A Study on Fractal Interpolation in Shape Preserving and in Numerical Methods for Ordinary Differential Equations

by

N. Balasubramani

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039, INDIA

February, 2019

A Study on Fractal Interpolation in Shape Preserving and in Numerical Methods for Ordinary Differential Equations

A thesis submitted

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

by

N. Balasubramani

(Roll Number: 126123020)

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039, INDIA

February, 2019

DECLARATION

It is certified that the work contained in the thesis titled **"A Study on Fractal Interpolation in Shape Preserving and in Numerical Methods for Ordinary Differential Equations"** has been done by me, a student in the Department of Mathematics, Indian Institute of Technology Guwahati under the guidance of **Prof. M. Guru Prem Prasad and Prof. S. Natesan**, Indian Institute of Technology Guwahati, for the award of **Doctor of Philosophy** and that this work has not been submitted elsewhere for a degree.

Guwahati
February 2019.

N. Balasubramani
Roll No: 126123020
Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati-781039, India

CERTIFICATE

It is certified that the work contained in the thesis titled **"A Study on Fractal Interpolation in Shape Preserving and in Numerical Methods for Ordinary Differential Equations"** by **N. Balasubramani (126123020)**, a student in the Department of Mathematics, Indian Institute of Technology Guwahati for the award of the degree of **Doctor of Philosophy** has been carried out under our supervision and this work has not been submitted elsewhere for a degree.

Prof. M. Guru Prem Prasad
Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati-781039, India

Prof. S. Natesan
Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati-781039, India

Dedicated

To

My Parents

Acknowledgements

Undertaking this PhD is a long journey and it would not have been possible to reach the destination without the support and guidance from many people. I would like to take this opportunity to express my deepest gratitude for their generous support.

First of all, I am greatful to my supervisors Prof. M. Guru Prem Prasad and Prof. S. Natesan for their guidance throughout my research period. During my Ph. D. period, they have shown extraordinary patience with me and I have been able to learn many things with their help. They carefully guided me through this research period and discussed all ideas in this thesis. They have put a tremendous amount of effort in this thesis from their busy schedule of different administrative responsibilities. Without their guidance and persistent help, this dissertation would not have been possible.

Besides my supervisors, I want to convey my sincere thanks to my doctoral committee members Prof. D. C. Dalal, Prof. R. K. Sinha and Prof. S. N. Bora for their valuable suggestions. I also convey my gratitude to all the faculty members who have offered their help at different stages during my research time.

I would like to thank my friends Anirudha Poria, Maneesh Kumar Singh, Anirban Majumdar, Abhishek Das, Md. Nasim Akhtar, Madhusudan Bera, Ranjan Kumar Das, Ramesh Prasad Panda, Ashish Kumar Kesarwany, Tamal Pramanick, Koyel Chakravarty, Gautam Singh, Sougata Biswas, Swapnendu Panda, Devanand, Vignesh Babu, Vigneshwaran and Johnney Mertens. My special thanks to my friends P. Muthukumar, Swarup Barik, Ananda Chandra Nayak, Ankur Kanaujiya and Ramesh Kumar Jallu for their help at different occasions.

My sincere thanks to the Ministry of Human Resource Development, Government of India, for providing me financial support for pursuing PhD at IIT Guwahati. I sincerely acknowledge IIT Guwahati for providing me the necessary facilities for conducting my research work smoothly.

Finally, spacial homage goes to my family for their continuous love, encouragement and support throughout my life. This work would not be possible without the strong moral support and the unfailing faith of my parents (my father Mr. K. Natesan and my mother Mrs. N. Arukkani) that has sustained me throughout this endeavour. I would like to express my heartfelt gratitude towards them for everything they have done for me over the years. I would also like to convey my deepest love to all my family members.

February, 2019 **N. Balasubramani**

Abstract

Fractal interpolation is an eminent tool to approximate a function whose derivative of a certain order is irregular. The interpolant obtained by the traditional interpolation methods are either infinitely differentiable or piecewise infinitely differentiable. In many situations, interpolation data contains the shape properties (positivity, monotonicity, convexity etc.) and also data representing a derivatives have irregularity in dense subsets of the interval. Though traditional non-recursive interpolant methods produces shape preserving interpolants, these methods are not suitable for the irregular representation of the derivatives. In this thesis, the shape preserving fractal interpolants are constructed and derivatives of these interpolants can have points of non-differentiability in finite or a dense subset of the interval. Apart from constructing shape preserving fractal interpolation functions, we proposed the numerical schemes with the help of fractal interpolation functions to get the numerical approximations for the boundary-value problems of ordinary differential equations.

Chapter 1, is the introductory chapter reviewing the basics and some useful results related to our work. In Chapter 2, we introduced the fractal rational cubic spline with two families of shape parameters and the fractal rational cubic spline with three families of shape parameters to interpolate the univariate data. Also, the scaling factors and shape parameters are constrained to preserve the shape properties of the interpolation data. Convergence analysis of the constructed fractal rational cubic splines are established. In Chapter 3, a fractal interpolation surface is constructed with the help of the rational cubic splines that contains three families of shape parameters to interpolate the bivariate data that lies on the rectangular grid. Shape preserving aspects and convergence analysis of the fractal interpolation surface are presented.

In Chapter 4, a fractal cubic spline is used to get the numerical solutions of the self-adjoint singularly perturbed boundary-value problems with Dirichlet and Neumann boundary conditions. Convergence analysis of the developed methods are carried out. In Chapter 5, the approximate solutions of nonself-adjoint singularly perturbed boundary-value problems are obtained with the help of fractal cubic spline method. Also, linear and non-linear singular boundary-value problems are considered and these problems are solved using fractal cubic spline. Error analysis of the proposed methods are established. In Chapter 6, we provide numerical methods with the help of fractal non-polynomial cubic spline to solve the both self-adjoint and nonself-adjoint singularly perturbed boundary-value problems. Convergence analysis of the developed schemes are discussed.

In Chapter 7, the numerical solutions of the self-adjoint singularly perturbed boundary-value problems are obtained with the help of fractal quintic spline. Also, a fractal quintic spline is used to solve the non-linear boundary-value problems. Error analysis of the proposed methods are derived. In Chapter 8, the fourth-order boundary-value problems are solved with the help of fractal quintic splines. Truncation error of the developed methods are derived.

List of Figures

2.1 Positivity preserving interpolation. 41
2.2 Derivatives of FIFs given in Figure 2.1. 42
2.3 Monotonicity preserving interpolation. 44
2.4 Derivatives of FIFs given in Figure 2.3. 45
2.5 Convexity preserving interpolation. 47
2.6 Derivatives of FIFs given in Figure 2.5. 48
2.7 Constrained interpolation. 50
2.8 Positive preserving interpolation. 67
2.9 Derivatives of the FIFs given in Figure 2.8. 68
2.10 Monotonicity preserving interpolation. 70
2.11 Derivatives of the FIFs given in Figure 2.10. 71
2.12 Convexity preserving interpolation. 73
2.13 Derivatives of the FIFs given in Figure 2.12. 74

3.1 Surface interpolation. 85
3.2 Positive surface interpolation. 110
3.3 Constrained surface interpolation. 111
3.4 Monotone surface interpolation. 113
3.5 Convex surface interpolation. 115

4.1 Numerical results of Example 4.1.1. 131
4.2 Numerical results of Example 4.1.2. 131
4.3 Log-log plot of Example 4.2.1. 144

5.1 Numerical results of Example 5.1.1. 153
5.2 Numerical results of Example 5.1.2. 155
6.1 Log-log plots of Example 6.1.1. 186
6.2 Log-log plots of Example 6.1.2. 188
6.3 Numerical results of Example 6.2.1. 195
7.1 Log-log plots of Example 7.1.1. 211
7.2 Log-log plots of Example 7.1.2. 213

List of Tables

2.1 Parameters for positive FIFs with $u_i = 1.5$ for $i = 1, 2, \ldots, 7$. 43

2.2 Parameters for monotonic FIFs with $u_i = 1.5$ for $i = 1, 2, \ldots, 6$. 46

2.3 Parameters for convex FIFs with $u_i = 1.5$ for $i = 1, 2, 3, 4$. 49

2.4 Parameters for constrained FIFs with $u_i = 1.5$ for $i = 1, 2, \ldots, 6$. 49

2.5 Parameters for positive interpolation with $u_i = 1.5$ and $v_i = 1$ for $i = 1, 2, \ldots, 6$. 69

2.6 Parameters for monotonic interpolation with $u_i = 1.5$ and $v_i = 1$ for $i = 1, 2, \ldots, 8$. 72

2.7 Parameters for convex interpolation with $u_i = 1.5$ and $v_i = 1$ for $i = 1, 2, \ldots, 5$. 75

3.1 Surface data 85

3.2 Positive data 109

3.3 Scaling matrices for positive interpolation 109

3.4 Shape parameter matrices for positive interpolation with $\boldsymbol{u} = [1.5]_{3\times4}$, $\boldsymbol{u}^* = [1.5]_{4\times3}$ $\boldsymbol{v} = [1]_{3\times4}$ and $\boldsymbol{v}^* = [1]_{4\times3}$. 110

3.5 Constrained data 111

3.6 Scaling matrices for constrained interpolation. 112

3.7 Shape parameters for constrained interpolation with $\boldsymbol{u} = [1.5]_{5\times6}$, $\boldsymbol{u}^* = [1.5]_{6\times5}$, $\boldsymbol{v} = [1]_{5\times6}$ and $\boldsymbol{v}^* = [1]_{6\times5}$. 112

3.8 Monotonic data 113

3.9 Scaling matrices for monotone interpolation 113

3.10 Shape parameter matrices for monotone interpolation with $\boldsymbol{u} = [1.5]_{3\times4}$, $\boldsymbol{u}^* = [1.5]_{4\times3}$ $\boldsymbol{v} = [1]_{3\times4}$ and $\boldsymbol{v}^* = [1]_{4\times3}$. 114
3.11 Convex data. 114
3.12 Scaling matrices for convex interpolation 115
3.13 Shape matrices for convex interpolation with $\boldsymbol{u} = [1.5]_{4\times5}$, $\boldsymbol{u}^* = [1.5]_{5\times4}$, $\boldsymbol{v} = [1]_{4\times5}$ and $\boldsymbol{v}^* = [1]_{5\times4}$. 116

4.1 Maximum point-wise error and order of convergence corresponding to Example 4.1.1. 129
4.2 Maximum point-wise error and order of convergence corresponding to Example 4.1.1 with cubic spline. 130
4.3 Maximum point-wise error and order of convergence corresponding to Example 4.1.2. 132
4.4 Maximum point-wise error and order of convergence corresponding to Example 4.1.2 with cubic spline. 133
4.5 Maximum point-wise error and order of convergence corresponding to Example 4.2.1. 142
4.6 Maximum point-wise error and order of convergence corresponding to Example 4.2.1 with cubic spline. 143

5.1 Maximum point-wise error and order of convergence corresponding to Example 5.1.1. 153
5.2 Maximum point-wise error and order of convergence corresponding to Example 5.1.1 with $\alpha = 0$. 154
5.3 Maximum point-wise error and order of convergence corresponding to Example 5.1.2. 155
5.4 Maximum point-wise error and order of convergence corresponding to Example 5.1.2 with $\alpha = 0$. 156
5.5 Maximum point-wise error and order of convergence corresponding to Example 5.2.1. 164

5.6 Maximum point-wise error and order of convergence corresponding to Example 5.2.1 with $\alpha = 0$. 164
5.7 Maximum point-wise error and order of convergence corresponding to Example 5.2.2. 164
5.8 Maximum point-wise error and order of convergence corresponding to Example 5.2.2 with $\alpha = 0$. 165
5.9 Maximum point-wise error and order of convergence corresponding to Example 5.3.1. 170
5.10 Maximum point-wise error and order of convergence corresponding to Example 5.3.1 with $\alpha = 0$. 170

6.1 Scaling factor for Example 6.1.1. 185
6.2 Maximum point-wise error and order of convergence corresponding to Example 6.1.1. 185
6.3 Maximum point-wise error and order of convergence corresponding to Example 6.1.1 with $\alpha = 0$ 186
6.4 Scaling factor for Example 6.1.2. 187
6.5 Maximum point-wise error and order of convergence corresponding to Example 6.1.2. 187
6.6 Maximum point-wise error and order of convergence corresponding to Example 6.1.2 with $\alpha = 0$. 188
6.7 Scaling factor for Example 6.2.1. 195
6.8 Maximum point-wise error and order of convergence corresponding to Example 6.2.1. 196
6.9 Maximum point-wise error and order of convergence corresponding to Example 6.2.1 with $\alpha = 0$. 197

7.1 Maximum point-wise error and order of convergence corresponding to Example 7.1.1. 211
7.2 Maximum point-wise error and order of convergence corresponding to Example 7.1.1 with $\alpha = 0$. 212

7.3 Maximum point-wise error and order of convergence corresponding to Example 7.1.2. 213
7.4 Maximum point-wise error and order of convergence corresponding to Example 7.1.2 with $\alpha = 0$. 214
7.5 Maximum point-wise error and order of convergence of Example 7.2.1. . . 223
7.6 Maximum point-wise error an order of convergence of Example 7.2.1 with $\alpha = 0$. 223
7.7 Maximum point-wise error and order of convergence of Example 7.2.2. . . 224
7.8 Maximum point-wise error and order of convergence of Example 7.2.2 with $\alpha = 0$. 224

8.1 Maximum point-wise error and order of convergence corresponding to Example 8.1.1. 233
8.2 Maximum point-wise error and order of convergence corresponding to Example 8.1.1 with $\alpha = 0$. 233
8.3 Maximum point-wise error and order of convergence corresponding to Example 8.1.2. 234
8.4 Maximum point-wise error and order of convergence corresponding to Example 8.1.2 with $\alpha = 0$. 234
8.5 Maximum point-wise error and order of convergence corresponding to Example 8.2.1. 243
8.6 Maximum point-wise error and order of convergence corresponding to Example 8.2.1 with $\alpha = 0$. 243
8.7 Maximum point-wise error and order of convergence corresponding to Example 8.2.2. 244
8.8 Maximum point-wise error and order of convergence corresponding to Example 8.2.2 with $\alpha = 0$. 244

Contents

List of Figures xi

List of Tables xiii

1 Introduction **1**

1.1 Overview . . . 1
1.2 Iterated Function Systems . . . 2
1.3 Fractal Interpolation Functions . . . 3
1.4 Smooth Fractal Interpolation Functions . . . 4
1.5 Fractal Interpolation Surfaces . . . 6
1.6 Shape Preserving FIF and FIS . . . 6
1.7 Spline Functions . . . 10
1.8 Motivation of Present Work . . . 17
1.9 Organization of Present Work . . . 19

2 Shape Preserving Rational Cubic Fractal Interpolation Functions **23**

2.1 FIF with Two Families of Shape Parameters . . . 24
2.1.1 Convergence analysis . . . 28
2.1.2 Shape preserving aspects of FIF . . . 31
2.1.3 Numerical examples . . . 40
2.2 FIF with Three Families of Shape Parameters . . . 51
2.2.1 Convergence analysis . . . 53
2.2.2 Shape preserving interpolation . . . 59

2.2.3 Numerical examples 66
2.3 Conclusion 75

3 Shape Preserving Rational Cubic Fractal Interpolation Surfaces 77
3.1 Fractal Interpolation Surfaces 78
3.1.1 Construction of fractal boundary curves 78
3.1.2 Construction of rational cubic FIS 82
3.1.3 Error analysis 86
3.2 Shape Preserving FIS 90
3.2.1 Positivity preserving interpolation 91
3.2.2 Constrained interpolation 95
3.2.3 Monotonicity preserving interpolation 98
3.2.4 Convexity preserving interpolation 105
3.3 Numerical Examples 109
3.4 Conclusion 116

