonlinearity $\mathbf{F}^{\text{DMD}}(\vec{u}_1)$

$$\mathbf{F}^{\text{DMD}}(\vec{u}_1) = \Psi^{\text{DMD}} \text{diag}(\exp(w^{\text{DMD}}t))b, \quad (4.36)$$

where $b = (\Psi^{\text{DMD}})^\dagger \mathbf{F}((\vec{u}_1)_1)$.

Although the dimension of the system (4.35) can be the same with (4.34) when $k = r$, the main contribution of the system (4.35) is that the DMD makes the system linear, so the OCP becomes convex. Hence, it is not needed to use Newton method. Therefore, the POD-DMD is significantly much faster than the other ROM methods which will be supported with the numerical results in the last Section 4.5.

The algorithm for POD with DMD for the OCP is given below.

Algorithm 7 POD-DMD Optimization Algorithm for Reduced Order OCP

Input: u^0 , Tol_1 , Tol_2 (Stopping criteria tolerances for the optimization algorithm), k (Number of POD basis functions), r (Number of DMD basis functions)

Output: \vec{u} , \vec{u}_1 , \vec{u}_2

- 1: Solve the full order optimal control problem and save the solutions.
- 2: Collect snapshot data for state variables

$$(\vec{u}_{1i}, \vec{u}_{2i}) = (\vec{u}_1(t_i), \vec{u}_2(t_i)),$$

for $t_i = i\Delta t$, $i = 0, \dots, J$.

- 3: Construct the snapshot matrices U_1, U_2 for state variables $\{\vec{u}_1, \vec{u}_2\}$.
 - 4: Compute POD basis functions $\{\psi_{u_1,i}, \psi_{u_2,i}\}_{i=1}^k$ using snapshot matrices.
 - 5: Find the number of POD basis functions capturing at least 99.99% of the snapshot energy (3.11).
 - 6: **if** the state equation is nonlinear **then**
 - 7: compute DMD modes according to Algorithm 3.
 - 8: **end if**
 - 9: Calculate reduced order approximation of the nonlinear term in terms of DMD modes.
 - 10: Solve the reduced order optimal control problem using projected nonlinear CG algorithm [40, 41].
-

4.5 Numerical Results

In this section, we demonstrate numerical results related to OCP (4.1) governed by the convective FHNE in the whole space-time domain and terminal time domain, and also by the diffusive FHNE at the terminal time and in the whole-space time domain with or without sparse controls. We show the effectiveness of the POD-DEIM and the POD-DMD in contrast to the POD in terms of accuracy and speed up's factors for solving our OCPs. The projected nonlinear CG method [40] is used to solve the discrete optimality system. This method was successfully applied to the OCP of Schlögl and classical FHNEs in [21, 23, 24].

Algorithm 8 Nonlinear CG Algorithm

Input: Initial control guess u_0 , initial step size s_0 and stopping criteria tolerances Tol_1 and Tol_2 .

Output: Optimal states \vec{u}_1, \vec{u}_2 , optimal control u , and optimal objective value of $J(u)$.

1: Compute

- initial state variables: $(u_1^0, u_2^0) = (u_{1u_0}, u_{2u_0})$,
- initial adjoint variables: $(p_1^0, p_2^0) = (p_{1u_1^0, u_2^0}, p_{2u_1^0, u_2^0})$,
- sub-gradient of j : $\mu_0 = \mu_{u_0, p_1^0}$,
- sub-gradient of J : $g_0 = \lambda u_0 + p_0 + \mu \mu_0$,
- anti-sub-gradient of J : $d_0 = -g_0$.

Set $k := 0$.

2: Update new gradient

- control: $u_{k+1} = u_k + s_k d_k$,
- state variables: $(u_1^{k+1}, u_2^{k+1}) = (u_{1u_{k+1}}, u_{2u_{k+1}})$,
- adjoint variables $(p_1^{k+1}, p_2^{k+1}) = (p_{1u_1^{k+1}, u_2^{k+1}}, p_{2u_1^{k+1}, u_2^{k+1}})$,
- sub-gradient of j : $\mu_{k+1} = \mu_{u_{k+1}, p_1^{k+1}}$,
- sub-gradient of J : $g_{k+1} = \lambda u_{k+1} + p_1^{k+1} + \mu \mu_{k+1}$.

3: Stop if $\|g_{k+1}\| < \text{Tol}_1$ or $\|J_{k+1} - J_k\| \leq \text{Tol}_2$.

4: Compute the conjugate direction β_{k+1} according to one of the update formulas such as Hestenes-Stiefel, Polak-Ribiere, Fletcher-Reeves, and Hager-Zhang, see e.g., [21, 40] for details.

$$d_{k+1} = -g_{k+1} + \beta_{k+1} d_k.$$

5: Select step size s_{k+1} according to some standard options such as bisection, strong Wolfe-Powell, see e.g., [21, 41] for details. Set $k := k + 1$ and go to Step 2.

In order to compute of the reduced Hessian for FOM of OCPs governed by the convective FHNE, we use the BFGS algorithm [44]:

- Set $H_0 = I$.
- Update for $k = 1, 2, \dots$

$$H_{k+1} = H_k + \frac{q_k q_k^T}{q_k^T r_k} - \frac{(H_k r_k)(H_k r_k)^T}{r_k^T H_k r_k},$$

where $r_k = u_{k+1} - u_k$ and $q_k = g_{k+1} - g_k$.

In all numerical tests, we begin with the underlying control $u = 0$. If it is not specified, the tolerances Tol_1 , Tol_2 for the stopping criteria of the optimization algorithm are chosen as 10^{-3} and 10^{-5} , respectively. The penalty parameter γ in (4.3.1) is chosen as 6, 12 on the interior edges and boundary edges, respectively.

4.5.1 FOM Results for OCP of Convective FHNE

In this subsection, we give the FOM results of the OCP for convective FHNE. If it is not specified, our parameter spaces are chosen as following in all numerical examples

$$c_1 = 9, \quad c_2 = 0.02, \quad c_3 = 5, \quad \epsilon = 0.1, \quad d_1 = d_2 = 1$$

on a rectangular box $\Omega = [0, L] \times [0, H]$ with $L = 100$ and $H = 5$ as in [35]. In addition, the final time is taken as $T = 1$.

4.5.1.1 Optimal Control in the Space-Time Domain

First of all, we consider the OCP (4.1) with the desired state functions defined in the whole space-time domain Q_T with $\omega_Q = 1$, $\omega_T = 0$, and Tikhonov regularization parameter $\lambda = 10^{-5}$. The step sizes are taken as $\Delta x_1 = \Delta x_2 = 0.5$ in spatial, and $\Delta t = 0.05$ in temporal domain. The desired states are chosen as the solution of uncontrolled FHNE

$$u_{1Q}(x, t) = \begin{cases} u_{1\text{nat}}(x, t), & \text{if } t \leq T/2, \\ 0, & \text{otherwise,} \end{cases} \quad u_{2Q}(x, t) = \begin{cases} u_{2\text{nat}}(x, t), & \text{if } t \leq T/2, \\ 0, & \text{otherwise,} \end{cases}$$

with the initial conditions given by

$$u_{10}(x, t) = \begin{cases} 0.1, & \text{if } 0 \leq x_1 \leq 0.1, 0 \leq x_2 \leq H, \\ 0, & \text{otherwise,} \end{cases} \quad u_{20}(x, 0) = 0.$$

Here, $(u_{1\text{nat}}, u_{2\text{nat}})$ shows the uncontrolled solutions of the convective FHNE.

Further, the set of admissible control is chosen as

$$\mathcal{U}_{ad} := \{u \in L^\infty(Q) : -0.2 \leq u(x, t) \leq 0 \text{ for a.e } (x, t) \in Q\}.$$

First, we examine the numerical solutions of our optimization problem (4.1) without sparsity, i.e., $\mu = 0$. As it is seen in Figure 4.1 with increasing values of V_{\max} , the computed solutions of the control u is bounded with our chosen box constraints and the corresponding state u_1 becomes more curved, as expected [35]. Additionally, optimal values of the cost functional J , the number of iterations, line searches and Newton steps in Table 4.1 occurring in the optimization algorithm are increasing with the increasing values of V_{\max} . This is due to the fact that the higher value of V_{\max} makes the character of our minimization problem (4.1) more stiff.

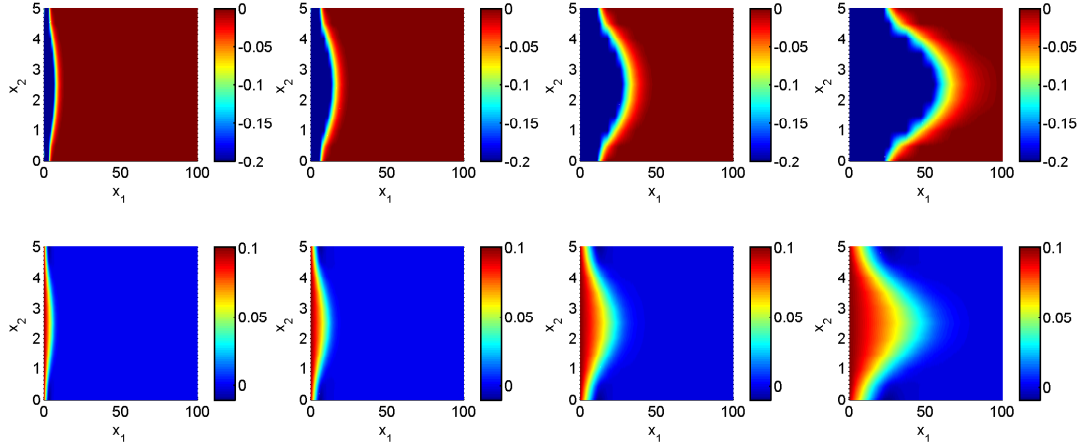


Figure 4.1: Optimal controls u (top) and associated states u_1 (bottom) at $t = 0.75$ without sparse control for $V_{\max}=16, 32, 64, 128$ (from left to right).

Table 4.1: Cost functional J , number of nonlinear CG iterations, line searches, Newton steps without sparse control.

V_{\max}	$J = I$	#ite.	#search	#Newton
16	2.91e-2	64	245	787
32	5.80e-2	77	297	956
64	1.16e-1	96	373	1203
128	2.30e-1	123	481	1554

Secondly, we test the effect of the sparsity parameter μ in our optimization problem (4.1). As it is observed in Figure 4.2, higher values of μ induces the sparsity of the optimal control with their associated states behaviours u_1 . As increasing values of μ , the smooth I and non-smooth μj cost functional increase in Table 4.2. In addition to this, the number nonlinear CG iterations, line searches and Newton iterations increase in Table 4.2 which is expected result since the larger value of μ increases the non-smoothness of the problem (4.1). However as it can be seen in the next terminal control problem, this situation can be just the opposite. That is, the number of nonlinear CG iterations, line searches and Newton iterations can diminish while the sparse parameter is increasing [23].

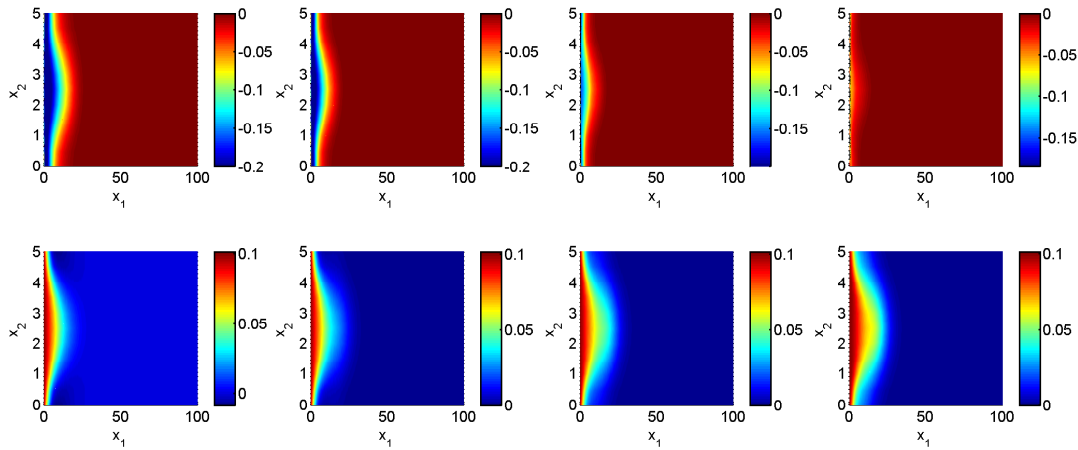


Figure 4.2: Optimal controls u (top) and associated states u_1 (bottom) for the sparsity parameters $\mu = 1/500, 1/100, 1/50, 1/35$ (from left to right) and $V_{\max} = 32$ at $t = 0.75$.

Table 4.2: Optimal values of I , μj , and number of iterations, line searches, and Newton steps for $V_{\max} = 32$.

μ	I_{opt}	μj_{opt}	#ite.	#search	#Newton
1/35	1.10e-0	1.30e-1	170	666	2193
1/50	7.31e-1	8.21e-2	156	610	1971
1/100	3.60e-1	3.47e-2	121	470	1516
1/500	1.21e-1	5.93e-3	88	341	1099
0	5.80e-2	0	77	297	956

Eventually for vanishing Tikhonov parameter λ for $\mu = 0$ and $\mu \neq 0$, the convergence of the optimal control and corresponding state are examined in Figure 4.3, which assumes an essential part in order to check SSCs for OCPs without exact solution [24, 26]. In Figure 4.3, our reference solutions for states and control are $\bar{u}_{1\text{ref}} := u_{11e-10}$, $\bar{u}_{2\text{ref}} := u_{21e-10}$, and $\bar{u}_{\text{ref}} := u_{1e-10}$ with fixed $V_{\max} = 64$. As it is observed in Figure 4.3, for both sparsity parameters $\mu = 0$, $\mu = 1/100$ the errors decay as $\lambda \downarrow 0$.

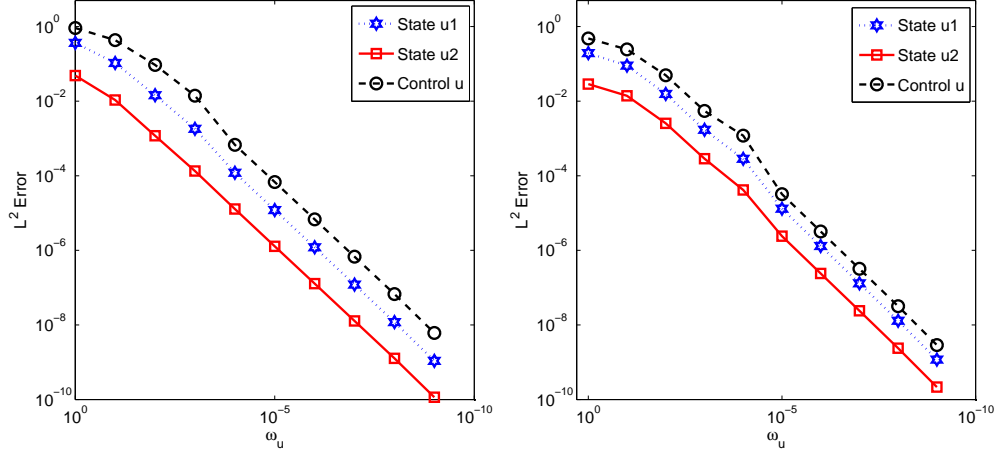


Figure 4.3: $L^2(Q)$ errors for $\|\bar{u}_{1\lambda} - \bar{u}_{1\text{ref}}\|$, $\|\bar{u}_{2\lambda} - \bar{u}_{2\text{ref}}\|$, and $\|\bar{u}_\lambda - \bar{u}_{\text{ref}}\|$ for the sparsity parameter $\mu = 0$ (left), $\mu = 1/100$ (right).

In Table 4.3, $L^2(Q)$ errors for the desired functions are shown. With decreasing values of Tikhonov regularization parameter λ the errors are decreasing. Hence, desired functions are quite achieved with decreasing values of λ .

Table 4.3: $L^2(Q)$ errors for $\|\bar{u}_{1\lambda} - u_{1Q}\|$ and $\|\bar{u}_{2\lambda} - u_{2Q}\|$.