4 Fractal Cubic Spline Solutions for Self-Adjoint Boundary-Value Problems 117
4.1 Dirichlet Boundary-Value Problem 118
4.1.1 Convergence analysis 121
4.1.2 Numerical examples 128
4.2 Neumann Boundary-Value Problem 134
4.2.1 Convergence analysis 135
4.2.2 Numerical examples 141
4.3 Conclusion 144

5 Fractal Cubic Spline Solutions for Nonself-Adjoint Boundary-Value Problems 145
5.1 Singularly Perturbed Nonself-Adjoint Boundary-Value Problems 146
5.1.1 Error analysis 148
5.1.2 Numerical examples 152

5.2 Linear Singular Boundary-Value Problems 156
5.2.1 Convergence analysis 158
5.2.2 Numerical examples 163
5.3 Non-Linear Singular Boundary-Value Problems 165
5.3.1 Error analysis 167
5.3.2 Numerical examples 169
5.4 Conclusion 171

6 Fractal Non-Polynomial Cubic Spline Solutions for Singularly Perturbed Boundary-Value Problems 173
6.1 Self-Adjoint Boundary-Value Problems 174
6.1.1 Convergence analysis 177
6.1.2 Numerical examples 184
6.2 Nonself-Adjoint Boundary-Value Problems 189
6.2.1 Convergence analysis 190
6.2.2 Numerical examples 195
6.3 Conclusion 197

7 Fractal Quintic Spline Solutions for Boundary-Value Problems 199
7.1 Self-Adjoint Boundary-Value Problems 200
7.1.1 Convergence analysis 206
7.1.2 Numerical examples 210
7.2 Nonlinear Boundary-Value Problem 214
7.2.1 Numerical examples 222
7.3 Conclusion 224

8 Fractal Quintic Spline Solutions for Fourth-Order Boundary-Value Problems 225
8.1 Numerical Method 226
8.1.1 Truncation error 230
8.1.2 Numerical examples 232

8.2 Numerical Scheme 235
8.2.1 Truncation error 240
8.2.2 Numerical examples 243
8.3 Conclusion 245

Bibliography **247**

List of published and communicated papers **259**

Abbreviation and Notation

IFS	Iterated Function System
FIF	Fractal Interpolation Function
FIS	Fractal Interpolation Surface
BVP	Boundary-Value Problem
$\|.\|$	Norm
$\mathcal{C}(I)$	Space of real valued continuous functions on I
$\mathcal{C}^r(I)$	Space of real valued functions that are r times differentiable with continuous r-th derivative on I
$f(x) = \mathcal{O}(g(x))$ as $x \to a$	$\lvert f(x)\rvert \leq k\lvert g(x)\rvert$ for x in a neighborhood of a, where k is a positive constant

Chapter 1

Introduction

1.1 Overview

The notion of fractals occupies an important place to understand the structures of objects found in nature and to study of non-linearity. Mandelbrot [91] coined the term fractals to describe the objects that were too irregular to fit into the traditional geometric setting. Fractal geometry emerged as an extension of classical Euclidean geometry and has permeated many areas of science, such as astrophysics [47], biological sciences [59], medical sciences [66] and computer graphics [84]. Two of the most important properties of fractals are self-similarity and non-integer dimension. Mandelbrot [92] recognized that many objects in nature are nothing but recurrence of patterns at all scales, no matter how small. Mandelbrot defined a fractal to be a set with Hausdorff dimension strictly greater than its topological dimension. Falconer [56] suggested that a set should be called fractal, if it has some or all of the following features:

- It has fine structure at arbitrary small scales.
- It is too irregular to be described by traditional Euclidean geometry.
- It is self-similar, at least approximately.
- It has the Hausdorff dimension greater than its topological dimension.
- It has a simple and recursive definition.

The last property *simple and recursive definition* is very useful in constructing fractals from a given interpolation data as graphs of functions.

1.2 Iterated Function Systems

Let (X, d) be a complete metric space. Let $H(X)$ be the set of all nonempty compact subsets of X. Then $H(X)$ is a complete metric space with respect to the Hausdorff metric ρ, where ρ is defined as

$$\rho(A, B) = \max\Big\{d(A, B), d(B, A)\Big\},\ \forall A, B \in H(X),$$

where $d(A, B) = \max_{x \in A}\ \min_{y \in B} d(x, y)$. Let $w_i : X \to X$, $i = 1, 2, \ldots, M$ be continuous functions on X. Then $\{X; w_i, i = 1, 2, \ldots, M\}$ is called an iterated function system (IFS). If the maps w_i, $i = 1, 2, \ldots, M$, are contractive, i.e., $d(w_i(x), w_i(y)) \le c_i\ d(x, y)$ for all x, $y \in X$, where $0 \le c_i < 1$, then the IFS $\{X; w_i, i = 1, 2, \ldots, M\}$ is called hyperbolic. The contractive factor of the hyperbolic IFS is defined as $c = \max\{c_i : i = 1, 2, \ldots, M\}$. Associated with the hyperbolic IFS $\{X; w_i, i = 1, 2, \ldots, M\}$, there is a set valued map W from $H(X)$ into itself, i.e.,

$$W : H(X) \to H(X), \quad W(A) := \bigcup_{i=1}^{M} w_i(A).$$

The map W is a contraction map on $\Big(H(X), \rho\Big)$ with the contractive factor c.

Theorem 1.2.1. *Let $\{X; w_i, i = 1, 2, \ldots, M\}$ be a hyperbolic IFS on the complete metric space (X, d). Then there exists a set A is called the attractor of IFS $\{X; w_i, i = 1, 2, \ldots, M\}$ such that $A = \lim_{m\to\infty} W^m(B)$ for every $B \in H(X)$ and $W(A) = A$.*

Theorem 1.2.1 is called the Banach fixed-point theorem. In Theorem 1.2.1, the limit is taken with respect to the Hausdorff metric and W^m is the m-fold composition of W with itself. Next, we address the construction of functions that interpolate a given data set and whose graphs are attractors of suitable IFSs.

1.3 Fractal Interpolation Functions

Let $\{(x_i, y_i) \in \mathbb{R}^2 : i = 1, 2, \ldots, N\}$ be a data set such that $x_1 < x_2 < \cdots < x_N$. Let $I = [x_1, x_N]$ and $I_i = [x_i, x_{i+1}]$. For each $i = 1, 2, \ldots, N-1$, let $L_i : I \to I_i$ be a contraction homeomorphisms such that

$$L_i(x_1) = x_i, \quad L_i(x_N) = x_{i+1}, \ |L_i(x) - L_i(x^*)| \leq l_i |x - x^*|, \tag{1.1}$$

where $x, \ x^* \in I$, $0 \leq l_i < 1$. For each $i = 1, 2, \ldots, N-1$, let $F_i : I \times \mathbb{R} \to \mathbb{R}$ be a continuous function satisfy the conditions

$$F_i(x_1, y_1) = y_i, \quad F_i(x_N, y_N) = y_{i+1}, \ |F_i(x, y) - F_i(x, y^*)| \leq \alpha_i |y - y^*|, \tag{1.2}$$

where $x \in I, \ y, \ y^* \in \mathbb{R}, \ 0 \leq \alpha_i < 1$. Now, define functions $W_i : I \times \mathbb{R} \to I \times \mathbb{R}$ by

$$W_i(x, y) = (L_i(x), F_i(x, y)), \ i = 1, 2, \ldots, N-1.$$

Then $\{I \times \mathbb{R}; W_i : i = 1, 2, \ldots, N-1\}$ is an IFS.

Theorem 1.3.1. *[10] The IFS $\{I \times \mathbb{R}; W_i : i = 1, 2, \ldots, N-1\}$ has a unique attractor G such that G is the graph of a continuous function $f : I \to \mathbb{R}$ which obeys $f(x_i) = y_i$, $i = 1, 2, \ldots, N$.*

The function f given in Theorem 1.3.1 is called a fractal interpolation function (FIF) and it can be seen as a fixedpoint of an operator that defined in some complete metric space as follows:

Let $\mathscr{F} := \{g \in \mathcal{C}(I) : g(x_1) = y_1, \ g(x_N) = y_N\}$ be endowed with the uniform metric $d(g_1, g_2) := \max\{|g_1(x) - g_2(x)| : x \in I\}$. Let $T : \mathscr{F} \to \mathscr{F}$ be defined by

$$(Tg)(L_i(x)) = F_i(x, g(x)), \quad x \in I, \ i = 1, 2, \ldots, N-1. \tag{1.3}$$

Then T has a unique fixedpoint f and the function f satisfies the following implicit relation:

$$f(L_i(x)) = F_i(x, f(x)), \quad x \in I, \ i = 1, 2, \ldots, N-1.$$

The operator defined in (1.3) is called a Read-Bajraktarević operator and f satisfies the conditions given in Theorem 1.3.1. The widely studied FIFs are defined by the IFS $\{I \times \mathbb{R}; W_i(x,y) = (L_i(x), F_i(x,y)) : i = 1,2,\ldots,N-1\}$ with

$$L_i(x) = a_i x + b_i, \quad F_i(x,y) = \alpha_i y + r_i(x),\ i = 1,2,\ldots,N-1, \tag{1.4}$$

where α_i is called the scaling factor and $|\alpha_i| < 1$, a_i, b_i are constants and they are evaluated using the conditions given in (1.1) as

$$a_i = \frac{x_{i+1} - x_i}{x_N - x_1}, \quad b_i = \frac{x_i x_N - x_{i+1} x_1}{x_N - x_1},$$

and r_i is a suitable continuous function so that F_i satisfy the conditions given in (1.2). If r_i, $i = 1,2,\ldots,N-1$ are affine in (1.4), then the corresponding FIF is called the affine FIF. By taking r_i, $i = 1,2,\ldots,N-1$ as suitable polynomials, polynomial FIFs can be obtained. The parameter $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{N-1})$ is called the scale vector of the IFS.

The main differences between fractal interpolation method and traditional interpolation method are as follows:

- The FIFs are constructed through iteration not using an analytic formula.
- The functional equation of the FIFs provides a self-similarity in small scales.
- The scaling factors provide more flexibility in the choices of interpolants.

1.4 Smooth Fractal Interpolation Functions

In general, FIFs obtained through the IFS given in (1.4) need not be differentiable, for example one can see in the Reference [99]. However, restricting the scaling factors and choosing special types of r_i in (1.4), Barnsley and Harrington [11] constructed smooth FIFs. With the help of Barnsley and Harrington results, different types of smooth FIFs were constructed [30, 100]. The following theorem shows that the integral of a FIF passing through a given interpolation data is also a FIF.

Theorem 1.4.1. *[11] Let f be the FIF associated with the IFS $\{I\times\mathbb{R};(L_i(x),F_i(x,y)):i=1,2,\dots,N-1\}$ where $L_i(x)=a_ix+b_i$, $F_i(x,y)=\alpha_iy+r_i(x)$ and let $\hat{f}(x)=\hat{y}_1+\int_{x_1}^{x}f(t)dt$. Then, $\hat{f}$ is the FIF associated with the IFS $\{I\times\mathbb{R};(L_i(x),\hat{F}_i(x,y)):i=1,2,\dots,N-1\}$, where $\hat{F}_i(x,y)=a_i\alpha_iy+\hat{y}_i-a_i\alpha_i\hat{y}_1+a_i\int_{x_1}^{x}r_i(t)dt$, $\hat{y}_1$ is arbitrary and the values $\hat{y}_2,\hat{y}_3,\dots,\hat{y}_N$ are computed as*

$$\begin{aligned}\hat{y}_i=&\hat{y}_1+\sum_{j=1}^{i}a_j\Big[\alpha_j(\hat{y}_N-\hat{y}_1)+\int_{x_1}^{x_N}r_i(t)dt\Big],\ i=2,3,\dots,N-1,\\ \hat{y}_N=&\hat{y}_1+\Big(\sum_{j=1}^{N-1}a_j\int_{x_1}^{x_N}r_i(t)dt\Big)\Big(1-\sum_{j=1}^{N-1}a_j\alpha_j\Big)^{-1}.\end{aligned}$$

The FIF $\hat{f}$ given in Theorem 1.4.1 interpolates the data $\{(x_i,\hat{y}_i):i=1,2,\dots,N\}$. The following proposition is an immediate consequence of the above theorem and it gives the relation between the IFS of f and the IFS of $\hat{f}$.

Proposition 1.4.1. *[11] Let f be the FIF corresponding to the IFS defined by the maps $L_i(x)=a_ix+b_i$ and $F_i(x,y)=\alpha_iy+r_i(x)$. Then $\hat{f}'=f$ if and only if $\hat{f}$ is the FIF associated with the IFS $\{I\times\mathbb{R};(L_i(x),\hat{F}_i(x,y)):i=1,2,\dots,N-1\}$, where $\hat{F}_i(x,y)=\hat{\alpha}_iy+\hat{r}_i(x)$, where $\hat{\alpha}_i=a_i\alpha_i$, $\hat{r}_i'=a_ir_i$, $i=1,2,\dots,N-1$.*

By iterating the Proposition 1.4.1, the conditions on α_i and r_i of the IFS (1.4) are obtained so that the corresponding FIF belongs to the class $\mathcal{C}^r$.

Theorem 1.4.2. *[11] Let $\{(x_i,y_i):i=1,2,\dots,N\}$ be the set of interpolation data, where $x_1<x_2<\dots<x_N$. Let $L_i(x)=a_ix+b_i$, $i=1,2,\dots,N-1$, satisfy (1.1) and $F_i(x,y)=\alpha_iy+r_i(x)$, $i=1,2,\dots,N-1$, satisfy (1.2). Suppose that for some integer $r\geq 0$, $|\alpha_i|<a_i^r$ and $r_i\in\mathcal{C}^r(I)$, $i=1,2,\dots,N-1$. Let*

$$F_{i,k}=\frac{\alpha_iy+r_i^{(k)}(x)}{a_i^k},\ y_{1,k}=\frac{r_1^{(k)}(x_1)}{a_1^k-\alpha_1},\ y_{N,k}=\frac{r_{N-1}^{(k)}(x_N)}{a_{N-1}^k-\alpha_{N-1}},\ k=1,2,\dots,r.$$

If $F_{i-1,k}(x_N,y_{N,k})=F_{i,k}(x_1,y_{1,k})$ for $i=2,3,\dots,N-1$ and $k=1,2,\dots,r$, then the IFS $\{I\times\mathbb{R};(L_i(x),F_i(x,y)):i=1,2,\dots,N-1\}$ determines a FIF $f\in\mathcal{C}^r(I)$, and $f^{(k)}$ is the FIF determined by $\{I\times\mathbb{R};(L_i(x),F_{i,k}(x,y)):i=1,2,\dots,N-1\}$, for $k=1,2,\dots,r$.

1.5 Fractal Interpolation Surfaces

Fractal surfaces are found in abundance in nature. The surface of mountains, rocks, clouds are irregular in nature and they can be approximated by fractal surfaces. Different methods have been developed to construct fractal surfaces, for example, Mandelbrot [90] described fractal Brownian surface through random translation of fractal curves and Voss [147,148] generated fractal landscapes using successive random addition algorithm.

In order to have the fractal surface that passes through a prescribed data, Massopust [93] initiated the construction of fractal interpolation surface (FIS). He constructed the FIS on the triangular domain and interpolation points on the boundary are coplanar. Later, Geronimo and Hardin [62] constructed the FIS on a polygonal domain with arbitrary interpolation points. Zhao [151] gave two algorithms that generalize the construction of Geronimo and Hardin. Xie and Sun [149] constructed a bivariate FIS on rectangular grids with arbitrary boundary data and without restriction on the scaling factors. Dalla [51] constructed FIS using collinear boundary points and proved that the attractor is a continuous FIS. Also, various kinds of FISs by various authors can be found in [24–26, 29, 58].