λ	$\mu = 0$		$\mu = 1/100$	
	$\ \bar{u}_{1\lambda} - u_{1Q}\ $	$\ \bar{u}_{2\lambda} - u_{2Q}\ $	$\ \bar{u}_{1\lambda} - u_{1Q}\ $	$\ \bar{u}_{2\lambda} - u_{2Q}\ $
1	5.77e-1	1.09e-1	5.29e-1	9.91e-2
1e - 1	3.84e-1	7.32e-2	4.48e-1	8.48e-2
1e - 2	3.42e-1	6.53e-2	4.02e-1	7.44e-2
1e - 3	3.39e-1	6.47e-2	3.95e-1	7.25e-2
1e - 4	3.39e-1	6.46e-2	3.95e-1	7.23e-2
1e - 5	3.39e-1	6.46e-2	3.95e-1	7.23e-2
1e - 6	3.39e-1	6.46e-2	3.95e-1	7.23e-2
1e - 7	3.39e-1	6.46e-2	3.95e-1	7.23e-2
1e - 8	3.39e-1	6.46e-2	3.95e-1	7.23e-2
1e - 9	3.39e-1	6.46e-2	3.95e-1	7.23e-2
1e - 10	3.39e-1	6.46e-2	3.95e-1	7.23e-2

4.5.1.2 Terminal Control

Here, we consider terminal control, that is, $\omega_Q = 0$, $\omega_T = 1$. The regularization parameter is chosen $\lambda = 10^{-3}$, and we consider the sparse optimal control, i.e., $\mu \neq 0$. The step sizes are $\Delta x_1 = \Delta x_2 = 0.125$ in spatial and $\Delta t = 0.05$ in temporal domain. The desired state functions are constructed as

$$u_{1T}(x, T) = u_{1\text{nat}}(x, T/2) \quad \text{and} \quad u_{2T}(x, T) = u_{2\text{nat}}(x, T/2),$$

where $(u_{1\text{nat}}, u_{2\text{nat}})$ is the solution of the uncontrolled convective FHNE at the final time $T = 1$ with the initial conditions

$$u_{10}(x, t) = \begin{cases} 1, & \text{if } 2 \leq x_1 \leq 2.2, 0 \leq x_2 \leq H, \\ 0, & \text{otherwise,} \end{cases} \quad u_{20}(x, 0) = 0.$$

The admissible set of controls is chosen as

$$\mathcal{U}_{ad} := \{u \in L^\infty(Q) : 0 \leq u(x, t) \leq 0.2 \text{ for a.e. } (x, t) \in Q\}.$$

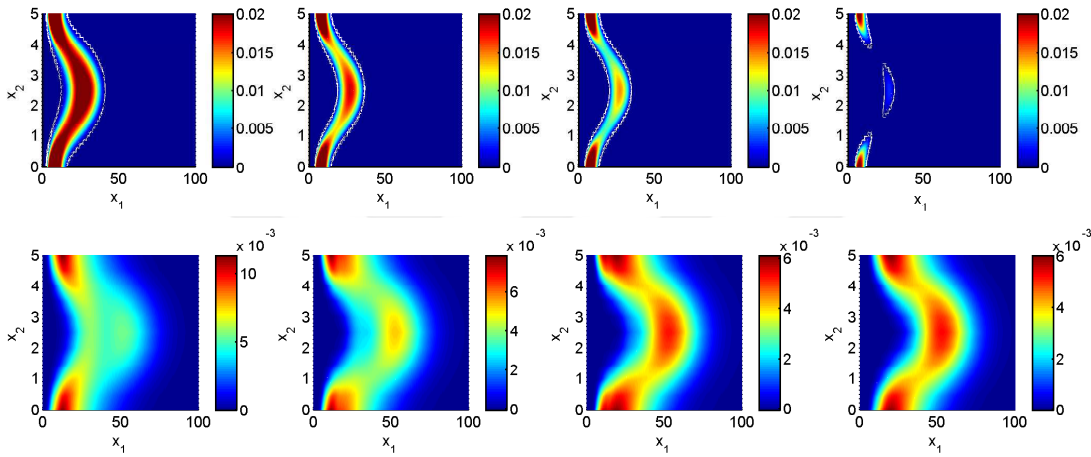


Figure 4.4: Optimal controls u (top) and associated states u_1 (bottom) for the sparse parameters $\mu = 1/500, 1/100, 1/50, 1/35$ (from left to right) and $V_{\max} = 32$ at $t = 0.75$.

Table 4.4: Optimal value of I , μj , and number of nonlinear CG iterations, line searches, and Newton steps for $V_{\max} = 64$.

μ	I_{opt}	μj_{opt}	#ite.	#search	#Newton
1/100	1.37e-1	1.15e-2	2	5	16
1/150	9.24e-2	7.70e-3	3	7	18
1/200	7.00e-2	5.88e-3	4	8	19
1/2000	8.79e-3	6.43e-4	19	74	238
0	1.42e-3	0	22	89	289

While the value of the sparsity parameter μ is increasing, the sparsity is seen more obvious in the optimal control u and corresponding optimal states u_1, u_2 in Figure 4.4. As we said in the previous part, the sparsity parameter μ increases while the number of nonlinear CG iterations, line searches, and Newton steps decrease Table 4.4, similar to the result in [23].

Table 4.5: Numerical errors of $\|u - u_h\|$ and error estimates $\frac{2}{\delta}\|\zeta\|$ for sparse controls with $\mu = 1/200$.

$\Delta x_1 = \Delta x_2$	$\ u - u_h\ $	$\frac{2}{\delta}\ \zeta\ $
2.5	2.43e-2	5.92e-2
1.25	2.43e-2	5.60e-2
0.625	2.46e-2	4.78e-2

In the Table 4.5, the SSCs are checked. Since the exact solution is not known, we consider the reference solution computed with $\Delta x_1 = \Delta x_2 = 0.3125$ to show that the result (4.22) of the Theorem 4.3 is satisfied.

4.5.2 ROM Results for OCP of Convective FHNE

Here, we consider the OCP with desired state functions defined at the final time $T = 1$ only, that is, $\omega_Q = 0$, $\omega_T = 1$. There is no sparse control, i.e. $\mu = 0$, and $\lambda = 10^{-3}$. The maximum wave speed of the velocity field is $V_{max} = 128$. We use a uniform step size in space $\Delta x_1 = \Delta x_2 = 0.5$ and in time $\Delta t = 0.05$. Stopping criteria for this problem is different than before: the criteria $|J_{old} - J| \leq 1e - 5$ is changed with $|J_{old} - J|/|J_{old}| \leq 1e - 3$ for the FOM solutions. Then, the corresponding sub-gradient of J : $g_{k+1} := \text{Tol}_1$ is saved and used as a stopping criteria as $\|g_{k+1}\| < \text{Tol}_1$ for the ROM solutions.

The desired states are chosen as

$$u_{1T}(x, T) = u_{1\text{nat}}(x, T/2) \quad \text{and} \quad u_{2T}(x, T) = u_{2\text{nat}}(x, T/2).$$

Initial conditions are chosen as

$$u_{10}(x, t) = \begin{cases} 1, & \text{if } 2 \leq x_1 \leq 2.2, 0 \leq x_2 \leq H, \\ 0, & \text{otherwise,} \end{cases} \quad u_{20}(x, 0) = 0.$$

The admissible set of controls is chosen as

$$\mathcal{U}_{ad} := \{u \in L^\infty(Q) : -0.01 \leq u(x, t) \leq 0.01 \text{ for a.e. } (x, t) \in Q\}.$$

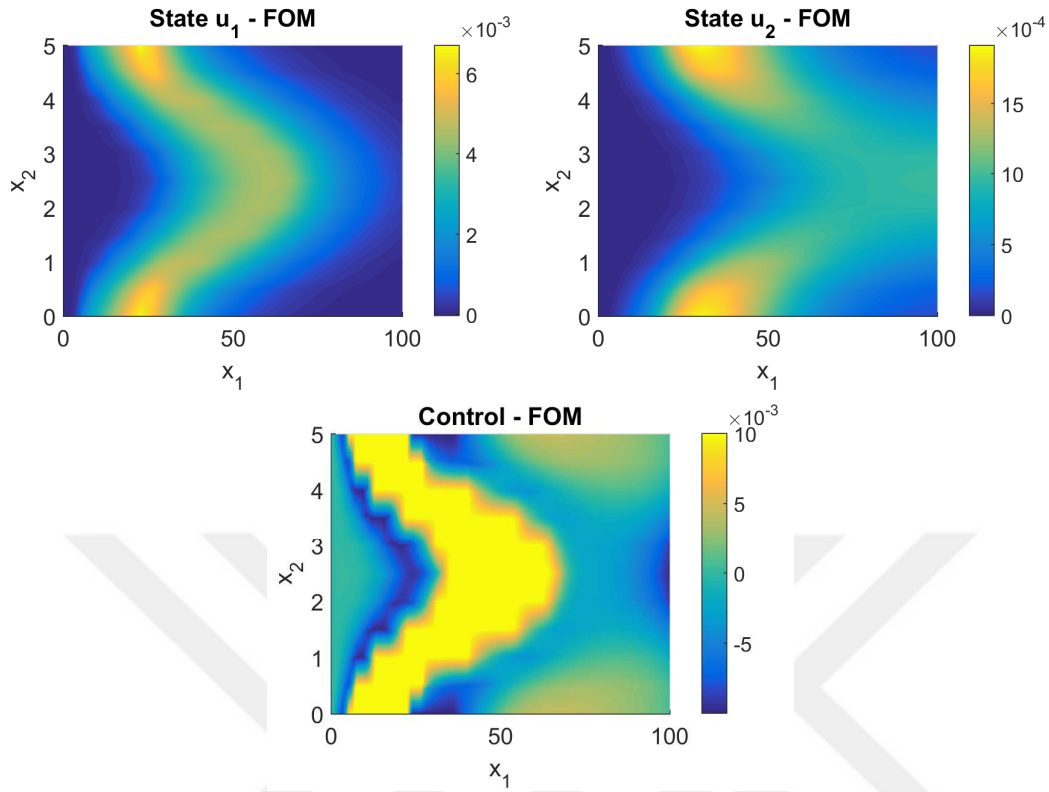


Figure 4.5: Optimal states u_1 and u_2 (top), optimal control (bottom) for the full system at the final time.