1.6 Shape Preserving FIF and FIS

When data arises from various natural and scientific phenomena, it may contains certain geometric properties which may be expressed mathematically in terms of positivity, monotonicity, convexity, etc. For example, positive data arises in monthly rainfall amounts, discharge of gas in certain chemical reactions etc. The negative graphical representation is meaningless for these physical quantities. Rate of dissemination of drug in blood [16], empirical option of pricing models in finance [69], approximation of couples and quasi couples in statistics [16] are some examples of monotonic data. Convexity plays vital role in the theory of nonlinear optimization. The problem of searching sufficiently smooth functions that preserve these shape properties hidden in the data set is generally referred to as shape preserving interpolation.

Definition 1.6.1. *The univariate data* $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ *is said to be positive data if* $y_i > 0$ *for* $i = 1, 2, \ldots, N$.

Definition 1.6.2. *The univariate data* $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ *is said to be monotonically decreasing data if* $y_i \geq y_{i+1}$ *for* $i = 1, 2, \ldots, N-1$ *and monotonically increasing data if* $y_i \leq y_{i+1}$ *for* $i = 1, 2, \ldots, N-1$.

Definition 1.6.3. *The univariate data* $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ *is said to be convex data if* $\Delta_1 \leq \Delta_2 \leq \ldots \leq \Delta_{i-1} \leq \Delta_i \leq \Delta_{i+1} \leq \ldots \leq \Delta_{N-1}$ *where* $\Delta_i = \frac{y_{i+1}-y_i}{h_i}$ *and* $h_i = x_{i+1} - x_i$ *for* $1 \leq i \leq N-1$.

The following result is used to show the interpolantion function is convex in Chapters 2 and 3:

Three Chords Lemma: Let f be a convex function on the domain I and x^*, y^*, z^* $\in I$, where $x^* < y^* < z^*$. Then,

$$\frac{f(y^*) - f(x^*)}{y^* - x^*} \leq \frac{f(z^*) - f(x^*)}{z^* - x^*} \leq \frac{f(z^*) - f(y^*)}{z^* - y^*}.$$

Definition 1.6.4. *The bivariate data* $\{(x_i, y_j, z_{i,j}) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *is said to be positive surface data if* $z_{i,j} > 0$ *for* $i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N$.

Definition 1.6.5. *The bivariate data* $\{(x_i, y_j, z_{i,j}) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *is said to be monotonic surface data if monotonic in x-direction, i.e.,* $z_{i,j} \leq z_{i+1,j}$ $(z_{i,j} \geq z_{i+1,j})$ *for* $i = 1, 2, \ldots, M-1,\ j = 1, 2, \ldots, N$ *and monotonic in y-direction, i.e.,* $z_{i,j} \leq z_{i,j+1}$ $(z_{i,j} \geq z_{i,j+1})$ *for* $i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N-1$.

Definition 1.6.6. *The bivariate data* $\{(x_i, y_j, z_{i,j}) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *is said to be convex surface data if* $\Delta_{1,j} \leq \Delta_{2,j} \leq \ldots \leq \Delta_{i-1,j} \leq \Delta_{i,j} \leq \Delta_{i+1,j} \leq \ldots \leq \Delta_{M-1,j}$, $j = 1, 2, \ldots, N$, *and* $\Delta^1_{i,1} \leq \Delta^1_{i,2} \leq \ldots \leq \Delta^1_{i,j-1} \leq \Delta^1_{i,j} \leq \Delta^1_{i,j+1} \leq \ldots \leq \Delta^1_{i,N-1}$, $i = 1, 2, \ldots, M$, *where* $\Delta_{i,j} = \frac{z_{i+1,j}-z_{i,j}}{h_i}$, $h_i = x_{i+1} - x_i$, $\Delta^1_{i,j} = \frac{z_{i,j+1}-z_{i,j}}{h^*_j}$, $h^*_j = y_{j+1} - y_j$.

In order to preserve the shapes of the univariate data, various non-recursive traditional interpolation methods and fractal interpolation methods are available. We first present a brief survey on the univariate shape preserving interpolation. Construction of shape preserving interpolation was initiated by Schweikert [128] and he introduced tension splines through the solutions of suitable differentiable equations. Carl de Boor and Swartz [53] proved a number of theorems concerning the piecewise monotone interpolation of data by splines. Passow and Roulier [103] considered the possibility of a spline interpolant of pre-determined smoothness which is monotone or convex for a monotone and convex data respectively. Costantini [48] considered the problem of existence of monotone or convex splines that having degree n and order of continuity k. Also, the interpolating splines are obtained by using Bernstein polynomials. Fritsch and Carlson [61] proposed an algorithm which constructs a $\mathcal{C}^1$- monotone piecewise cubic interpolant to a monotonic data. Fritsch and Butland [60] described a method for producing piecewise cubic interpolant to a monotonic data. Also, the method is local and extremely simple to implement. Schumaker [127] designed an algorithm to construct $\mathcal{C}^1$-quadratic splines in such a way that monotonicity or convexity of the data is preserved. Lamberti [86] described a global method for construction of a $\mathcal{C}^2$-shape preserving interpolating function based on parametric cubic curve. He used step length as tension parameters and he selected tension parameters suitably so that the interpolation function preserves the positivity, monotonicity and convexity of the data. Also, we can see various types of shape preserving methods were developed by different authors in [49, 78, 79, 122].

The aforementioned shape preserving methods were developed using polynomials or tension splines. Alternative to the polynomial or tensional spline interpolation is rational spline interpolation and it was introduced by Späth [131]. Gregory and Delbourgo [54, 63] constructed rational quadratic splines which does not involve shape parameters to preserve the shapes of the monotonic data. By introducing shape parameter on each interval, Gregory and Delbourgo [55] have developed the rational cubic spline. Shape of the interpolation curves can be controlled by these shape parameters and shapes of the data can be preserved by selecting suitable choices of parameters. Recently, huge number of shape preserving rational interpolants were developed, for example [1, 123]. Using fractal methodology, variety of rational FIFs with shape parameters were developed by Chand and co-authors, to preserve the shapes of the univariate data [34, 142–146].

In order to preserve the shapes of the bivariate data, variety of numerical schemes have been developed. For example, Carlson and Fritsch [27] constructed a monotone $\mathcal{C}^1$-piecewise bicubic spline on a rectangular grid and the interpolant is determined by the first partial derivatives and mixed partial derivatives at the mesh points. Beatson and Ziegler [15] proposed an algorithm for $\mathcal{C}^1$-quadratic spline surface to preserve the monotonicity of the data that lies on the rectangular grid. Renka [119] developed an algorithm for construction of $\mathcal{C}^1$-convex surface that interpolates a convex data. Zhang et al. [150] constructed bivariate rational interpolation surface based on function values to preserve the convexity. Kouibia and Pasadas [85] presented an approximation problem of parametric curves and surfaces from a Lagrange or Hermite data set. Also, one can see in references [28, 50, 68], a variety of numerical schemes have been developed to preserve the shapes of the bivariate data. Also, using fractal techniques, Chand et al. [32, 33, 35, 37, 38, 141] constructed different kinds of FISs in the domain of shape preserving.

In this thesis, $\mathcal{C}^1$-FIFs and $\mathcal{C}^1$-FISs are constructed using the IFS that involves rational functions. To construct FIFs or FISs, derivative values or partial derivative values at the knot points are needed. Many situations derivative values or partial derivative values may not be supplied and only data points are available. In that situations estimation of the derivative values or partial derivative values are necessary. To compute derivative values, the arithmetic mean method [34] is commonly used and they are given by the following: At the interior point x_i, $i = 2, 3, \dots, N-1$, set

$$d_i = \begin{cases} 0, & \text{if } \Delta_i = 0 \text{ or } \Delta_{i-1} = 0, \\ \frac{h_i \Delta_{i-1} + h_{i-1}\Delta_i}{h_i + h_{i-1}}, & \text{otherwise.} \end{cases}$$

At the end points x_1 and x_N, set

$$d_1 = \begin{cases} 0, & \text{if } \Delta_1 = 0 \text{ or } sgn(d_1^*) \neq sgn(\Delta_1), \\ d_1^* = \Delta_1 + \frac{(\Delta_1 - \Delta_2)h_1}{h_1 + h_2}, & \text{otherwise,} \end{cases}$$

$$d_N = \begin{cases} 0, & \text{if } \Delta_{N-1} = 0 \text{ or } sgn(d_N^*) \neq sgn(\Delta_{N-1}), \\ d_N^* = \Delta_{N-1} + \frac{(\Delta_{N-1} - \Delta_{N-2})h_{N-1}}{h_{N-1} + h_{N-2}}, & \text{otherwise.} \end{cases}$$

To compute the partial derivatives, we used the arithmetic mean method [68] for the bivariate data as follows: For each fixed $j = 1, 2, \ldots, N$, we compute the x-direction partial derivatives as

$$z^x_{1,j} = \Delta_{1,j} + (\Delta_{1,j} - \Delta_{2,j})\frac{h_1}{h_1 + h_2}, z^x_{i,j} = \frac{\Delta_{i,j} + \Delta_{i-1,j}}{2}, \; i = 2, 3, \ldots, M-1,$$
$$z^x_{M,j} = \Delta_{M-1,j} + (\Delta_{M-1,j} - \Delta_{M-2,j})\frac{h_{M-1}}{h_{M-1} + h_{M-2}}.$$

For each fixed $i = 1, 2, \ldots, M$, we compute the y-direction partial derivatives as

$$z^y_{i,1} = \Delta^1_{i,1} + (\Delta^1_{i,1} - \Delta^1_{i,2})\frac{h^*_1}{h^*_1 + h^*_2}, z^y_{i,j} = \frac{\Delta^1_{i,j} + \Delta^1_{i,j-1}}{2}, \; j = 2, 3, \ldots, N-1,$$
$$z^y_{i,N} = \Delta^1_{i,N-1} + (\Delta^1_{i,N-1} - \Delta^1_{i,N-2})\frac{h^*_{N-1}}{h^*_{N-1} + h^*_{N-2}}.$$

1.7 Spline Functions

The study of any physical phenomena is involved with two major tasks.

- Modelling: Mathematical formulation of a physical process,
- Analysis: Numerical analysis of the mathematical model.

The mathematical formulation of physical process results in mathematical statements, mostly in differential equations. In numerical analysis of physical process, it is attempted to solve the governing equations by some numerical methods and an analysis of numerical method is carried out. The numerical methods developed may broadly be classified into the following three types:

- Finite difference methods,
- Finite element methods,
- Spline function approximation methods.

We give a brief introduction about finite difference, finite element and spline function approximation methods.

Finite difference method

There are four main methods of deriving finite difference methods as given in [80].

- The replacement of each term of the differential operator by a Taylor series approximation.
- The integration of the differential equation over a finite difference block and the subsequent replacement of each term by a Taylor series approximation.
- Formulation of the problems in variational form and the subsequent replacement of each term of the variational formulation by a Taylor series approximation.
- Derivation of a finite difference equation whose solution is identical to that of the differential equation with constant coefficients.

The first three methods share a common defect, namely that the individual terms of differential operator are approximated in isolation from the remaining terms of the operator. Consequently, the interaction between the terms of the differential operator are ignored. This is a fundamental cause for the existence of instability in both ordinary and partial differential equations. The last method is called the unified difference representation. In this case the term interactions are included, and no possibility of instability can exist.

Finite element method

The finite element method is characterized by three features as stated in [118]:

- The domain of the problem is represented by the collection of subdomain, called finite elements. The collection of finite elements are called the finite element mesh.
- Over each finite element, physical process is approximated by the functions of desired type (polynomials or otherwise), and algebraic equations relating physical quantities at selective points, called nodes, of the elements are developed.

- The element equations are assembled using continuity and/or "balance" of physical quantities.

In general, in the finite element method, we seek an approximate solution u of the differential equation in the form

$$u \approx \sum_{j=1}^{n} u_j \psi_j + \sum_{j=1}^{m} c_j \phi_j$$

where u_j are the values of u at the element nodes, ψ_j are the interpolation functions, c_j are the nodeless coefficients and ϕ_j are the associated approximation functions. Direct substitution of such approximations into the governing differential equations does not always result, for an arbitrary choice of the data of problem, in a necessary and sufficient number of equations for the undetermined coefficients u_j and c_j. Therefore, a procedure where by a necessary and sufficient number of equations can be obtained is needed. One such procedure is provided by a weighted-integral form of the governing differential equations.

Spline function approximation method

Let $a = x_0 < x_1 < \cdots < x_N = b$, be a partition of an interval $[a, b]$. A spline function of degree m with nodes at the points x_j, $j = 0, 1, \ldots, N$, is a function S with the following properties:

- S is a polynomial of degree m in each subinterval $[x_j, x_{j+1}]$, $j = 0, 1, \ldots, N-1$.
- S and its first $(m-1)$ derivatives are continuous on $[a, b]$.

If a spline S has only $(m-k)$ continuous derivatives, then k is defined as the deficiency of the polynomial spline and is usually denoted by $S(m, k)$. Usually, spline function is a piecewise polynomial satisfying certain conditions of continuity of the function and its derivatives.

Solving governing differential equations using spline has more advantage than the finite difference and finite element methods. For example, cubic spline method for solving the second-order differential equation can be described as follows:

Let us consider the second-order differential equation of the form

$$\begin{cases} -u''(x) + q(x)u(x) = f(x), \ x \in (0,1), \\ -u'(0) = \eta_0, \ u'(1) = \eta_1. \end{cases}$$

Let us consider the uniform grid $0 = x_0 < x_1 < \cdots < x_N = 1$ on $[0,1]$ where $x_i = ih$, $h = 1/N$. Let U_i be the approximation of $u(x_i)$, U_i' be the approximation of $u'(x_i)$ and M_i be the approximation of $u''(x_i)$.

On $[x_{i-1}, x_i]$, we define

$$S(x) = A_i\Big(\frac{x - x_{i-1}}{h}\Big)^3 + B_i\Big(\frac{x - x_{i-1}}{h}\Big)^2 + C_i\Big(\frac{x - x_{i-1}}{h}\Big) + D_i.$$

In order to get the cubic spline that passes through the approximations of the differential equation, we evaluate the constants using the conditions $S(x_{i-1}) = U_{i-1}$, $S(x_i) = U_i$, $S''(x_{i-1}) = M_{i-1}$ and $S''(x_i) = M_i$. We get

$$A_i = \frac{h^2}{6}[M_i - M_{i-1}], \ B_i = \frac{h^2}{2}M_{i-1}, C_i = (U_i - U_{i-1}) - \frac{h^2}{3}M_{i-1} - \frac{h^2}{6}M_i, \ D_i = U_{i-1}.$$

It can be seen that the function S and its derivative S'' are continuous on $[0,1]$. To get the continuity of S' at interior nodes x_i, $i = 1, 2, \ldots, N-1$, we need $S'(x_i^-) = S'(x_i^+)$ which leads to

$$\frac{h^2}{6}M_{i-1} + \frac{2h^2}{3}M_i + \frac{h^2}{6}M_{i+1} = U_{i-1} - 2U_i + U_{i+1}. \tag{1.5}$$

At end point x_0 and x_N, the derivative S' satisfies

$$S'(x_0) = \frac{1}{h}(U_1 - U_0) - \frac{h}{3}M_0 - \frac{h}{6}M_1, \tag{1.6}$$

$$S'(x_N) = \frac{1}{h}(U_N - U_{N-1}) + \frac{h}{6}M_{N-1} + \frac{h}{3}M_N. \tag{1.7}$$

Now the differential equation is discretized at $x = x_i$ as $-M_i + q(x_i)U_i = f(x_i)$. The boundary conditions are discretized as $-U_0' = \eta_0$ and $U_N' = \eta_1$. Now substituting

$$M_i = q(x_i)U_i - f(x_i), \ S'(x_0) = U_0' = -\eta_0, \ S'(x_N) = U_n' = \eta_1$$

in (1.5)-(1.7), we obtain the system of $(N+1)$ equations with $(N+1)$ unknowns namely U_0, U_1, $\ldots$, U_N. By solving the system, we get the approximate solution

$U = (U_0, U_1, \ldots, U_N)$ of the differential equation. Once we get the approximate solution, we can compute the values M_i as $M_i = q(x_i)U_i - f(x_i)$, $i = 0, 1, \ldots, N$.