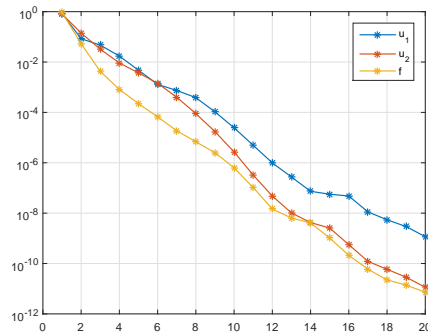


Figure 4.6: Decay of the singular values for u_1, u_2 and the nonlinearity f .

For the FOM, optimal control and corresponding optimal states are given in Figure 4.5, which coincides with the results in [92]. In Figure 4.6, the decay of the singular values of the snapshot matrices for both state variables u_1, u_2 , and the nonlinearity $f(u_1)$ are shown. As it is seen, all of them are linearly decreasing almost with the same rate.

For the ROM solutions to approximate FOM solutions as in Figure 4.5, we rely on the relative information content (3.11). We choose number of POD modes k according to the solution of $\min_k \epsilon(k) \geq 99.99\%$. Here, for this problem k is chosen as 8. We

choose also $m = r = 8$, the number of DEIM and DMD modes. The errors for the ROM solutions for both state components u_1 , u_2 and control u , which are calculated according to the difference between FOM and ROM solutions, are given in Figure 4.7 with almost the same rate.

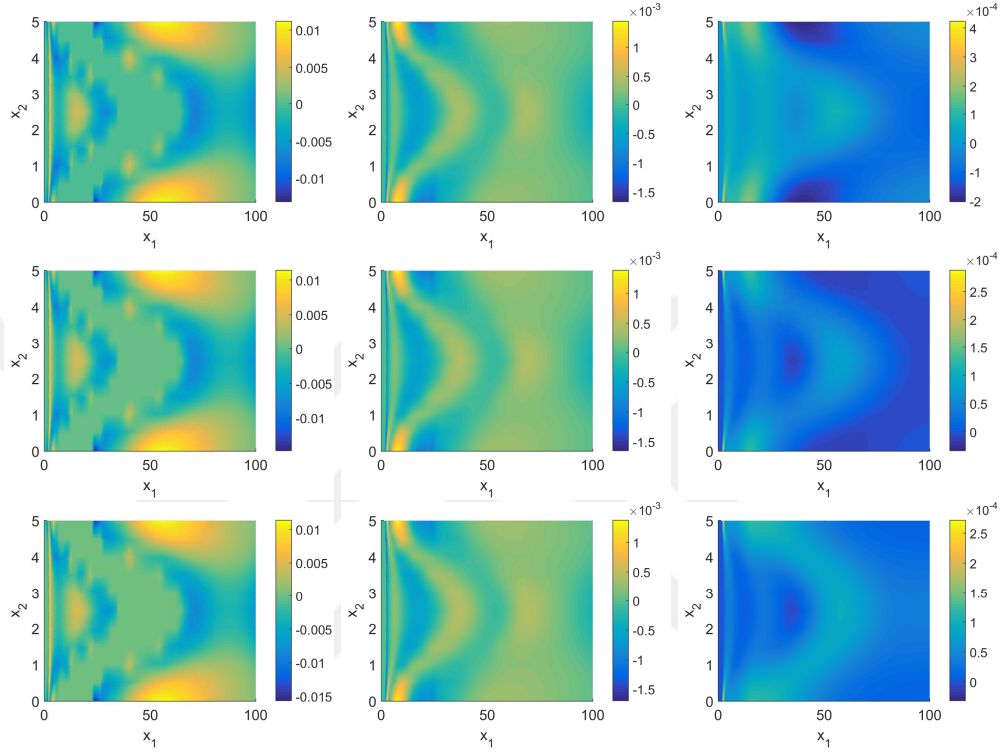


Figure 4.7: Errors of POD, POD-DEIM and POD-DMD with 8 basis functions (from top to bottom): optimal control errors u (left), associated states u_1 (middle), u_2 (right) errors at the final time.

The results in Table 4.6 shows optimal values of the cost functional J , the number of iterations, line searches, Newton steps, CPU times (in seconds) with speed-up factors for FOM and our using ROM techniques. The speed-up factors for ROMs are defined in Chapter 3.

Here, the important contribution in optimization problem with ROM is POD-DMD technique, which is a new approach in optimization problems. There is no literature searches related to this as far as we know. As one can see, the speed-up factor for POD-DMD is really promising, of course, with accepted accuracy and efficiency. This is due to the fact that POD-DMD reduced state system (4.35) is linear which makes the resulting OCP problem (4.1) convex. Instead of using nonlinear CG, we prefer to use linear CG. Therefore, there is no Newton iteration as it is seen in Table 4.6. In addition, the values of the cost functional are really close to each other.

Table 4.6: Optimal value of cost functional J , average number of nonlinear CG iterations, line searches, Newton steps, and speed-up's factors.

	$J_{opt} + \mu j_{opt}$	#ite.	#search	#Newton	CPU (s)	Speed-up
Full	1.72788e-04	54	205	11.91	6994.80	-
POD	2.14017e-04	13	44	9.38	275.80	25.36
POD-DEIM	1.99148e-04	14	48	9.64	171.69	40.74
POD-DMD	1.98408e-04	14	48	-	50.44	138.69

We also check $L^2(Q)$ errors in the whole space-time domain and $L^2(\Omega)$ errors at the final time for the control u and state variables u_1, u_2 with the same increasing number of POD-DEIM-DMD modes in Figures 4.8-4.9. As expected, the POD reduced sub-optimal solutions are the most accurate one. After around $k = m = r = 8$, the whole domain and terminal domain errors, shown in Figures 4.8-4.9, do not change much for the ROMs. Therefore, it is not needed to use greater number of modes.

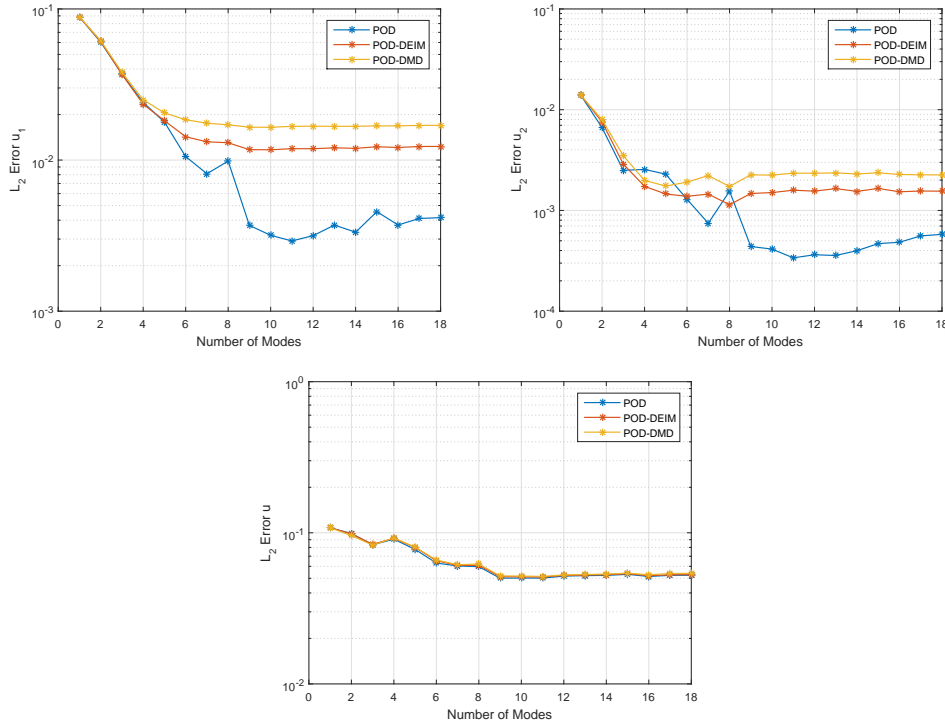


Figure 4.8: Errors for states u_1, u_2 (top), and control u (bottom) in the whole domain with increasing number of modes.