Since the values U_i and M_i are known, now we can evaluate the coefficients A_i, B_i, C_i and D_i. We obtain the cubic spline that passes through the solutions of the differential equation. Thus, the solution of the differential equation u is approximated by a cubic spline.

Some important features of the spline approximation methods are as follows:

- The solution of the differential equation u is approximated by different polynomial in each interval and hence we get not only the approximation of u but also for the derivatives of u at every point of the interval $[a, b]$. But in finite difference method, we get the approximation of u at the finite set of grid points only in $[a, b]$.

- Derivative boundary conditions can in many cases be applied more accurately and with less difficulty than with conventional finite difference schemes. Derivative boundary conditions are imposed directly without incurring large local discretization errors.

- Unlike finite element method, the evaluation of large numbers of quadratures is unnecessary.

In this thesis spline function approximation methods, particularly fractal spline approximation methods have been used to obtain the numerical solutions of ordinary differential equations. The use of spline functions dates back at least to the beginning of this century. Piecewise linear functions had been used already in connection with Peano's existence proof for solution to the initial-value problem of ordinary differential equations, although these functions were not called splines [95]. Spline functions was first introduced by Schoenberg [126]. The applications of spline as interpolating, approximating and curve fitting have been very successful [3,64,95,105]. It is also interesting to note that the cubic spline is a close mathematical approximation to the draughtsman's spline which is widely used manual curve drawing tool.

Loscalzo and Talbot [89] constructed the approximate solution of the initial-value problem through spline function of degree m. Also this approach is equivalent to the multi-step method. Trapezoidal rule and the Milne-Simpson methods are the special cases of the constructed method. Spline functions have been used by a number of authors to solve initial-value problems of ordinary differential equations, for example one can see the references [94, 101, 102]. Bickley [21] suggested the method of solving linear two-point boundary-value problems (BVPs) through cubic spline. His main idea was to use the "conditions of continuity" as a discretization equation for the linear two-point BVPs. Later, Fyfe [6] discussed the application of deferred corrections to the method suggested by Bickley by considering linear BVPs.

However, for the two-point BVPs, the discretization given by the continuity conditions of cubic spline are only the second-order discretization. Hence the cubic spline method is only second-order. But the cubic spline interpolation process itself fourth-order. In order to get higher-order scheme for solving two-point linear BVPs, non-polynomial spline that containing a parameter $\mu > 0$ (say) can be considered. These non-polynomial splines can be defined through the solution of a differential equation in each subinterval. When $\mu \to 0$, non-polynomial splines reduces into polynomial splines.

For example, non-polynomial cubic spline has the basis $\{1, x, \cos \mu x, \sin \mu x\}$ or $\{1, x, e^{-\mu x}, e^{\mu x}\}$ where μ is the frequency of trigonometric part of the spline function which can be real or pure imaginary. The parameter μ will be used to raise the accuracy of the method. It can be seen that

$$\begin{aligned}\lim_{\mu \to 0} \operatorname{span}\{1, x, \cos \mu x, \sin \mu x\} &= \lim_{\mu \to 0} \operatorname{span}\Big\{1, x, \frac{2}{\mu^2}[\cos(\mu x) - 1], \frac{6}{\mu^3}[\mu x - \sin(\mu x)]\Big\} \\ &= \lim_{\mu \to 0} \operatorname{span}\{1, x, x^2, x^3\}.\end{aligned}$$

In order to get the numerical solutions for the BVPs of ordinary differential equations, various kinds of splines and non-polynomial splines have been used. For example, De Boor [52] and Sakai [121] used cubic spline to get the numerical approximations of the BVPs. Ahlberg [3] used cardinal splines to solve the linear BVPs and nonlinear BVPs. Tewarson [133] used cubic and quintic splines to obtain the numerical solutions for the non-linear BVPs. Russell and Shampine [120] used collocation procedure based on piecewise polynomial functions for solving BVPs. Usmani and Sakai [138] used quartic splines to solve the second-order BVPs and the proposed method has fourth-order convergence. A cubic spline function has been applied to the solution of second-order nonlinear two-point BVPs with significant first derivatives by Jain and Aziz [72]. Iyengar and Jain [71] developed the numerical method using splines to solve the singular BVPs and the proposed method has second-order convergence. Chawla and Subramanian [45] proposed the numerical method using quintic spline to solve the fourth-order BVPs and proposed method has sixth-order convergence. Irodotou-Ellina and Houstis [70] constructed an sixth-order convergent collocation method for general fourth-order linear two-point BVPs. Ramadan et al. [106] developed the sixth-order convergent method using non-polynomial quintic spline to get the numerical approximations for the BVPs. Khan et al. [83] used non-polynomial sextic spline to get the numerical solutions for the fifth-order BVPs. Also one can see the References [81, 82, 107, 130].

In this thesis we have proposed the numerical methods using fractal cubic splines, fractal non-polynomial cubic splines and fractal quintic splines to solve the various types of the BVPs namely singularly perturbed BVPs, singular BVPs, non-linear BVPs and fourth-order BVPs. The theorems which have been used in this thesis are given as follows:

Let W be set of the first n integers, i.e., $W = \{1, 2, \ldots, n\}$.

Theorem 1.7.1. *A matrix A of order $n \geq 2$ is irreducible if and only if for any two integers i and j, $i \in W$, $j \in W$, there exists a sequence of nonzero elements of A of the form*

$$\{a_{i,i_1}, a_{i_1,i_2}, \ldots, a_{i_{m-1},j}\}. \tag{1.8}$$

A sequence of nonzero elements of A of the form (1.8) is called a *chain*. If $a_{i,j} \neq 0$, the chain (1.8) may be assumed to consist of the one element $a_{i,j}$.

Theorem 1.7.2. *A matrix A is monotone if and only if the elements of the inverse matrix A^{-1} are nonnegative.*

Theorem 1.7.3. *Let the matrix $A = (a_{i,j})$ be irreducible and satisfy the conditions*

- $a_{i,j} \leq 0$, $i \neq j$; $i, j = 1, 2, \ldots, n$,

- $\sum_{j=1}^{n} a_{i,j} \begin{cases} \geq 0, & i = 1, 2, \ldots, n, \\ > 0, & \textit{for at least one } i. \end{cases}$

Then A is monotone.

Theorem 1.7.4. *[18] Let A_1 and A_2 be any two $n \times n$ matrices. Let $\|.\|$ be any operator norm on the space of the matrices. Then eigenvalues of A_1 and A_2 can be enumerated as $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $\mu_1, \mu_2, \ldots, \mu_n$ in such a way that*

$$\max_i |\lambda_i - \mu_i| \leq 2^{\frac{2n-1}{n}} n^{\frac{1}{n}} (2Q)^{\frac{n-1}{n}} \|A_1 - A_2\|^{\frac{1}{n}}$$

where $Q = \max\{\|A_1\|, \|A_2\|\}$.

1.8 Motivation of Present Work

Interpolation deals with the problem of reconstructing an unknown function in a continuum from its availability in some grid points. In general, a traditional non-recursive interpolation scheme produces an interpolant with infinitely differentiable or piecewise infinitely differentiable. Many of the situations data comes from a function whose derivative of certain order may be irregular and giving smooth or piecewise smooth representations for these data may not be suitable.

Fractal interpolation is a recent technique that generalizes the traditional interpolation methods. Using IFS, Barnsley [10] constructed the FIF that is used to approximate the functions in nature which display some kind of geometrical complexity under magnification. Fractal interpolation provides an efficient method to model an experimental or a natural data set using a smooth or non-smooth function depending on the applications. FIFs are constructed as graphs of the attractors of specific IFSs. Fractal interpolants that are non-smooth can represent the rough aspect of real world signals, and hence it has applications in areas such as multiwavelets, computer graphics, financial series, acoustics, and sociology. If the situation demands a smooth interpolant, then the parameters of the corresponding IFS can be chosen appropriately to construct smooth fractal interpolants.

In many situations, the data set possesses certain shape properties and also data representing a certain derivative may be irregular. There are classical interpolation methods that are well-suited for the shape preserving interpolation problems but not well-suited for the irregular representation of a certain derivative of the unknown function. In contrast to traditional interpolation, by controlling the scaling factors inherent with the fractal splines, we can get a certain derivative of a fractal splines with non-differentiability in a finite or a dense subset of the interval. Also fractal splines can be used effectively in the shape preserving interpolation problems.

The motivation of this work is to construct shape preserving interpolants for the univariate and bivariate data. In this thesis, we introduced new fractal interpolants using rational splines to preserve the positivity, monotonicity and convexity nature of the univariate and the bivariate data sets.

Apart from using fractal splines in the field of shape preserving interpolation problems, we have used fractal splines to get the numerical solutions of two-point BVPs. Though finite difference discretizations are invariably used, one important advantage of solving BVPs using splines is that once the spline solution has been computed, the information required for spline interpolation between the mesh points is already available. This is particularly significant when the solution of the BVPs is required at various locations of the interval. An important instance also is the use of an automatic plotter that frequently requires interpolation at great many intermediate points [43]. The order of convergence of the numerical method developed by the fractal splines and the order of convergence of the numerical method developed by the corresponding non-recursive splines are same. By varying the scaling factors inherent in the numerical method developed by fractal splines, we can get infinitely many numerical solutions for the same BVP. In particular, when we take all the scaling factors are zero, the numerical method developed by the fractal splines reduces to numerical method of corresponding non-recursive spline.

1.9 Organization of Present Work

Chapter 1 consists of brief introduction about IFS, FIF, FIS and brief literature of shape preserving interpolation of univariate and bivariate data. Also, a brief literature review of the numerical approximation methods of BVPs of ordinary differential equations is given.

In **Chapter** 2, at beginning, we construct the rational cubic FIF that contains two families of shape parameters to interpolate the univariate data. For an interpolation method to be effective, the interpolant corresponding to a data set should converge to the corresponding data defining function. For this purpose, we have derived the uniform error bound between the FIF and the original function of the class $\mathcal{C}^3$ that defines the data set. To preserve the geometric properties of the data, namely, positivity, monotonicity and convexity, the scaling factors and shape parameters are restricted so that constructed FIF would preserve these properties. Also, constrained interpolation problem of the developed FIF is also discussed. When all the scaling factors are zero, the fractal interpolation method reduces into the method developed in [124].

Next, we provide the construction of rational cubic FIF that contains three families of shape parameters to interpolate the univariate data. To study the effectiveness of the constructed FIF, the uniform error bound between the constructed FIF and the original function that belongs to the class $\mathcal{C}^2$ is derived. The scaling factors and shape parameters are constrained so that the constructed FIF preserves the geometric natures of the univariate data namely positivity, monotonicity and convexity. The presented shape preserving interpolation scheme generalizes the traditional shape preserving interpolation scheme studied in [1].

In **Chapter** 3, to interpolate the bivariate data that lies on the rectangular grid, a FIS is constructed. First, the fractal boundary curves are constructed along the grid lines. Then with the help these fractal boundary curves and the blending functions, the FIS is constructed. On each rectangular patch, the constructed FIS contains four scaling factors and 12 shape parameters. The scaling factors and shape parameters are restricted to get the positivity, monotonicity and convexity preserving FIS. Also the shape parameters and scaling factors are restricted so that constructed FIS would lie above the plane whenever the surface data lies above the plane. Also to see the effectiveness of the constructed FIS, the uniform error bound between the FIS and the original function that belongs to the class $\mathcal{C}^4$ is developed.

In **Chapter 4**, we consider the self-adjoint singularly perturbed BVPs of the form

$$\begin{cases} -\varepsilon u''(x) + q(x)u(x) = f(x), \ x \in (0,1), \\ u(0) = \eta_0, \ u(1) = \eta_1, \end{cases}$$

and

$$\begin{cases} -\varepsilon u''(x) + q(x)u(x) = f(x), \ x \in (0,1), \\ -u'(0) = \eta_0, \ u'(1) = \eta_1, \end{cases}$$

where $0 < \varepsilon \leq 1$, q and f are sufficiently smooth functions in $[0,1]$ and $q(x) > 0$, $x \in [0,1]$. These BVPs exhibit boundary layer at both ends of the interval.

Continuity conditions of the fractal cubic spline are used to obtain the numerical methods for these BVPs. To see the efficiency, error analysis of the developed methods are established. The discretized equations given by continuity conditions are second-order and hence the resulting fractal cubic spline methods are also second-order. The numerical results are tabulated and compared with the numerical results of the cubic spline method.

In **Chapter 5**, we consider the nonself-adjoint singularly perturbed BVPs of the form

$$\begin{cases} \varepsilon u''(x) = p(x)u'(x) + q(x)u(x) + f(x), \ x \in (0,1), \\ u(0) = \eta_0, \ u(1) = \eta_1, \end{cases}$$

where $0 < \varepsilon \leq 1$, p, q, f are sufficiently smooth functions, $p(x) > 0$ or $p(x) < 0$, $q(x) > 0$ in $[0,1]$. Solutions of these BVPs exhibit boundary layer at $x = 0$ or $x = 1$ if $p(x) < 0$ or $p(x) > 0$ respectively. We used the continuity conditions of the fractal cubic spline to get the numerical method for these BVPs and convergence analysis of the developed method is established and it is shown that the proposed method has second-order convergence.

Next, we consider the linear singular BVPs of the form

$$\begin{cases} u''(x) + \dfrac{k}{x}u'(x) - q(x) = f(x),\ x \in (0,1), \\ u'(0) = 0,\ u(1) = \eta_1. \end{cases}$$

where $k = 1, 2$, the functions q, f are sufficiently smooth, $q(x) > 0$ in $[0, 1]$. This problem has singularity at $x = 0$. Second-order convergent numerical method obtained via fractal cubic spline has been used to get the numerical solutions for these BVPs.

Next, we consider the non-linear singular BVPs of the form

$$\begin{cases} u''(x) + \dfrac{k}{x}u'(x) = F(x, u(x)),\ x \in (0,1), \\ u'(0) = 0,\ u(1) = \eta_1, \end{cases}$$

where $k = 1, 2$. Also for $(x, u(x)) \in D = \{0 \leq x \leq 1,\ -\infty < u(x) < \infty\}$, the functions F, $\partial F/\partial u$ are continuous, $\partial F/\partial u \geq 0$ on D and $\partial F/\partial u > 0$ on $D^\circ = \{0 < x < 1,\ -\infty < u(x) < \infty\}$. Fractal cubic spline and quasi-linearization technique are used to get the numerical approximations for these BVPs. Convergence analysis of the proposed methods are established and it tells that the constructed method has second-order convergence.

In **Chapter 6**, we consider the self-adjoint singularly perturbed BVPs and the nonself-adjoint singularly perturbed BVPs. The method developed to get the numerical approximations for the self-adjoint singularly perturbed BVPs in **Chapter 4** and the method proposed in **Chapter 5** to get the numerical solutions for the nonself-adjoint singularly perturbed BVPs are second-order convergent. In order to obtain higher-order convergent methods, in this chapter we have used the fractal non-polynomial cubic spline to get the numerical solutions for these BVPs. The continuity conditions of the fractal non-polynomial cubic spline are used to construct the numerical method for these BVPs. Convergence analysis of the developed methods are carried out and it is shown that the proposed methods have fourth-order convergence.

In **Chapter 7**, at the beginning, we consider the BVPs of the form

$$\begin{cases} -\varepsilon u''(x) + q\, u(x) = f(x),\ x \in (0,1), \\ u(0) = \eta_0,\ u(1) = \eta_1, \end{cases}$$

where $0 < \varepsilon \leq 1$, $q > 0$, f is sufficiently smooth function in $[0, 1]$. Continuity conditions of the fractal quintic spline are used to get the numerical scheme and the developed numerical scheme has fourth-order convergence.