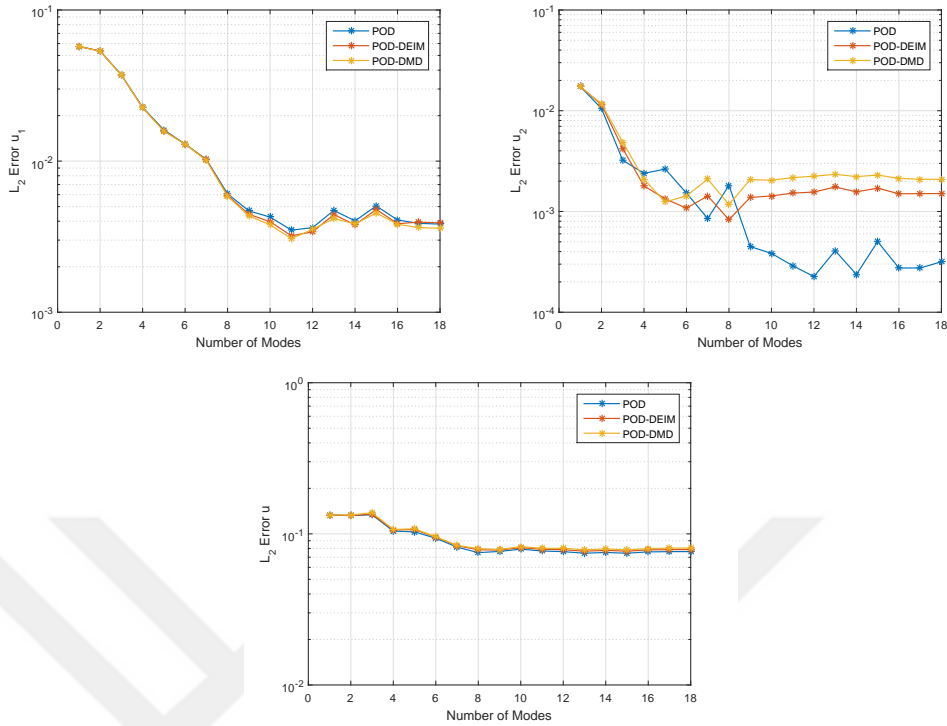


Figure 4.9: Errors for states u_1 , u_2 (top), and control u (bottom) at the final time with increasing number of modes.

In Figure 4.10, we check the cost functional values. After almost 9 basis functions for all ROM techniques, the objective function value does not change much and they all approach the FOM optimal control. Besides, the CPU times are shown in the same Figure 4.10. As expected, the CPU-time for the POD-DMD is always less than the other ROM techniques, and the CPU-time for all three reduced sub-optimal is always less than for the FOM optimal control.

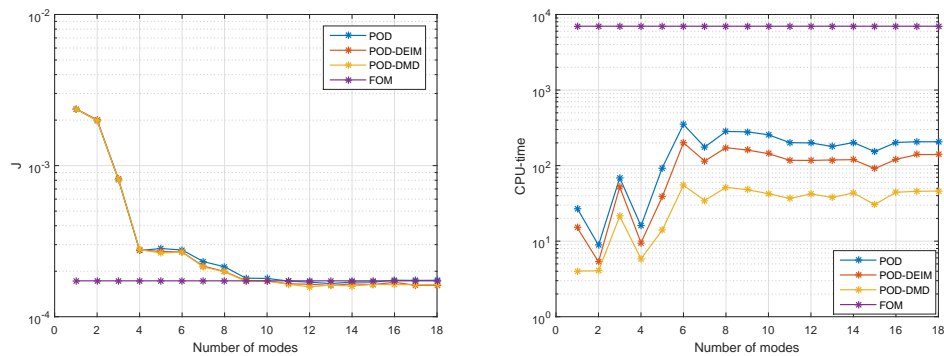


Figure 4.10: Objective function values (left) and CPU times (right) with increasing number of modes.



CHAPTER 5

CONCLUSION AND DISCUSSION

In the first part of the thesis, we have investigated pattern formation by the diffusive and convective FHNEs. We have shown that patterns like pulses, fronts, spots and labyrinths for the diffusive FHNE are computed accurately using the skew gradient structure preserving time integrator AVF in combination with DG finite elements in space. In addition, the wave type solutions of the convective FHNE have been captured in an efficient way with DG method. For convection dominated problems, DG methods are more stable than the continuous finite element methods. Due to the large amount of computing time of the high fidelity DG solutions, we have applied projection based MOR techniques. MOR can dramatically reduce simulation costs by preserving the behavior of parametrized PDEs, which was demonstrated here for the diffusive FHNE with Turing patterns and for the convective FHNE with wave types solutions. The DEIM and DMD are especially used to reduce the dimension of the nonlinear term in the activator equation. In the literature, DEIM is applied usually for the finite differences and the continuous finite elements discretized PDEs. We have shown that the DG discretization in space can save the computational cost and due its local structure DG is more efficient than the continuous finite elements. However, in terms of error convergence, the performance of the nonlinear dimension reduction techniques needs to be improved. For POD-DEIM solutions, increasing the number of DEIM modes leads an oscillation in the error. More accurate solutions can be obtained using exponential time integrators, like the exponential AVF. Pattern formation problems require much more POD mode than the parabolic problems. New approaches are needed to overcome this drawback. For POD-DMD solutions, a drawback of this method is that although the number of DMD modes increases, the resulting error does not decrease beyond some tolerance. This is because the DMD modes are not orthogonal. To overcome this issue, the DMD modes should be orthogonalized.

In the second part of the thesis we solved the OCP governed by the convective FHNE in blood coagulation. Numerical results of the optimal control governed by the convective FHNE model with travelling waves reveal different aspects of the parabolic semi-linear optimal control problems investigated. The second order optimality conditions for local solutions in form of 2D travelling waves are verified numerically for vanishing Tikhonov regularization parameter with sparse controls as done for one dimensional waves of the classical FHNE in the literature. By using the second order optimality conditions, we estimate the measure between the discrete solution and the local minima. Since the control of the convection dominated problems with wave so-

lutions requires a large amount of computing time, the MOR techniques have been applied. Among the three MOR techniques the POD is the most accurate as expected. The POD-DEIM and POD-DMD ROM errors are close, but in contrast to DEIM, POD-DMD ROM results are linear reduced order systems. Therefore in the computation of the reduced optimal control of the convective FHNE, the POD solutions are the most accurate, but the POD-DMD solutions are the fastest. Usually the optimality conditions are not satisfied exactly in the reduced model, therefore the called "reduced optimality system" approach should be applied.

As a future work, in order to compensate of deviations from the optimized trajectory, to diminish unwanted disturbances and errors in the optimal control, it is important to utilize a feedback mechanism like model predictive control. Since this would take more time than solving the optimal control problem, MOR techniques can be applied to speed up this method. As an entirely different subject, to capture some patterns of the reaction-diffusion equations we can also study parameter estimation based on the OCP.

REFERENCES

- [1] T. Akman and B. Karasözen, Variational time discretization methods for optimal control problems governed by diffusion-convection-reaction equations, *Journal of Computational and Applied Mathematics*, 272, pp. 41–56, 2014.
- [2] 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, *Computational Optimization and Applications*, 57, pp. 703–729, 2014.
- [3] A. Alla and J. N. Kutz, Nonlinear model order reduction via dynamic mode decomposition, arXiv:1602.05080 to appear *SIAM Journal Scientific Computing*, 2016.
- [4] A. Alla and J. N. Kutz, Randomized model order reduction, arXiv preprint arXiv:1611.02316, 2016.
- [5] H. Antil, M. Heinkenschloss, and D. C. Sorensen, Application of the discrete empirical interpolation method to reduced order modeling of nonlinear and parametric systems, in *Reduced Order Methods for Modeling and Computational Reduction*, volume 9 of *MS & A - Modeling, Simulation and Applications*, pp. 101–136, Springer International Publishing, 2014.
- [6] A. C. Antoulas and D. C. Sorensen, Approximation of large-scale dynamical systems: An overview, *Applied Mathematics and Computer Science*, 11(5), pp. 1093–1122, 2001.
- [7] N. Arada, E. Casas, and F. Tröltzsch, Error estimates for the numerical approximation of a semilinear elliptic control problem, *Computational Optimization and Applications*, 23(2), pp. 201–229, 2002.
- [8] D. N. Arnold, An interior penalty finite element with discontinuous elements, *SIAM Journal on Numerical Analysis*, 19, pp. 724–760, 1982.
- [9] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous galerkin methods for elliptic problems, *SIAM Journal on Numerical Analysis*, 39(5), pp. 1749–1779, 2002.
- [10] F. I. Ataullakhanov, V. I. Zarnitsina, A. V. Pokhilko, A. I. Lobanov, and O. L. Morozova, Spatio-temporal dynamics of blood coagulation and pattern formation: A theoretical approach, *International Journal of Bifurcation and Chaos*, 12(09), pp. 1985–2002, 2002.

- [11] B. Ayuso and L. D. Marini, Discontinuous Galerkin methods for advection-diffusion-reaction problems, *SIAM Journal on Numerical Analysis*, 47, pp. 1391–1420, 2009.
- [12] I. Babuška and M. Zlámal, Nonconforming elements in the finite element method with penalty, *SIAM Journal on Numerical Analysis*, 10, pp. 45–59, 1973.
- [13] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations, *Comptes Rendus Mathématique*, 339(9), pp. 667–672, 2004.
- [14] W. Barthel, C. John, and F. Tröltzsch, Optimal boundary control of a system of reaction diffusion equations, *ZAMM - Journal of Applied Mathematics and Mechanics*, 90(12), pp. 966–982, 2010, ISSN 1521-4001.
- [15] P. Benner, E. Sachs, and S. Volkwein, Model order reduction for PDE constrained optimization, in G. Leugering, P. Benner, S. Engell, A. Griewank, H. Harbrecht, M. Hinze, R. Rannacher, and S. Ulbrich, editors, *Trends in PDE Constrained Optimization*, pp. 303–326, Springer International Publishing, Cham, 2014.
- [16] W. J. Beyn, D. Otten, and J. Rottmann-Matthes, Stability and computation of dynamic patterns in pdes, in *Current challenges in stability issues for numerical differential equations*, pp. 89–172, Springer, 2014.
- [17] W. J. Beyn, D. Otten, and J. Rottmann-Matthes, *Stability and Computation of Dynamic Patterns in PDEs*, pp. 89–172, Springer International Publishing, Cham, 2014, ISBN 978-3-319-01300-8.
- [18] D. A. Bistrián and I. M. Navon, An improved algorithm for the shallow water equations model reduction: Dynamic mode decomposition vs pod, *International Journal for Numerical Methods in Fluids*, 78(9), pp. 552–580, 2015.
- [19] A. Borzi and R. Griesse, Distributed optimal control of lambda–omega systems, *Journal of Numerical Mathematics*, 14(1), pp. 17–40, 2006.
- [20] F. Brezzi, G. Manzini, D. Marini, P. Pietra, and A. Russo, Discontinuous Galerkin approximations for elliptic problems, *Numerical Methods for Partial Differential Equations*, 16, pp. 365–378, 2000.
- [21] R. Buchholz, H. Engel, E. Kammann, and F. Tröltzsch, On the optimal control of the Schlögl-model, *Computational Optimization and Applications*, 56(1), pp. 153–185, 2013.
- [22] E. Casas, R. Herzog, and G. Wachsmuth, Optimality conditions and error analysis of semilinear elliptic control problems with ℓ_1 cost functional, *SIAM Journal on Optimization*, 22(3), pp. 795–820, 2012.
- [23] E. Casas, C. Ryll, and F. Tröltzsch, Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems, *Computational Methods in Applied Mathematics*, 13(4), pp. 415–442, 2013.

- [24] E. Casas, C. Ryll, and F. Tröltzsch, Second order and stability analysis for optimal sparse control of the FitzHugh-Nagumo equation, *SIAM Journal on Control and Optimization*, 53(4), pp. 2168–2202, 2015.
- [25] E. Casas and F. Tröltzsch, Second order optimality conditions and their role in PDE control, *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 117(1), pp. 3–44, 2015, ISSN 0012-0456.
- [26] E. Casas and F. Tröltzsch, Second-order optimality conditions for weak and strong local solutions of parabolic optimal control problems, *Vietnam Journal of Mathematics*, 44, pp. 181–202, 2016.
- [27] E. Celledoni, V. Grimm, R. I. McLachlan, D. I. McLaren, D. J. O’Neale, B. Owren, and G. R. W. Quispel, Preserving energy resp. dissipation in numerical PDEs using the ”Average Vector Field” method, *Journal of Computational Physics*, 231, pp. 6770–6789, 2012.
- [28] S. Chaturantabut and D. C. Sorensen, Nonlinear model reduction via discrete empirical interpolation, *SIAM Journal on Scientific Computing*, 32(5), pp. 2737–2764, 2010.
- [29] C. N. Chen and Y. S. Choi, Traveling pulse solutions to FitzHugh–Nagumo equations, *Calculus of Variations and Partial Differential Equations*, 54(1), pp. 1–45, 2015.
- [30] C. N. Chen and X. Hu, Stability analysis for standing pulse solutions to FitzHugh-Nagumo equations, *Calculus of Variations and Partial Differential Equations*, 49, pp. 827–845, 2014.
- [31] K. K. Chen, J. H. Tu, and C. W. Rowley, Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses, *Journal of Nonlinear Science*, 22(6), pp. 887–915, 2012.
- [32] B. Cockburn and C. W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, *SIAM Journal on Numerical Analysis*, 35, pp. 2440–2463, 1999.
- [33] J. Douglas and T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, in R. Glowinski and J. L. Lions, editors, *Computing Methods in Applied Sciences*, volume 58 of *Lecture Notes in Phys*, Springer Berlin Heidelberg.
- [34] E. A. Ermakova, M. A. Panteleev, and E. E. Shnol, Blood coagulation and propagation of autowaves in flow, *Pathophysiol Haemost Thromb*, 34, pp. 135–142, 2005.
- [35] E. A. Ermakova, E. E. Shnol, M. A. Panteleev, A. A. Butylin, V. Volpert, and F. I. Ataulkhanov, On propagation of excitation waves in moving media: The FitzHugh-Nagumo model, *PLoSOne*, 4(2), p. E4454, 2009.
- [36] R. FitzHugh, Impulses and physiological states in theoretical models of nerve membrane, *Biophysical Journal*, 1(6), pp. 445 – 466, 1961.

- [37] A. Gierer, Some physical, mathematical and evolutionary aspects of biological pattern formation, *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 295(1078), pp. 429–440, 1981.
- [38] V. Girault, B. Rivière, and M. F. Wheeler, A discontinuous Galerkin method with nonoverlapping domain decomposition for the Stokes and Navier-Stokes problems, *Mathematics of Computation*, 74(249), pp. 53–84 (electronic), 2005.
- [39] T. Gudi and A. K. Pani, Discontinuous Galerkin methods for quasi-linear elliptic problems of nonmonotone type, *SIAM Journal on Numerical Analysis*, 45(1), pp. 163–192, 2007.
- [40] W. W. Hager and H. Zhang, Algorithm 851: CG_DESCENT, a conjugate gradient method with guaranteed descent, *ACM Transactions on Mathematical Software*, 32(1), pp. 113–137, 2006, ISSN 0098-3500.
- [41] W. W. Hager and H. Zhang, A survey of nonlinear conjugate gradient methods, *Pac. J. Optim.*, 2(1), pp. 35–58, 2006.
- [42] E. Hairer and C. Lubich, Energy-diminishing integration of gradient systems, *IMA Journal of Numerical Analysis*, 34(2), pp. 452–461, 2014.
- [43] N. Halko, P. G. Martinsson, and J. A. Tropp, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions, *SIAM review*, 53(2), pp. 217–288, 2011.
- [44] R. Herzog and K. Kunisch, Algorithms for PDE-constrained optimization, *GAMM-Mitteilungen*, 33(2), pp. 163–176, 2010.
- [45] M. Hinze and S. Volkwein, Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition, *Computational Optimization and Applications*, 39(3), pp. 319–345, 2008.
- [46] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, coherent structures, dynamical systems and symmetry*, Cambridge university press, 1998.
- [47] D. E. Jackson, Existence and regularity for the FitzHugh-Nagumo equations with inhomogeneous boundary conditions, *Nonlinear Analysis: Theory, Methods & Applications*, 14(3), pp. 201–216, 1990.
- [48] M. R. Jovanović, P. J. Schmid, and J. W. Nichols, Sparsity-promoting dynamic mode decomposition, *Physics of Fluids*, 26(2), p. 024103, 2014.
- [49] E. Kammann, F. Tröltzsch, and S. Volkwein, A posteriori error estimation for semilinear parabolic optimal control problems with application to model reduction by POD, *ESAIM: Mathematical Modelling and Numerical Analysis*, 47, pp. 555–581, 3 2013, ISSN 1290-3841.
- [50] B. Karasözen, T. Küçükseyhan, and M. Uzunca, Structure preserving integration and model order reduction of skew-gradient reaction-diffusion systems, *Annals of Operations Research*, pp. 1–28, 2015.