Next, we consider the BVP of the form

$$\begin{cases} u''(x) + F(x, u(x)) = 0, \ x \in (0, 1), \\ u(0) = \eta_0, \ u(1) = \eta_1. \end{cases}$$

We assume that for $(x, u(x)) \in D = \{0 \leq x \leq 1, \ -\infty < u(x) < \infty\}$, the functions F and $\partial F/\partial u$ are continuous. This problem possess a unique solution provided $\sup_{(x,u)\in D} \partial F/\partial u < \pi^2$ [42]. We assume that $\partial F/\partial u \leq 0$ on D and $\partial F/\partial u < 0$ on $D^\circ = \{0 < x < 1, \ -\infty < u(x) < \infty\}$. Using quasilinearization technique, we convert the non-linear BVP into sequence of linear BVPs. Then each of these linear BVPs have been solved using the method obtained from fractal quintic spline. The developed numerical method has fourth-order convergence.

In **Chapter 8**, we consider the BVPs of the forms

$$\begin{cases} u^{(4)}(x) + q(x)u(x) = f(x), \ x \in (0, 1), \\ u(0) = \eta_0, \ u(1) = \eta_1, \ u''(0) = \hat{\eta}_0, \ u''(1) = \hat{\eta}_1, \end{cases}$$

and

$$\begin{cases} u^{(4)}(x) + q(x)u(x) = f(x), \ x \in (0, 1), \\ u(0) = \eta_0, \ u(1) = \eta_1, \ u'(0) = \hat{\eta}_0, \ u'(1) = \hat{\eta}_1, \end{cases}$$

where the functions q and f are continuous in $[0, 1]$. Continuity conditions of the fractal quintic spline are used to develop the numerical methods for these problems. Truncation errors corresponding to these methods are derived. The developed methods have second-order convergence.

Chapter 2

Shape Preserving Rational Cubic Fractal Interpolation Functions

In this chapter, new FIFs in the field of shape preserving interpolation are constructed. The developed FIFs are $\mathcal{C}^1$-continuous functions and they can be utilized for preserving all the three fundamental shape properties namely positivity, monotonicity and convexity. The derivatives of the developed FIFs having irregularity in a finite subset or a dense subset of the interpolation interval. The proposed schemes have some interesting features.

- The proposed methods are best tool to approximate a function which is continuous and its derivatives are irregular.
- When all the scaling factors are zero, FIFs obtained from proposed methods, reduces into a classical rational cubic splines.
- The proposed methods are equally applicable for the data with derivatives or data without derivatives.
- In the developed methods, extra knots are not needed to get the shape preserving interpolants.
- In the proposed schemes, shape preserving fractal interpolant is unique for the fixed scaling factors and the fixed shape parameters. By changing the scaling factors

and the shape parameters, infinitely many shape preserving fractal interpolants can be obtained.

The FIF with two families of shape parameters is constructed and shape preserving aspects of the constructed FIF are analyzed in Section 2.1. To construct the FIF, we consider the IFS that contains the rational functions where the numerator of the rational functions contain cubic polynomials and the denominator of rational functions contain the preassigned quadratic polynomials each containing 2 shape parameters. When all the scaling factors are zero, the developed rational cubic FIF reduces to the classical rational cubic interpolant introduced by Sarfraz et al. [124]. The convergence analysis of the constructed FIF to an original function which belongs to the class $\mathcal{C}^3$ is derived. Sufficient conditions on the shape parameters and the scaling factors are derived to preserve the positivity, monotonicity and convexity nature of a univariate data. Also, the constrained interpolation problem of the constructed FIF is discussed. We consider some numerical examples to test the applicability of the developed interpolation scheme.

In Section 2.2, to preserve the shapes of the univariate data, rational cubic FIF with three families of the shape parameters is constructed. The rational cubic FIF is constructed with the help of the IFS involving rational functions. The numerator of the rational functions contain cubic polynomials and the denominator of rational functions contain the preassigned quadratic polynomials each containing 3 shape parameters. To see effectiveness of the interpolantion scheme, error analysis of the constructed FIF to an original function that belongs to $\mathcal{C}^2$ is derived. Parameters are constrained so that the constructed FIF would preserve the positivity, monotonicity and convexity of a prescribed set of data points. The shape preserving aspects of the FIF are implemented through numerical examples. We get the classical rational cubic interpolant introduced by Abbas et al. [1] when all the scaling factors are zero.

2.1 FIF with Two Families of Shape Parameters

In this section, a new rational cubic FIF with two families of the shape parameter is constructed. Let $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ be a univariate data such that $x_1 < x_2 <$

$\cdots < x_N$. Let d_i be the derivative value at the knot x_i. Consider the IFS given in (1.4) with

$$r_i(x) = \frac{p_i(x)}{q_i(x)} \equiv \frac{P_i(\theta)}{Q_i(\theta)} = \frac{A_{1,i}(1-\theta)^3 + A_{2,i}\theta(1-\theta)^2 + A_{3,i}\theta^2(1-\theta) + A_{4,i}\theta^3}{u_i + v_i\theta(1-\theta)}, \quad (2.1)$$

$\theta = (x - x_1)/(x_N - x_1)$, $x \in [x_1, x_N]$. Here, u_i and v_i are the shape parameters. It is assumed that $u_i > 0$ and $v_i > 0$ to avoid the singularity in the denominator.

Let $\mathscr{F} = \{\phi : I \to \mathbb{R} \mid \phi \in \mathcal{C}^1(I),\ \phi(x_1) = y_1,\ \phi(x_N) = y_N,\ \phi^{(1)}(x_1) = d_1$ and $\phi^{(1)}(x_N) = d_N\}$. Then $(\mathscr{F}, \rho)$ is a complete metric space, where ρ is the metric induced by the $\mathcal{C}^1$-norm $\|\phi\| = \|\phi\|_\infty + \|\phi^{(1)}\|_\infty$ on $\mathcal{C}^1(I)$. Define the Read-Bajraktarević operator on $(\mathscr{F}, \rho)$ as

$$T\phi(L_i(x)) = \alpha_i\phi(x) + r_i(x),\ x \in I,\ i = 1, 2, \ldots, N-1.$$

Let $|\alpha_i| < a_i$, $i = 1, 2, \ldots, N-1$. The operator T is a contraction and hence the operator T has a fixed-point Φ (say) which satisfies the functional equation

$$\Phi(L_i(x)) = \alpha_i\Phi(x) + r_i(x), \quad x \in I,\ i = 1, 2, \ldots, N-1. \quad (2.2)$$

The derivative $\Phi^{(1)}$ satisfies the functional equation

$$\Phi^{(1)}(L_i(x)) = \frac{\alpha_i\Phi^{(1)}(x) + r_i^{(1)}(x)}{a_i}, \quad x \in I,\ i = 1, 2, \ldots, N-1. \quad (2.3)$$

The constants $A_{1,i}$, $A_{2,i}$, $A_{3,i}$ and $A_{4,i}$ appearing in (2.1) are evaluated based on the interpolation conditions $\Phi(x_i) = y_i$, $\Phi(x_{i+1}) = y_{i+1}$, $\Phi^{(1)}(x_i) = d_i$ and $\Phi^{(1)}(x_{i+1}) = d_{i+1}$ (these conditions are equivalent to $F_i(x_1, y_1) = y_i$, $F_i(x_N, y_N) = y_{i+1}$, $F_{i,1}(x_1, d_1) = d_i$ and $F_{i,1}(x_N, d_N) = d_{i+1}$). Let $h_i = x_{i+1} - x_i$. By substituting $x = x_1$ in (2.2), we can evaluate the constant $A_{1,i}$ as

$$\Phi(L_i(x_1)) = \alpha_i\Phi(x_1) + \frac{A_{1,i}}{u_i}$$

and hence we have

$$A_{1,i} = u_i[y_i - \alpha_i y_1].$$

By substituting $x = x_N$ in (2.2), we can evaluate the constant $A_{4,i}$ as

$$A_{4,i} = u_i[y_{i+1} - \alpha_i y_N].$$

We can evaluate $A_{2,i}$ by substituting $x = x_1$ in (2.3)as

$$d_i = \frac{\alpha_i}{a_i} d_1 + \frac{A_{2,i} u_i - A_{1,i}(3u_i + v_i)}{h_i u_i^2}$$

and hence we have

$$A_{2,i} = (3u_i + v_i) y_i + u_i h_i d_i - \alpha_i [(3u_i + v_i) y_1 + u_i (x_N - x_1) d_1].$$

Similarly, substituting $x = x_N$ in (2.3), we can evaluate $A_{3,i}$ as

$$A_{3,i} = (3u_i + v_i) y_{i+1} - u_i h_i d_{i+1} - \alpha_i [(3u_i + v_i) y_N - u_i (x_N - x_1) d_N].$$

Thus, the rational cubic FIF Φ is given by

$$\Phi(L_i(x)) = \alpha_i \Phi(x) + \frac{P_i(\theta)}{Q_i(\theta)}, \tag{2.4}$$

where

$$\begin{aligned} P_i(\theta) = & (u_i[y_i - \alpha_i y_1])(1-\theta)^3 + (u_i[y_{i+1} - \alpha_i y_N])\theta^3 \\ & + ((3u_i + v_i) y_i + u_i h_i d_i - \alpha_i [(3u_i + v_i) y_1 + u_i (x_N - x_1) d_1])\theta(1-\theta)^2 \\ & + ((3u_i + v_i) y_{i+1} - u_i h_i d_{i+1} - \alpha_i [(3u_i + v_i) y_N - u_i (x_N - x_1) d_N])\theta^2(1-\theta) \\ Q_i(\theta) = & u_i + v_i \theta(1-\theta), \ \theta = (x - x_1)/(x_N - x_1), \ x \in [x_1, x_N], \ i = 1, 2, \ldots, N-1. \end{aligned}$$

Remark 2.1.1. *The graph of the rational cubic FIF (2.4) is the attractor of the following IFS:*

$$\{I \times \mathbb{R}; w_i(x, y) = (L_i(x), \ F_i(x, y)), \ i = 1, 2, \ldots, N-1\}$$

with

$$L_i(x) = a_i x + b_i, \ F_i(x, y) = \alpha_i y + \frac{P_i(\theta)}{Q_i(\theta)}, \tag{2.5}$$

where L_i, $P_i(\theta)$ and $Q_i(\theta)$ are given as in (2.4).

Remark 2.1.2. *If $\alpha_i = 0$ for all $i = 1, 2, \ldots, N-1$, then the rational cubic FIF given in (2.4) reduces to the classical rational cubic spline C as*

$$C(x) = \frac{U_i(\varphi)}{V_i(\varphi)}, \tag{2.6}$$

where

$$U_i(\varphi) = u_i y_i (1-\varphi)^3 + [u_i h_i d_i + y_i(3u_i + v_i)]\varphi(1-\varphi)^2$$
$$+ [-u_i h_i d_{i+1} + y_{i+1}(3u_i + v_i)]\varphi^2(1-\varphi) + u_i y_{i+1}\varphi^3,$$
$$V_i(\varphi) = u_i + v_i\varphi(1-\varphi),\ \varphi = \frac{x - x_i}{x_{i+1} - x_i},\ x \in [x_i, x_{i+1}].$$

Remark 2.1.3. *If $\alpha_i = 0,\ u_i = 1$ and $v_i = 0$ then the rational cubic FIF (2.4) reduces to the standard cubic Hermite spline.*

Remark 2.1.4. *Using the scaling factor α_i, shape parameters u_i and v_i, shape of the curve can be modified according to the user. In particular, the scaling factor α_i and the shape parameter v_i are playing vital role on visualizing shape of the data while u_i can take any positive value. When $\alpha_i \to 0$ and $v_i \to \infty$, the rational cubic FIF (2.4) converges to a straight line in $[x_i, x_{i+1}]$. To see this, rewrite the rational cubic FIF (2.4) in the following form*

$$\Phi(L_i(x)) = \alpha_i \Phi(x) + \left\{\left[(1-\theta)y_i + \theta y_{i+1} + \frac{R_{1,i}(\theta)}{Q_i(\theta)}\right] - \alpha_i\left[(1-\theta)y_1 + \theta y_N + \frac{R_{2,i}(\theta)}{Q_i(\theta)}\right]\right\},$$

where

$$R_{1,i}(\theta) = u_i h_i \theta(1-\theta)[(\Delta_i - d_{i+1})\theta + (d_i - \Delta_i)(1-\theta)],$$
$$R_{2,i}(\theta) = u_i\theta(1-\theta)[\{(y_N - y_1) - (x_N - x_1)d_N\}\theta$$
$$+ \{(x_N - x_1)d_1 - (y_N - y_1)\}(1-\theta)].$$

If $v_i \to \infty$, then Φ converges to affine FIF as

$$\Phi(L_i(x)) = \alpha_i\Phi(x) + (y_i - \alpha_i y_1)(1-\theta) + (y_{i+1} - \alpha_i y_N)\theta.$$

Also, if $v_i \to \infty$ and $\alpha_i \to 0$, then Φ converges to the straight line segment in the interval $[x_i, x_{i+1}]$, i.e.,

$$\Phi(L_i(x)) = y_i(1-\theta) + y_{i+1}\theta.$$

2.1.1 Convergence analysis

Let the data $\{(x_i, y_i, d_i) : i = 1, 2, \dots, N\}$ be generated from the function S which belongs to $\mathcal{C}^3$. In this section, the uniform error bound between the original function S and the rational cubic FIF Φ defined in (2.4) is estimated.

Theorem 2.1.1. *Let Φ be a rational cubic FIF given in (2.4) and C be a classical rational cubic spline given in (2.6) with respect to the data $\{(x_i, y_i, d_i) : i = 1, 2, \dots, N\}$ which is generated from an original function $S \in \mathcal{C}^3(I)$. Then*

$$\|S - \Phi\|_\infty \leq \frac{1}{(1 - |\boldsymbol{\alpha}|_\infty)}\Big[|\boldsymbol{\alpha}|_\infty\Big([M + \frac{h}{4}\overline{M}] + [M^* + \frac{|I|}{4}M_*]\Big)\Big] + \|S^{(3)}\|_\infty h^3 c^*,$$

where $M = \max\{|y_i| : i = 1, 2, \dots, N\}$, $\overline{M} = \max\{|d_i| : i = 1, 2, \dots, N\}$, $M^* = \max\{|y_1|, |y_N|\}$, $M_* = \max\{|d_1|, |d_N|\}$, $|I| = x_N - x_1$, $h = \max\{h_i : i = 1, 2, \dots, N-1\}$, $|\boldsymbol{\alpha}|_\infty = \max\{|\alpha_i| : i = 1, 2, \dots, N-1\}$,

$$\sigma_1(u_i, v_i, \varphi) = -\frac{\varphi^3}{3} + \frac{(1-\varphi)^3\varphi^2\{(v_i + \varphi(u_i - v_i)\}}{3V_i(\varphi)} + \frac{8u_i^2(1-\varphi)^3\varphi^2}{3V_i(\varphi)[(2u_i + v_i)(1-\varphi) + u_i]^2},$$

$$\sigma_2(u_i, v_i, \varphi) = -\frac{\varphi^3}{3} + \frac{\varphi^2\{(3u_i + v_i) - \varphi(2u_i + v_i) - 3u_i(1-\varphi)\}}{3V_i(\varphi)}$$
$$+ \frac{2\varphi^3[\varphi(2u_i + v_i) - \sqrt{H_i}]^3}{3[u_i + \varphi(2u_i + v_i)]^3} + \frac{6u_i\varphi^2(1-\varphi)\{u_i + \varphi(u_i + v_i) - \varphi\sqrt{H_i}\}^2}{3V_i(\varphi)\{u_i + \varphi(2u_i + v_i)\}^2}$$
$$- \frac{2\varphi^2\{(3u_i + v_i) - \varphi(2u_i + v_i)\}\{u_i + \varphi(u_i + v_i) - \varphi\sqrt{H_i}\}^3}{3V_i(\varphi)\{u_i + \varphi(2u_i + v_i)\}^3},$$

$H_i = u_i(u_i + v_i) + \varphi v_i(2u_i + v_i)$, $\sigma(u_i, v_i, \varphi) = \begin{cases} \max \sigma_1(u_i, v_i, \varphi), & 0 \leq \varphi \leq \frac{u_i + v_i}{2u_i + v_i}, \\ \max \sigma_2(u_i, v_i, \varphi), & \frac{u_i + v_i}{2u_i + v_i} \leq \varphi \leq 1, \end{cases}$

$c_i^* := \max\{\sigma(u_i, v_i, \varphi) : 0 \leq \varphi \leq 1\}$, $i = 1, 2, \dots, N-1$ *and* $c^* = \max\{c_i^* : i = 1, 2, \dots, N-1\}$.