- [51] B. Karasözen, M. Uzunca, and T. Küçükseyhan, *Model Order Reduction for Pattern Formation in FitzHugh-Nagumo Equations*, pp. 369–377, Springer International Publishing, Cham, 2016, ISBN 978-3-319-39929-4.
- [52] B. O. Koopman, Hamiltonian Systems and Transformations in Hilbert Space, *Proceedings of the National Academy of Science*, 17, pp. 315–318, 1931.
- [53] K. Kunisch and M. Müller, Uniform convergence of the POD method and applications to optimal control, *Discrete and Continuous Dynamical Systems*, 35(9), pp. 4477–4501, 2015.
- [54] K. Kunisch and S. Volkwein, Galerkin proper orthogonal decomposition methods for parabolic problems, *Numerische Mathematik*, 90(1), pp. 117–148, 2001.
- [55] K. Kunisch and S. Volkwein, Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics, *SIAM Journal on Numerical analysis*, 40(2), pp. 492–515, 2002.
- [56] K. Kunisch and S. Volkwein, Proper orthogonal decomposition for optimality systems, *ESAIM: Mathematical Modelling and Numerical Analysis*, 42(1), pp. 1–23, 2008.
- [57] 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*, pp. 235–273, Springer, 2014.
- [58] I. Lengyel and I. Epstein, Modeling of turing structures in the chlorite-iodide-malononic acid-starch reaction system, *Science*, 251(4994), pp. 650–652, 1991.
- [59] A. I. Lobanov and T. K. Starozhilova, The effect of convective flows on blood coagulation processes, *Pathophysiol Haemost Thromb*, 34, pp. 121–134, 2005.
- [60] J. Mann and J. N. Kutz, Dynamic mode decomposition for financial trading strategies, *Quantitative Finance*, 16(11), pp. 1643–1655, 2016.
- [61] T. T. Marquez-Lago and P. Padilla, A selection criterion for patterns in reaction–diffusion systems, *Theoretical Biology and Medical Modelling*, 11, p. 7, 2014.
- [62] P. G. Martinsson, Randomized methods for matrix computations and analysis of high dimensional data, arXiv preprint arXiv:1607.01649, 2016.
- [63] I. Mezić, Analysis of fluid flows via spectral properties of the Koopman operator, *Annual Review of Fluid Mechanics*, 45, pp. 357–378, 2013.
- [64] H. D. Mittelmann, Verification of second-order sufficient optimality conditions for semilinear elliptic and parabolic control problems, *Comput. Optim. Appl.*, 20(1), pp. 93–110, 2001.
- [65] H. Müller and S. Volkwein, Model reduction by proper orthogonal decomposition for lambda-omega systems, in *ECCOMAS CFD 2006: Proceedings of the European Conference on Computational Fluid Dynamics, Egmond aan Zee*,

- The Netherlands, September 5-8, 2006*, Delft University of Technology; European Community on Computational Methods in Applied Sciences (ECCOMAS), 2006.
- [66] M. Müller, *Uniform convergence of the POD method and Applications to Optimal Control*, Ph.D. thesis, Karl-Franzens-Universität Graz, 2011.
- [67] J. D. Murray, *Mathematical Biology: I. An Introduction*, Springer-Verlag New York, 2002.
- [68] J. D. Murray and J. Dickson, *Mathematical biology*, Biomathematics, Springer-Verlag, 1990, Berlin, New York, 1989.
- [69] J. Nagumo, S. Arimoto, and S. Yoshizawa, An Active Pulse Transmission Line Simulating Nerve Axon, *Proceedings of the IRE*, 50(10), pp. 2061–2070, 1962.
- [70] C. V. Pao, *Nonlinear parabolic and elliptic equations*, Springer Science & Business Media, 2012.
- [71] J. Y. Park and S. H. Park, Optimal control problems for anti-periodic quasi-linear hemivariational inequalities, *Optimal Control Applications and Methods*, 28(4), pp. 275–287, 2007.
- [72] J. Peraire and P. O. Persson, The compact discontinuous Galerkin (CDG) method for elliptic problems, *SIAM Journal on Scientific Computing*, 30, pp. 1806–1824, 2008.
- [73] W. H. Reed and T. R. Hill, *Triangular mesh methods for the neutron transport equation*, 1973.
- [74] B. Rivière, *Discontinuous Galerkin methods for solving elliptic and parabolic equations. Theory and implementation*, volume 35 of *Frontiers in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.
- [75] A. Rösch and D. Wachsmuth, A-posteriori error estimates for optimal control problems with state and control constraints, *Numerische Mathematik*, 120, pp. 733–762, 2012.
- [76] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, Spectral analysis of nonlinear flows, *Journal of Fluid Mechanics*, 641, pp. 115–127, 2009, ISSN 1469-7645.
- [77] C. Ryll, J. Löber, S. Martens, H. Engel, and F. Tröltzsch, Analytical, optimal, and sparse optimal control of traveling wave solutions to reaction-diffusion systems, in E. Schöll, S. H. L. Klapp, and P. Hövel, editors, *Control of Self-Organizing Nonlinear Systems*, pp. 189–210, Springer International Publishing, 2016.
- [78] C. Ryll and F. Tröltzsch, Proper orthogonal decomposition in sparse optimal control of some reaction diffusion equations using model predictive control, *PAMM*, 14(1), pp. 883–884, 2014.

- [79] C. P. Schenk, M. Or-Guil, M. Bode, and H. G. Purwins, Interacting pulses in three-component reaction-diffusion systems on two-dimensional domains, *Physical Review Letters*, 78, pp. 3781–3784, 1997.
- [80] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, *Journal of Fluid Mechanics*, 656, pp. 5–28, 2010.
- [81] S. K. Scott and P. Gray, Chemical reactions in isothermal systems: oscillations and instabilities, in *Nonlinear phenomena and chaos (Malvern, 1985)*, Malvern Phys. Ser., pp. 70–95, Hilger, Bristol, 1986.
- [82] L. Sirovich, Turbulence and the dynamics of coherent structures. i. coherent structures, *Quarterly of applied mathematics*, 45(3), pp. 561–571, 1987.
- [83] G. Stadler, Elliptic optimal control problems with L^1 -control cost and applications for the placement of control devices, *Computational Optimization and Applications*, 44(1), pp. 159–181, 2009.
- [84] M. Stich, A. S. Mikhailov, and Y. Kuramoto, Self-organized pacemakers and bistability of pulses in an excitable medium, *Physical Review E*, 79, p. 026110, 2009.
- [85] M. Stoll, J. M. Pearson, and P. K. Maini, Fast solvers for optimal control problems from pattern formation, *Journal of Computational Physics*, 304, pp. 27–45, 2016.
- [86] A. Studinger and S. Volkwein, Numerical analysis of POD a-posteriori error estimation for optimal control, in K. Bredies, C. Clason, K. Kunisch, and G. von Winckel, editors, *Control and Optimization with PDE Constraints*, Springer Basel, Basel, 2013.
- [87] G. Tissot, L. Cordier, N. Benard, and B. R. Noack, Model reduction using dynamic mode decomposition, *Comptes Rendus Mécanique*, 342(6), pp. 410–416, 2014.
- [88] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, volume 112 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, 2010.
- [89] F. Tröltzsch and S. Volkwein, POD a-posteriori error estimates for linear-quadratic optimal control problems, *Computational Optimization and Applications*, 44(1), pp. 83–115, 2008.
- [90] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, On dynamic mode decomposition: Theory and applications, *Journal of Computational Dynamics*, 1(2), pp. 391–421, 2014.
- [91] A. Turing, The chemical basis of morphogenesis, *Philosophical Transactions of the Royal Society B*, 237, pp. 37–72, 1952.
- [92] M. Uzunca, T. Küçükseyhan, H. Yücel, and B. Karasözen, Optimal control of convective FitzHugh-Nagumo equation, *Computers & Mathematics with Applications*, 73(9), pp. 2151 – 2169, 2017.