Proof. The Read-Bajraktarević operator corresponding to the rational cubic FIF Φ can be written as $T_{\boldsymbol{\alpha}} : \mathscr{F} \to \mathscr{F}$ such that

$$(T_{\boldsymbol{\alpha}}\phi)(x) = \alpha_i \phi(L_i^{-1}(x)) + \frac{p_i(L_i^{-1}(x), \alpha_i)}{q_i(L_i^{-1}(x))}, \quad x \in I_i,\ i = 1, 2, \dots, N-1,$$

where $p_i(x, \alpha_i) \equiv P_i(\theta)$ and $q_i(x) \equiv Q_i(\theta)$ are as given in (2.4). It is known that Φ is the fixed-point of the operator $T_{\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha} \neq \mathbf{0}$. Also, classical rational cubic spline C is the fixed-point of $T_{\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha} = \mathbf{0} = (0, 0, \dots, 0) \in \mathbb{R}^{N-1}$. Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_{N-1})$ be a scale vector such that $|\alpha_i| < a_i$ for all $i = 1, 2, \dots, N-1$ and with at least one $\alpha_i \neq 0$. For $\boldsymbol{\alpha} \neq \mathbf{0}$, $T_{\boldsymbol{\alpha}}$ is a contraction map with uniform metric. Hence

$$\|T_{\boldsymbol{\alpha}}\Phi - T_{\boldsymbol{\alpha}}C\|_\infty \leq |\boldsymbol{\alpha}|_\infty \|\Phi - C\|_\infty. \tag{2.7}$$

Also for $x \in I_i$, we have

$$\begin{aligned} |T_{\boldsymbol{\alpha}}C(x) - T_{\mathbf{0}}C(x)| &= \left| \alpha_i \, C \circ L_i^{-1}(x) + \frac{p_i(L_i^{-1}(x),\ \alpha_i)}{q_i(L_i^{-1}(x))} - \frac{p_i(L_i^{-1}(x),\ 0)}{q_i(L_i^{-1}(x))} \right| \\ &\leq |\alpha_i| \|C\|_\infty + \frac{|p_i(L_i^{-1}(x),\ \alpha_i) - p_i(L_i^{-1}(x),\ 0)|}{q_i(L_i^{-1}(x))}. \end{aligned} \tag{2.8}$$

Using the mean-value theorem for functions of several variables, there exists a $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_{N-1})$ such that $|\beta_i| < |\alpha_i|$ and

$$p_i(L_i^{-1}(x),\ \alpha_i) - p_i(L_i^{-1}(x),\ 0) = \left(\frac{\partial}{\partial \alpha_i} (p_i(L_i^{-1}(x),\ \beta_i)) \right) \alpha_i. \tag{2.9}$$

From (2.8) and (2.9), we can obtain that

$$|T_{\boldsymbol{\alpha}}C(x) - T_{\mathbf{0}}C(x)| \leq |\alpha_i| \left(\|C\|_\infty + \left| \frac{\partial}{\partial \alpha_i} \left(\frac{p_i(L_i^{-1}(x), \beta_i)}{q_i(L_i^{-1}(x))} \right) \right| \right). \tag{2.10}$$

To find the bound for the right hand side of (2.10), the classical rational cubic interpolant C can be written as

$$C(x) = w_1(u_i, v_i, \varphi) y_i + w_2(u_i, v_i, \varphi) y_{i+1} + w_3(u_i, v_i, \varphi) d_i - w_4(u_i, v_i, \varphi) d_{i+1}, \tag{2.11}$$

where

$$w_1(u_i, v_i, \varphi) = \frac{u_i(1-\varphi)^3 + (3u_i + v_i)\varphi(1-\varphi)^2}{V_i(\varphi)}, \quad w_3(u_i, v_i, \varphi) = \frac{u_i h_i \varphi (1-\varphi)^2}{V_i(\varphi)},$$

$$w_2(u_i, v_i, \varphi) = \frac{(3u_i + v_i)\varphi^2(1-\varphi) + u_i\varphi^3}{V_i(\varphi)} \quad \text{and} \quad w_4(u_i, v_i, \varphi) = \frac{u_i h_i \varphi^2 (1-\varphi)}{V_i(\varphi)}.$$

It can be observed that

$$w_1(u_i, v_i, \varphi) + w_2(u_i, v_i, \varphi) = 1.$$

Also, we get

$$w_3(u_i, v_i, \varphi) + w_4(u_i, v_i, \varphi) = \frac{u_i h_i \varphi (1-\varphi)^2 + u_i h_i \varphi^2 (1-\varphi)}{u_i + v_i \varphi (1-\varphi)}$$

$$\leq \frac{u_i h_i \varphi(1-\varphi)^2 + u_i h_i \varphi^2(1-\varphi)}{u_i} = h_i \varphi(1-\varphi).$$

From (2.11), we obtain

$$|C(x)| \leq \max\{|y_i|, |y_{i+1}|\} + \frac{h_i}{4}\max\{|d_i|, |d_{i+1}|\} = M_i + \frac{h_i}{4}\overline{M_i},$$

where $M_i = \max\{|y_i|, |y_{i+1}|\}$ and $\overline{M_i} = \max\{|d_i|, |d_{i+1}|\}$.

Thus, we have

$$\|C\|_\infty \leq M + \frac{h}{4}\overline{M}. \tag{2.12}$$

It can be noticed that

$$\frac{\partial}{\partial \alpha_i}\left(\frac{p_i(L_i^{-1}(x), \beta_i)}{q_i(L_i^{-1}(x))}\right) = - w_1^*(u_i, v_i, \varphi)y_1 - w_2^*(u_i, v_i, \varphi)y_N - w_3^*(u_i, v_i, \varphi)d_1 + w_4^*(u_i, v_i, \varphi)d_N,$$

where

$$w_1^*(u_i, v_i, \varphi) = \frac{u_i(1-\varphi)^3 + (3u_i + v_i)\varphi(1-\varphi)^2}{V_i(\varphi)},$$
$$w_2^*(u_i, v_i, \varphi) = \frac{(3u_i + v_i)\varphi^2(1-\varphi) + u_i\varphi^3}{V_i(\varphi)},$$
$$w_3^*(u_i, v_i, \varphi) = \frac{u_i(x_N - x_1)\varphi(1-\varphi)^2}{V_i(\varphi)} \quad \text{and}$$
$$w_4^*(u_i, v_i, \varphi) = \frac{u_i(x_N - x_1)\varphi^2(1-\varphi)}{V_i(\varphi)}.$$

Thus, we have

$$\left|\frac{\partial}{\partial \alpha_i}\left(\frac{p_i(L_i^{-1}(x), \beta_i)}{q_i(L_i^{-1}(x))}\right)\right| \leq |w_1^*(u_i, v_i, \varphi)y_1| + |w_2^*(u_i, v_i, \varphi)y_N| + |w_3^*(u_i, v_i, \varphi)d_1| + |w_4^*(u_i, v_i, \varphi)d_N|.$$

By using a similar procedure for finding the bound for $\|C\|_\infty$, it is easy to see that

$$\left|\frac{\partial}{\partial \alpha_i}\left(\frac{p_i(L_i^{-1}(x), \beta_i)}{q_i(L_i^{-1}(x))}\right)\right| \leq M^* + \frac{|I|}{4}M_*. \tag{2.13}$$

From (2.10), (2.12) and (2.13), we obtain

$$|T_{\boldsymbol{\alpha}}C(x) - T_{\mathbf{0}}C(x)| \leq |\boldsymbol{\alpha}|_\infty\Big([M + \frac{h}{4}\overline{M}] + [M^* + \frac{|I|}{4}M_*]\Big), \text{ for } x \in I_i.$$

Since this inequality does not depend on the interval, we get

$$\|T_{\boldsymbol{\alpha}}C - T_{\mathbf{0}}C\|_\infty \leq |\boldsymbol{\alpha}|_\infty\Big([M + \frac{h}{4}\overline{M}] + [M^* + \frac{|I|}{4}M_*]\Big). \tag{2.14}$$

From (2.7) and (2.14), we get

$$\|\Phi - C\|_\infty = \|T_{\boldsymbol{\alpha}}\Phi - T_{\mathbf{0}}C\|_\infty \leq \|T_{\boldsymbol{\alpha}}\Phi - T_{\boldsymbol{\alpha}}C\|_\infty + \|T_{\boldsymbol{\alpha}}C - T_{\mathbf{0}}C\|_\infty,$$

which implies that

$$\|\Phi - C\|_\infty \leq \frac{1}{(1 - |\boldsymbol{\alpha}|_\infty)}\Big[|\boldsymbol{\alpha}|_\infty\Big([M + \frac{h}{4}\overline{M}] + [M^* + \frac{|I|}{4}M_*]\Big)\Big]. \tag{2.15}$$

From [124], the error bound between the original function S and the classical rational cubic spline C is

$$\|S - C\|_\infty \leq \frac{1}{2}\|S^{(3)}\|_\infty h^3 c^*. \tag{2.16}$$

Using (2.15) and (2.16) with the following inequality

$$\|S - \Phi\|_\infty \leq \|S - C\|_\infty + \|C - \Phi\|_\infty$$

we get the required bound for $\|S - \Phi\|_\infty$. This completes the proof. □

Remark 2.1.5. *Since $|\alpha_i| < a_i = h_i/(x_N - x_1)$, $i = 1, 2, \dots, N-1$, from Theorem 2.1.1, it can be seen that $\|S - \Phi\|_\infty = \mathcal{O}(h)$.*

- *If $|\alpha_i| < a_i^2$, then $\|S - \Phi\|_\infty = \mathcal{O}(h^2)$.*
- *If $|\alpha_i| < a_i^3$, then $\|S - \Phi\|_\infty = \mathcal{O}(h^3)$.*

2.1.2 Shape preserving aspects of FIF

In this section, shape preserving aspects of the rational cubic FIF Φ are developed. Shape parameters and scaling factors are restricted suitably for rational cubic FIF to preserve the positivity, monotonicity and convexity. Also, parameters are restricted so that rational cubic FIF would lie above the line whenever the data lies above the line.

Positivity

Let $\{(x_i, y_i, d_i) : i = 1, 2, \dots, N\}$ be a positive data. In the following theorem, the sufficient conditions on the scaling factors and the shape parameters are obtained to ensure the positivity of rational cubic FIF. It can be seen that, the rational cubic FIF Φ is positive if $\Phi(x) > 0$ for all $x \in [x_1, x_N]$.

Theorem 2.1.2. *Let $\{(x_i, y_i) : i = 1, 2, \dots, N\}$ be a data such that $y_i > 0$, $i = 1, 2, \dots, N$. Let d_i, $i = 1, 2, \dots, N$ be the derivative value at the knot x_i. Then the following conditions on the scaling factors and the shape parameters are sufficient to the rational cubic FIF (2.4) to be positive.*

$$0 \leq \alpha_i < \min\Big\{a_i,\ \frac{y_i}{y_1},\ \frac{y_{i+1}}{y_N}\Big\},$$

$$u_i > 0 \text{ and } v_i > \max\Big\{0,\ \gamma_{1i},\ \gamma_{2i}\Big\},$$

where $\gamma_{1i} = \dfrac{-u_i h_i d_i + \alpha_i u_i (x_N - x_1) d_1}{y_i - \alpha_i y_1}$, $\gamma_{2i} = \dfrac{u_i h_i d_{i+1} - \alpha_i u_i (x_N - x_1) d_N}{y_{i+1} - \alpha_i y_N}$, $i = 1,\ 2,\ \dots,\ N-1$.

Proof. For each node x_j, $j = 1, 2, \dots, N$, we obtain

$$\Phi(L_i(x_j)) = \alpha_i \Phi(x_j) + \frac{P_i(\theta_j)}{Q_i(\theta_j)},\quad \theta_j = \frac{x_j - x_1}{x_N - x_1},\ i = 1, 2, \dots, N-1.$$

Assume that $\alpha_i \geq 0$, $i = 1, 2, \dots, N-1$. Also $u_i > 0$ and $v_i > 0$ gives $Q_i(\theta_j) > 0$. So $\Phi(L_i(x_j)) > 0$, $i = 1, 2, \dots, N-1$, $j = 1, 2, \dots, N$, if $P_i(\theta_j) > 0$. Now, we have

$$P_i(\theta_j) = A_{1,i}(1-\theta_j)^3 + A_{2,i}\theta_j(1-\theta_j)^2 + A_{3,i}{\theta_j}^2(1-\theta_j) + A_{4,i}{\theta_j}^3.$$

It can be seen that $P_i(\theta_j) > 0$ if $A_{1,i} > 0$, $A_{2,i} > 0$, $A_{3,i} > 0$ and $A_{4,i} > 0$. We get

$$A_{1,i} > 0 \text{ if } \alpha_i < \frac{y_i}{y_1} \text{ and } A_{4,i} > 0 \text{ if } \alpha_i < \frac{y_{i+1}}{y_N}.$$

Let $0 \leq \alpha_i < \Big\{\frac{y_i}{y_1},\ \frac{y_{i+1}}{y_N}\Big\}$. Then we have

$$A_{2,i} > 0 \quad \text{if } v_i > \frac{-u_i h_i d_i + \alpha_i u_i (x_N - x_1) d_1}{y_i - \alpha_i y_1}.$$

Also, we get

$$A_{3,i} > 0 \quad \text{if } v_i > \frac{u_i h_i d_{i+1} - \alpha_i u_i (x_N - x_1) d_N}{y_{i+1} - \alpha_i y_N}.$$

From the above conditions, it is clear that $\Phi(L_i(x_j)) > 0$ for all $i = 1, 2, \ldots, N-1$, $j = 1, 2, \ldots, N$ if the scaling factors and the shape parameters satisfy the sufficient conditions given in the statement of Theorem 2.1.2. Since the rational cubic FIF has recursive nature, therefore, $\Phi(L_i(x_j)) > 0$, for all $i = 1, 2, \ldots, N-1$, $j = 1, 2, \ldots, N$ which in turn gives $\Phi(x) > 0$ for all $x \in [x_1, \ x_N]$. □

Remark 2.1.6. *If $\alpha_i = 0$, $i = 1, 2, \ldots, N-1$, then the sufficient conditions for the classical rational spline C to be positive are*

$$u_i > 0 \text{ and } v_i > \max\left\{0, \frac{-u_i h_i d_i}{y_i}, \frac{u_i h_i d_{i+1}}{y_{i+1}}\right\}.$$