- [93] P. van Heijster, A. Doelman, and T. J. Kaper, Pulse dynamics in a three-component system: Stability and bifurcations, *Physica D: Nonlinear Phenomena*, 237(24), pp. 3335 – 3368, 2008.
- [94] P. van Heijster, A. Doelman, T. J. Kaper, and K. Promislow, Front interactions in a three-component system, *SIAM Journal on Applied Dynamical Systems*, 9(2), pp. 292–332, 2010.
- [95] C. Vilas, M. R. García, J. R. Banga, and A. A. Alonso, Robust feed-back control of travelling waves in a class of reaction-diffusion distributed biological systems, *Physica D: Nonlinear Phenomena*, 237(18), pp. 2353 – 2364, 2008.
- [96] S. Volkwein, Proper orthogonal decomposition: Theory and reduced-order modelling, *Lecture Notes*, University of Konstanz, 4(4), 2013.
- [97] G. Wachsmuth and D. Wachsmuth, Convergence and regularization results for optimal control problems with sparsity function, *ESAIM: Control, Optimisation and Calculus of Variations*, 17(3), pp. 858–886, 2010.
- [98] M. F. Wheeler, An elliptic collocation-finite element method with interior penalties, *SIAM Journal on Numerical Analysis*, 15, pp. 152–161, 1978.
- [99] A. Wynn, D. S. Pearson, B. Ganapathisubramani, and P. J. Goulart, Optimal mode decomposition for unsteady flows, *Journal of Fluid Mechanics*, 733, p. 473, 2013.
- [100] E. Yanagida, Mini-maximizers for reaction-diffusion systems with skew-gradient structure, *Journal of Differential Equations*, 179, pp. 311–335, 2002.
- [101] E. Yanagida, Standing pulse solutions in reaction-diffusion systems with skew-gradient structure, *Journal of Dynamics and Differential Equations*, 14, pp. 189–205, 2002.
- [102] H. Yücel and P. Benner, Distributed optimal control problems governed by coupled convection dominated pdes with control constraints, in A. Abdulle, S. Deparis, D. Kressner, F. Nobile, and M. Picasso, editors, *Numerical Mathematics and Advanced Applications - ENUMATH 2013*, volume 103 of *Lecture Notes in Computational Science and Engineering*, pp. 469–478, Springer International Publishing, 2015.
- [103] H. Yücel and B. Karasözen, Adaptive symmetric interior penalty Galerkin (SIPG) method for optimal control of convection diffusion equations with control constraints, *Optimization*, 63, pp. 145–166, 2014.
- [104] H. Yücel, M. Stoll, and P. Benner, A discontinuous Galerkin method for optimal control problems governed by a system of convection-diffusion PDEs with nonlinear reaction terms, *Computers & Mathematics with Applications*, 70, pp. 2414–2431, 2015.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Küçükseyhan, Tuğba
Nationality: Turkish (TC)
Date and Place of Birth: 01.04.1989, Bolu
Marital Status: Married
Phone: +903122105365
Fax: +903122102972
e-mail: guney.tugba@metu.edu.tr

EDUCATION

Degree	Institution	Year of Graduation
Ph.D.	Scientific Computing, IAM, METU	2017
B.S.	Mathematics Department, METU	2012
High School	Bolu Atatürk High School (Foreign Language Intensive)	2007

LANGUAGE SKILLS

Turkish(native), English(fluently)

COMPUTER SKILLS

Proficient in MATLAB, LATEX

PROFESSIONAL EXPERIENCE

- 02/2017–present** Research Assistant, Department of Mathematics
Balıkesir University, Balıkesir, TURKEY
- 12/2013–02/2017** Research Assistant, Department of Mathematics
Middle East Technical University, Ankara, TURKEY

TEACHING EXPERIENCE

MATH 117 - Calculus with Analytic Geometry (METU)

Teaching Assistant for 2015-2016 Spring Semester

MATH 119 - Calculus with Analytic Geometry (METU)

Teaching Assistant for 2014-2015, 2016-2017 Fall Semesters

MATH 120 - Calculus of Functions of Several Variables (METU)

Teaching Assistant for 2014-2015 Spring Semester

MATH 480 - Numerical Methods for Differential Equations (METU)

Teaching Assistant for 2015-2016 Spring Semester

AWARDS AND SCHOLARSHIPS

- 09/2012–present** 2211-Doctorate Scholarship
Turkish Scientific and Technical Research Council (TÜBİTAK)
- 06/2015–12/2015** 2214/A-International Doctoral Research Fellowship Programme
Turkish Scientific and Technical Research Council (TÜBİTAK)
- 09/2007–08/2012** 2205-Domestic Undergraduate Scholarship
Turkish Scientific and Technical Research Council (TÜBİTAK)
- 06/2012** Graduated as a best student with CPGA 3.88/4.
Mathematic Department, METU.
- 06/2006** Graduated as a best student with CPGA 5/5.
Ataturk High School (Foreign Language Intensive), Bolu.

RESEARCH VISITS

- 06/2015–12/2015** Institute of Mathematics, TU Berlin, Berlin/Germany,
supervised by Prof. Dr. Fredi Tröltzsch

PUBLICATIONS

Publications in International Journals

B. Karasözen, T. Küçükseyhan, M. Uzunca, *Reduced Order Optimal Control of the Convective FitzHugh-Nagumo equation*. arXiv preprint arXiv:1703.00008, (2017).

M. Uzunca, T. Küçükseyhan, H. Yücel, B. Karasözen, *Optimal control of convective FitzHugh-Nagumo equation*. *Computers & Mathematics with Applications*, 73 (2017), 2151 - 2169.

DOI: <http://doi.org/10.1016/j.camwa.2017.02.028>

B. Karasözen, T. Küçükseyhan, M. Uzunca, *Structure preserving integration and model order reduction of skew-gradient reaction-diffusion systems*. *Annals of Operations Research* (2015), 1-28.

DOI: <http://dx.doi.org/10.1007/s10479-015-2063-6>

M. Uzunca, B. Karasözen, T. Küçükseyhan, *Moving mesh discontinuous Galerkin methods for PDEs with traveling waves*. *Applied Mathematics and Computation*, 292 (2017), 9-18.

DOI: <http://dx.doi.org/10.1016/j.amc.2016.07.034>

International Conference Papers

B. Karasözen, M. Uzunca, T. Küçükseyhan, *Model Order Reduction for Pattern Formation in FitzHugh-Nagumo Equations*. *Numerical Mathematics and Advanced Applications ENUMATH*, Springer (2015).

DOI: http://dx.doi.org/10.1007/978-3-319-39929-4_35

SCIENTIFIC MEETINGS

Presentations in International Scientific Meetings

T. Küçükseyhan, B. Karasözen, M. Uzunca, Çankaya University International Workshop on Mathematical Models in Engineering Ankara (Turkey), April 27-29, 2017.

Contributed Talk: Model Order Reduction for Pattern Formation of FitzHugh-Nagumo Equation

T. Küçükseyhan, B. Karasözen, M. Uzunca, 14th EUROPT Workshop on Advances in Continuous Optimization Warsaw (Poland), July 1-2, 2016.

Contributed Talk: Optimal Control of Convective FitzHugh-Nagumo Equation

The Third International School on Model Order reduction for Dynamical Control Systems, Dubrovnik, Croatia, October 5-10, 2015.

B. Karasözen, M. Uzunca, **T. Küçükseyhan**, European Conference on Numerical Mathematics and Advanced Applications (ENUMATH 2015), Ankara – Turkey, September 14-18, 2015.

Contributed Talk: Structure Preserving Integration and Model Order Reduction of Skew-Gradient Reaction Diffusion Systems

T. Güney, B. Karasözen, M. Uzunca, 27th IFIP TC7 Conference 2015 on System Modelling and Optimization, Sophia Antipolis-France, June 29 – July 3, 2015.

Contributed Talk: Optimal Control of the Gradient Systems

T. Güney, M. Uzunca, B. Karasözen, Numerical Approximation of PDEs., Zurich - Switzerland, September 8-10, 2014.

Contributed Poster: Structure Preserving Integration of Skew-Gradient Reaction Diffusion Systems

Presentations in National Scientific Meetings

T. Güney, B. Karasözen, M. Uzunca, 27. National Mathematics Symposium, Yeditepe University, Istanbul-Turkey, August 26-29, 2014.

Contributed Talk: Moving Mesh Solution of PDEs with Discontinuous Galerkin Finite Elements.