Monotonicity

Let $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ be a monotonic data. Let d_i be the derivative value at the knot x_i. Without loss of generality assume that the data is monotonically increasing i.e., $y_1 \leq y_2 \leq \ldots \leq y_N$. Then $\Delta_i = (y_{i+1} - y_i)/(x_{i+1} - x_i) \geq 0$, $i = 1, 2, \ldots, N-1$. From calculus, Φ is monotonically increasing in $[x_1, x_N]$ if $\Phi^{(1)}(x) \geq 0$, for all $x \in [x_1, x_N]$. We have

$$\Phi^{(1)}(L_i(x)) = \frac{\alpha_i \Phi^{(1)}(x)}{a_i} + \frac{\Psi_i(\theta)}{(Q_i(\theta))^2},$$

where $\Psi_i(\theta) = \sum_{k=1}^{5} B_{k,i} \theta^{k-1} (1-\theta)^{5-k}$,

$$B_{1,i} = u_i^2 d_i^*,$$

$$B_{2,i} = 2u_i\{(3u_i + v_i)\Delta_i^* - u_i d_{i+1}^*\},$$

$$B_{3,i} = 3u_i^2 \Delta_i^* + (3u_i + v_i)\{(3u_i + v_i)\Delta_i^* - u_i d_{i+1}^* - u_i d_i^*\},$$

$$B_{4,i} = 2u_i\{(3u_i + v_i)\Delta_i^* - u_i d_i^*\},$$

$$B_{5,i} = u_i^2 d_{i+1}^*,$$

$$d_i^* = d_i - \frac{\alpha_i d_1}{a_i}, \ d_{i+1}^* = d_{i+1} - \frac{\alpha_i d_N}{a_i} \text{ and } \Delta_i^* = \Delta_i - \alpha_i \frac{y_N - y_1}{h_i}.$$

Theorem 2.1.3. *Let $\{(x_i, y_i, d_i) : i = 1, 2, \dots, N\}$ be a monotonically increasing data. Let the derivative values satisfy the necessary condition for monotonicity, i.e., $sgn(d_i) = sgn(d_{i+1}) = sgn(\Delta_i)$. Then the following conditions on the scaling factors and the shape parameters are sufficient to the rational cubic FIF Φ defined in (2.4) to be monotonically increasing.*

$$0 \le \alpha_i < \left\{ a_i,\ \frac{a_i d_i}{d_1},\ \frac{a_i d_{i+1}}{d_N},\ \frac{y_{i+1} - y_i}{y_N - y_1} \right\},$$

$$u_i > 0 \textit{ and } v_i > \max\left\{0,\ \frac{u_i(d_i^* + d_{i+1}^*)}{\Delta_i^*}\right\},\ i = 1, 2, \dots, N-1.$$

Proof. For each node x_j, $j = 1, 2, \dots, N$, we have

$$\Phi^{(1)}(L_i(x_j)) = \frac{\alpha_i \Phi^{(1)}(x_j)}{a_i} + \frac{\Psi_i(\theta_j)}{(Q_i(\theta_j))^2}, \quad \theta_j = \frac{x_j - x_1}{x_N - x_1},\ i = 1, 2, \dots, N-1.$$

Assume $\alpha_i \ge 0,\ i = 1, 2, \dots, N-1$. It can be seen that $(Q_i(\theta_j))^2 > 0$. Therefore, $\Phi^{(1)}(L_i(x_j)) \ge 0$, if $\Psi_i(\theta_j) \ge 0$. We have

$$\Psi_i(\theta_j) = B_{1,i}(1-\theta_j)^4 + B_{2,i}\theta_j(1-\theta_j)^3 + B_{3,i}{\theta_j}^2(1-\theta_j)^2 + B_{4,i}{\theta_j}^3(1-\theta_j) + B_{5,i}\theta_j^4.$$

It can be seen that $\Psi_i(\theta_j) \ge 0$, if $B_{1,i} \ge 0,\ B_{2,i} \ge 0,\ B_{3,i} \ge 0,\ B_{4,i} \ge 0$ and $B_{5,i} \ge 0$. We have

$$B_{1,i} \ge 0 \text{ if } \alpha_i \le \frac{a_i d_i}{d_1},\ B_{5,i} \ge 0 \text{ if } \alpha_i \le \frac{a_i d_{i+1}}{d_N}.$$

Let $0 \le \alpha_i < \left\{\frac{a_i d_i}{d_1}, \frac{a_i d_{i+1}}{d_N}, \frac{y_{i+1} - y_i}{y_N - y_1}\right\}$. Then, we get

$$B_{2,i} \ge 0 \text{ if } v_i \ge \frac{u_i d_{i+1}^*}{\Delta_i^*},\ B_{3,i} \ge 0 \text{ if } v_i \ge \frac{u_i(d_i^* + d_{i+1}^*)}{\Delta_i^*},\ B_{5,i} \ge 0 \text{ if } v_i \ge \frac{u_i d_i^*}{\Delta_i^*}.$$

So according to the conditions prescribed in Theorem 2.1.3, it is clear that $\Phi^{(1)}(L_i(x_j)) \ge 0$ for all $j = 1, 2, \dots, N,\ i = 1, 2, \dots, N-1$. Since $\Phi^{(1)}$ is also a fractal function and it has recursive nature, the condition $\Phi^{(1)}(L_i(x_j)) \ge 0$, $i = 1, 2, \dots, N-1,\ j = 1, 2, \dots, N$ gives $\Phi^{(1)}(x) \ge 0$ for all $x \in [x_1, x_N]$. □

Remark 2.1.7. *If $\alpha_i = 0,\ i = 1, 2, \dots, N-1$ then the sufficient conditions for classical rational cubic spline (2.6) which preserves monotonicity are*

$$u_i > 0 \textit{ and } v_i > \max\left\{0,\ \frac{u_i(d_i + d_{i+1})}{\Delta_i}\right\},\ i = 1, 2, \dots, N-1.$$

Convexity

Assume that the data $\{(x_i, y_i) : i = 1, 2, \dots, N\}$ is strictly convex. To avoid the possibility of straight line segments, assume that $d_1 < \Delta_1 < \cdots < d_i < \Delta_i < d_{i+1} < \cdots < \Delta_{N-1} < d_N$. Since Φ belongs to $\mathcal{C}^1$, Φ is convex if $\Phi^{(2)}(x^+)$ or $\Phi^{(2)}(x^-)$ exists and non-negative for all $x \in (x_1, x_N)$ [96]. In this section, sufficient conditions on the scaling factors and the shape parameters will be determined to ensure the convexity of the rational cubic FIF (2.4).

Theorem 2.1.4. *Suppose $\{(x_i, y_i) : i = 1, 2, \dots, N\}$ is a strictly convex data. Let d_i be the derivative value at the knot x_i. Let the derivative values satisfy $d_1 < \Delta_1 < \cdots < d_i < \Delta_i < d_{i+1} < \cdots < \Delta_{N-1} < d_N$. Then sufficient conditions on the scaling factors and the shape parameters to ensure convexity of Φ in (2.4) are*

$$0 \le \alpha_i < \min\left\{a_i^2, \frac{h_i(d_{i+1}-\Delta_i)}{d_N(x_N - x_1) - (y_N - y_1)}, \frac{h_i(\Delta_i - d_i)}{(y_N - y_1) - d_1(x_N - x_1)}, \frac{a_i(d_{i+1} - d_i)}{(d_N - d_1)}\right\},$$

$$u_i > 0 \text{ and } v_i > \max\left\{0, \frac{u_i(d_{i+1}^* - \Delta_i^*)}{(\Delta_i^* - d_i^*)}, \frac{u_i(\Delta_i^* - d_i^*)}{(d_{i+1}^* - \Delta_i^*)}\right\}.$$

Proof. Informally, we have

$$\Phi^{(2)}(L_i(x)) = \frac{\alpha_i \Phi^{(2)}(x)}{a_i^2} + \frac{\Psi_i^*(\theta)}{h_i(Q_i(\theta))^3},$$

where $\Psi_i^*(\theta) = \sum_{k=1}^{6} C_{k,i}\theta^{k-1}(1-\theta)^{6-k}$,

$$C_{1,i} = 2u_i^2\{(2u_i + v_i)(\Delta_i^* - d_i^*) - u_i(d_{i+1}^* - \Delta_i^*)\},$$

$$C_{2,i} = 2u_i^2\{7u_i(\Delta_i^* - d_i^*) + 2v_i(\Delta_i^* - d_i^*) - 2u_i(d_{i+1}^* - \Delta_i^*)\},$$

$$C_{3,i} = 2u_i\{(6u_i^2 + u_iv_i)(\Delta_i^* - d_i^*) + 2u_i^2(d_{i+1}^* - d_i^*)\},$$

$$C_{4,i} = 2u_i\{(6u_i^2 + u_iv_i)(d_{i+1}^* - \Delta_i^*) + 2u_i^2(d_{i+1}^* - d_i^*)\},$$

$$C_{5,i} = 2u_i^2\{7u_i(d_{i+1}^* - \Delta_i^*) + 2v_i(d_{i+1}^* - \Delta_i^*) - 2u_i(\Delta_i^* - d_i^*)\},$$

$$C_{6,i} = 2u_i^2\{(2u_i + v_i)(d_{i+1}^* - \Delta_i^*) - u_i(\Delta_i^* - d_i^*)\}.$$

Now, we get

$$\Phi^{(2)}(x_1^+) = \frac{C_{1,1}}{h_1 u_1^3}\left[1 - \frac{\alpha_1}{a_1^2}\right]^{-1}, \tag{2.17}$$

$$\Phi^{(2)}(x_N^-) = \frac{C_{6,N-1}}{h_{N-1} u_{N-1}^3}\left[1 - \frac{\alpha_{N-1}}{a_{N-1}^2}\right]^{-1}, \tag{2.18}$$

$$\Phi^{(2)}(x_n^+) = \frac{\alpha_n}{a_n^2}\Phi^{(2)}(x_1^+) + \frac{C_{1,n}}{u_n^3 h_n}, \quad n = 2, 3, \ldots, N-1. \tag{2.19}$$

Let $0 \le \alpha_i < a_i^2,\ i = 1, 2, \ldots, N-1$. From (2.17), (2.18) and (2.19), it is evident that if $C_{1,i} \ge 0,\ i = 1, 2, \ldots, N-1$ and $C_{6,N-1} \ge 0$, then the right-handed second derivatives at the knots $x_i,\ i = 1, 2, \ldots, N-1$ and the left-handed second derivative at x_N are non-negative. For a knot point $x_n,\ n = 1, 2, \ldots, N-1$, we get

$$\Phi^{(2)}(L_i(x_n^+)) = \frac{\alpha_i \Phi^{(2)}(x_n^+)}{a_i^2} + R_i(x_n^+),\ i = 1, 2, \ldots, N-1,$$

where $R_i(x) = R_i(x_1 + \theta(x_N - x_1)) = \Psi_i^*(\theta)/(h_i(Q_i(\theta))^3)$. Assuming that $C_{1,i} \ge 0,\ i = 1, 2, \ldots, N-1$, we get $\Phi^{(2)}(L_i(x_n^+)) \ge 0$ if $R_i(x_n^+) \ge 0$. Note that $R_i(x_n^+) \ge 0$ if the coefficients $C_{j,i} \ge 0$, for $j = 1, 2, \ldots, 6$. From the Three Chords Lemma for the convex functions [145], it follows that for a strictly convex interpolant, the end point derivatives should necessarily satisfy $d_1 < (y_N - y_1)/(x_N - x_1) < d_N$. One can see that

$$(\Delta_i^* - d_i^*) > 0 \quad \text{if } \alpha_i < \frac{h_i(\Delta_i - d_i)}{(y_N - y_1) - d_1(x_N - x_1)},$$

$$(d_{i+1}^* - \Delta_i^*) > 0 \quad \text{if } \alpha_i < \frac{h_i(d_{i+1} - \Delta_i)}{d_N(x_N - x_1) - (y_N - y_1)},$$

$$(d_{i+1}^* - d_i^*) > 0 \quad \text{if } \alpha_i < \frac{a_i(d_{i+1} - d_i)}{(d_N - d_1)}.$$

Let

$$0 \le \alpha_i < \min\left\{\frac{h_i(d_{i+1} - \Delta_i)}{d_N(x_N - x_1) - (y_N - y_1)}, \frac{h_i(\Delta_i - d_i)}{(y_N - y_1) - d_1(x_N - x_1)}, \frac{a_i(d_{i+1} - d_i)}{(d_N - d_1)}\right\}. \tag{2.20}$$

It can be seen that

$$C_{1,i} \ge 0 \quad \Leftrightarrow \quad (2u_i + v_i)(\Delta_i^* - d_i^*) - u_i(d_{i+1}^* - \Delta_i^*) \ge 0.$$

$$(2u_i + v_i)(\Delta_i^* - d_i^*) - u_i(d_{i+1}^* - \Delta_i^*) \ge 0 \quad \text{if } v_i \ge \frac{u_i(d_{i+1}^* - \Delta_i^*)}{(\Delta_i^* - d_i^*)}.$$

$$C_{2,i} \ge 0 \quad \Leftrightarrow \quad 7u_i(\Delta_i^* - d_i^*) + 2v_i(\Delta_i^* - d_i^*) - 2u_i(d_{i+1}^* - \Delta_i^*) \ge 0.$$

$$7u_i(\Delta_i^* - d_i^*) + 2v_i(\Delta_i^* - d_i^*) - 2u_i(d_{i+1}^* - \Delta_i^*) \geq 0 \quad \text{if } v_i \geq \frac{u_i(d_{i+1}^* - \Delta_i^*)}{(\Delta_i^* - d_i^*)}.$$

$$C_{3,i} \geq 0 \quad \Leftrightarrow \quad (6u_i^2 + u_i v_i)(\Delta_i^* - d_i^*) + 2u_i^2(d_{i+1}^* - d_i^*) \geq 0.$$

The assumption on the scaling factors given in (2.20) ensures that $C_{3,i} \geq 0$.

$$C_{4,i} \geq 0 \quad \Leftrightarrow \quad (6u_i^2 + u_i v_i)(d_{i+1}^* - \Delta_i^*) + 2u_i^2(d_{i+1}^* - d_i^*) \geq 0.$$

The assumption on the scaling factors given in (2.20) shows that $C_{4,i} \geq 0$. Similarly, one can observe that

$$C_{5,i} \geq 0 \quad \text{if } v_i \geq \frac{u_i(\Delta_i^* - d_i^*)}{(d_{i+1}^* - \Delta_i^*)},$$
$$C_{6,i} \geq 0 \quad \text{if } v_i \geq \frac{u_i(\Delta_i^* - d_i^*)}{(d_{i+1}^* - \Delta_i^*)}.$$

Thus the conditions on the scaling factors and the shape parameters given in Theorem 2.1.4 ensure that $C_{j,i} \geq 0$, $i = 1, 2, \ldots, N-1$, $j = 1, 2, \ldots, 6$ and hence the non-negativity of $\Phi^{(2)}(L_i(x_n^+))$ for $i, n = 1, 2, \ldots, N-1$, $\Phi^{(2)}(x_N^-)$. The non-negativity of $\Phi^{(2)}(L_i(x_n^+))$ for $i, n = 1, 2, \ldots, N-1$, and $\Phi^{(2)}(x_N^-)$ ensure that the non-negativity of $\Phi^{(2)}(x^+)$ or $\Phi^{(2)}(x^-)$ for $x \in (x_1, x_N)$. □

Remark 2.1.8. *If $\alpha_i = 0$, then the conditions*

$$u_i > 0 \text{ and } v_i > \max\left\{0, \frac{u_i(d_{i+1} - \Delta_i)}{(\Delta_i - d_i)}, \frac{u_i(\Delta_i - d_i)}{(d_{i+1} - \Delta_i)}\right\},$$

are the sufficient conditions for the classical rational cubic spline C to be convex.

Remark 2.1.9. *If $\Delta_i - d_i = 0$ or $d_{i+1} - \Delta_i = 0$, then take $\alpha_i = 0$, $d_i = d_{i+1} = \Delta_i$. In this case, Φ becomes a straight line $\Phi(L_i(x)) = y_i(1-\theta) + y_{i+1}\theta$ in the interval $[x_i, x_{i+1}]$.*

Constrained Interpolation

Let $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ be a data such that it lies above the straight line $p = mx + c$, i.e., $y_i > p_i$ where $p_i = mx_i + c$. Then, in general, the rational cubic FIF Φ may not lie above the line $p = mx + c$ in $[x_1, x_N]$. In this section, sufficient conditions on the scaling factors and the shape parameters are derived such that Φ would lie above the straight line $p = mx + c$. The straight line $p = mx + c$ can be written as $p_1(1-\theta) + p_N\theta$, $\theta = (x - x_1)/(x_1 - x_N)$, $x \in [x_1, x_N]$. The rational cubic FIF Φ lies above the straight line p if $\Phi(x) > p_1(1-\theta) + p_N\theta$, for all $x \in [x_1, x_N]$. Since the graph of Φ is the attractor of the IFS (2.5), it is evident that $\Phi(x) > p_1(1-\theta) + p_N\theta$, for all $x \in [x_1, x_N]$ if $F_i(x, y) > p_i(1-\theta) + p_{i+1}\theta$ for all (x, y) such that $x \in [x_1, x_N]$, $y > p_1(1-\theta) + p_N\theta$, $i = 1, 2, \ldots, N-1$.

Theorem 2.1.5. *Let $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ be the data such that it lies above the straight line $p = mx + c$, i.e., $y_i > p_i$ where $p_i = mx_i + c$. Then the rational cubic FIF Φ defined in (2.4) lies above the straight line $p = mx + c$ in $[x_1, x_N]$, if the scaling factors and the shape parameters satisfy the following conditions:*

$$0 \le \alpha_i < \min\left\{a_i, \ \frac{y_i - p_i}{y_1 - p_1}, \ \frac{y_{i+1} - p_{i+1}}{y_N - p_N}\right\},$$

$$u_i > 0 \text{ and } v_i > \max\left\{0, \delta_{1,i}, \ \delta_{2,i}\right\},$$

where

$$\delta_{1,i} = \frac{-u_i[(y_i - p_{i+1}) - \alpha_i(y_1 - p_N)] - u_i[h_i d_i - \alpha_i(x_N - x_1)d_1]}{(y_i - p_i) - \alpha_i(y_1 - p_1)},$$
$$\delta_{2,i} = \frac{-u_i[(y_{i+1} - p_i) - \alpha_i(y_N - p_1)] - u_i[-h_i d_{i+1} + \alpha_i(x_N - x_1)d_N]}{(y_{i+1} - p_{i+1}) - \alpha_i(y_N - p_N)},$$

$i = 1, 2, \ldots, N-1$.

Proof. For each fixed $i = 1, 2, \ldots, N-1$, let $0 \le \alpha_i < a_i$ and (x, y) such that $x \in [x_1, x_N]$ and $y > p_1(1-\theta) + p_N\theta$. It is evident that $y\alpha_i > [p_1(1-\theta) + p_N\theta]\alpha_i$. We have

$$\alpha_i y + \frac{P_i(\theta)}{Q_i(\theta)} > [p_1(1-\theta) + p_N\theta]\alpha_i + \frac{P_i(\theta)}{Q_i(\theta)}.$$

To prove $F_i(x, y) > p_i(1-\theta) + p_{i+1}\theta$, it is enough to prove that

$$[p_1(1-\theta) + p_N\theta]\alpha_i + \frac{P_i(\theta)}{Q_i(\theta)} > p_i(1-\theta) + p_{i+1}\theta. \tag{2.21}$$

Now (2.21) can be written as

$$\frac{Q_i(\theta)[(\alpha_i p_1 - p_i)(1-\theta) + (\alpha_i p_N - p_{i+1})\theta] + P_i(\theta)}{Q_i(\theta)} > 0. \tag{2.22}$$

Since $Q_i(\theta) > 0$, to prove (2.22), it is sufficient to show that $Q_i(\theta)[(\alpha_i p_1 - p_i)(1-\theta) + (\alpha_i p_N - p_{i+1})\theta] + P_i(\theta) > 0$. The numerator $Q_i(\theta)[(\alpha_i p_1 - p_i)(1-\theta) + (\alpha_i p_N - p_{i+1})\theta] + P_i(\theta)$ can be written as $Q_i(\theta)[(\alpha_i p_1 - p_i)(1-\theta) + (\alpha_i p_N - p_{i+1})\theta] + P_i(\theta) = D_{1,i}(1-\theta)^3 + D_{2,i}\theta(1-\theta)^2 + D_{3,i}\theta^2(1-\theta) + D_{4,i}\theta^3$, where

$$
\begin{aligned}
D_{1,i} &= u_i[(y_i - p_i) - \alpha_i(y_1 - p_1)], \ D_{4,i} = u_i[(y_{i+1} - p_{i+1}) - \alpha_i(y_N - p_N)], \\
D_{2,i} &= (2u_i + v_i)[(y_i - p_i) - \alpha_i(y_1 - p_1)] + u_i[(y_i - p_{i+1}) - \alpha_i(y_1 - p_N)] \\
&\quad + u_i[h_i d_i - \alpha_i(x_N - x_1)d_1], \\
D_{3,i} &= (2u_i + v_i)[(y_{i+1} - p_{i+1}) - \alpha_i(y_N - p_N)] + u_i[(y_{i+1} - p_i) - \alpha_i(y_N - p_1)] \\
&\quad + u_i[-h_i d_{i+1} + \alpha_i(x_N - x_1)d_N].
\end{aligned}
$$

It can be seen that $Q_i(\theta)[(\alpha_i p_1 - p_i)(1-\theta) + (\alpha_i p_N - p_{i+1})\theta] + P_i(\theta) > 0$ if all the coefficients $D_{j,i} > 0$, $j = 1, 2, \dots, 4$. We have

$$
D_{1,i} > 0 \ \text{ if } \alpha_i < \frac{y_i - p_i}{y_1 - p_1}, \quad D_{4,i} > 0 \text{ if } \alpha_i < \frac{y_{i+1} - p_{i+1}}{y_N - p_N}.
$$

Let $0 \le \alpha_i < \min\left\{\frac{y_i - p_i}{y_1 - p_1}, \ \frac{y_{i+1} - p_{i+1}}{y_N - p_N}\right\}$. The coefficient $D_{2,i} > 0$ if

$$
v_i > \frac{-u_i[(y_i - p_{i+1}) - \alpha_i(y_1 - p_N)] - u_i[h_i d_i - \alpha_i(x_N - x_1)d_1]}{(y_i - p_i) - \alpha_i(y_1 - p_1)}.
$$

The coefficient $D_{3,i} > 0$ if

$$
v_i > \frac{-u_i[(y_{i+1} - p_i) - \alpha_i(y_N - p_1)] - u_i[-h_i d_{i+1} + \alpha_i(x_N - x_1)d_N]}{(y_{i+1} - p_{i+1}) - \alpha_i(y_N - p_N)}.
$$

Thus the sufficient conditions on the scaling factors and the shape parameters to satisfy (2.21) are

$$
0 \le \alpha_i < \min\left\{a_i, \ \frac{y_i - p_i}{y_1 - p_1}, \ \frac{y_{i+1} - p_{i+1}}{y_N - p_N}\right\},
$$

$$
u_i > 0 \text{ and } v_i > \max\left\{\delta_{1,i}, \ \delta_{2,i}\right\},
$$

$i = 1, 2, \dots, N-1$. □

2.1.3 Numerical examples

The effectiveness of the rational cubic FIF (2.4) towards the visualization of the shaped data is illustrated through numerical examples.

Let us consider the positive data $\{ (1,14), (2,8), (3,2), (8,0.8), (10,0.5), (11,0.25), (12,0.4), (14,0.37) \}$ as used in [124]. The derivatives are approximated using arithmetic mean method and are given by $d_1 = -6.0000$, $d_2 = -6.0000$, $d_3 = -5.0400$, $d_4 = -0.1757$, $d_5 = -0.2167$, $d_6 = -0.0500$, $d_7 = 0.0950$ and $d_8 = -0.1250$.

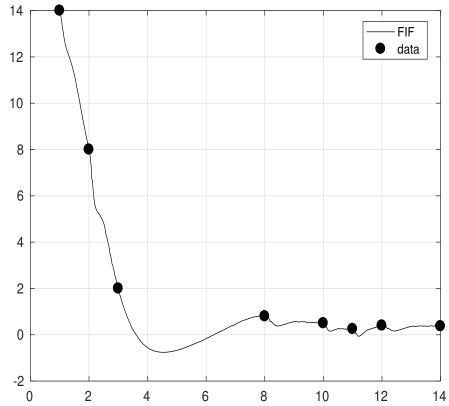
By taking the random scaling factors and shape parameters, FIF given in Figure 2.1(a) is drawn. It can be observed that, FIF (Figure 2.1(a)) is non-positive. Thus, the parameters are taken according to restrictions given in Theorem 2.1.2. Using these restricted parameters, FIF given in Figure 2.1(b) is drawn. We can see that, FIF (Figure 2.1(b)) is positive.

To see the effect of the scaling factors on the FIF given in Figure 2.1(b), by perturbing the scaling factors α_3 and α_4, the FIF given in Figure 2.1(c) is constructed. There are significant changes in Figure 2.1(c) in the intervals $[x_3, x_4]$ and $[x_4, x_5]$, whereas the changes in the other intervals are negligible.

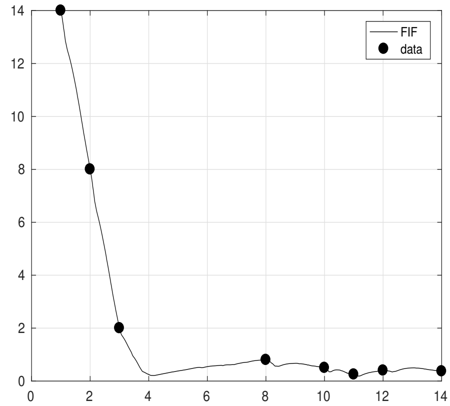
To demonstrate the effect of the shape parameters on the FIF given in Figure 2.1(b), the FIF in Figure 2.1(d) is constructed by changing the shape parameters v_3 and v_4. There is a little change in the intervals $[x_3, x_4]$ and $[x_4, x_5]$, but there is no significant changes in the other intervals.

By taking all the scaling factors as zero, the classical rational cubic spline is constructed and it is shown in Figure 2.1(e). Figure 2.2 represents the derivatives of FIFs given in Figure 2.1. The values of the scaling factors and the shape parameters used to construct the FIFs given in Figures 2.1 and 2.2 are tabulated in Table 2.1.

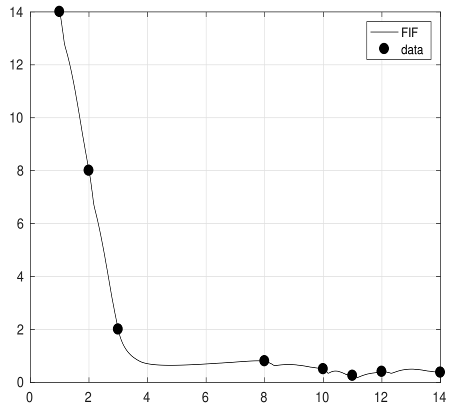
Next, we consider the monotonically increasing data $\{ (0,0.5), (2.5,1.61), (3,7.3891), (6,9.8696), (11,22.18), (15,27.3),(20,35.2) \}$ [31]. The derivatives are computed using arithmetic mean method and are given by $d_1 = 0$, $d_2 = 9.7058$, $d_3 = 10.0251$, $d_4 = 1.4401$, $d_5 = 1.8054$, $d_6 = 1.4133$ and $d_7 = 1.7467$. According to restrictions given in Theorem 2.1.3, the scaling factors and the shape parameters are restricted.



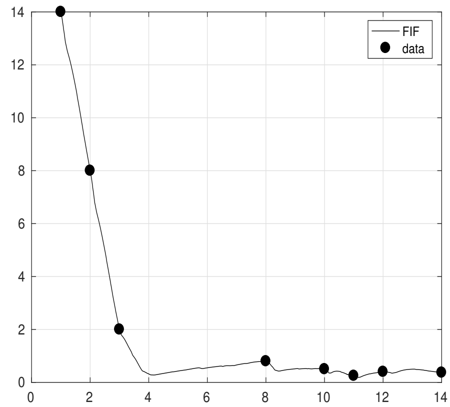
(a) Non-positive FIF.



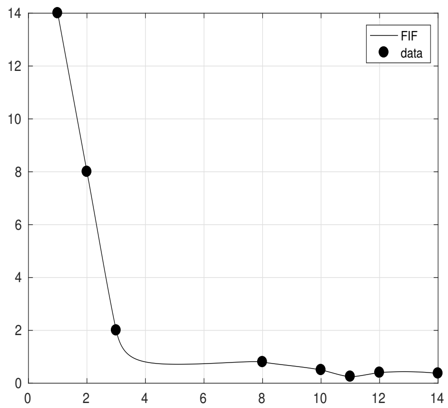
(b) Positive rational FIF.



(c) Effect of α_3 and α_4.

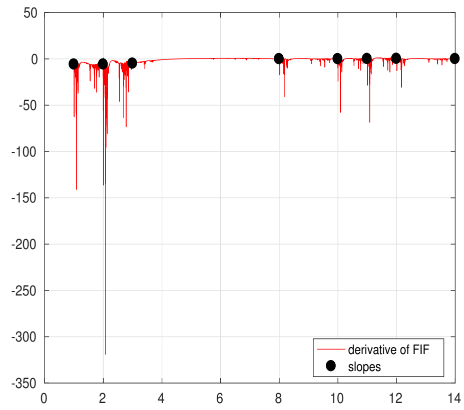


(d) Effect of v_3 and v_4.

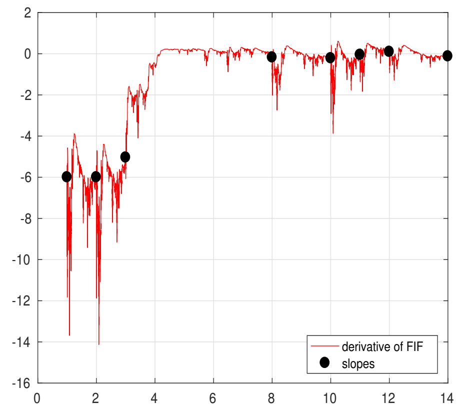


(e) Classical positive rational cubic spline.

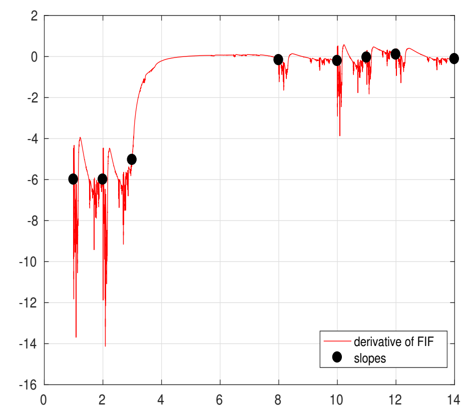
Figure 2.1: Positivity preserving interpolation.



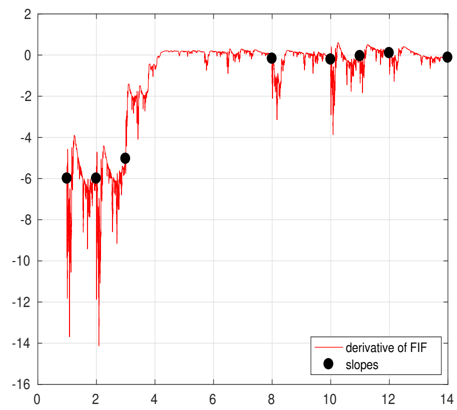
(a) Derivative of FIF given in Figure 2.1(a).



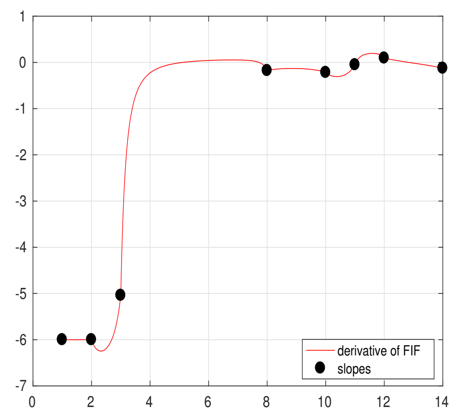
(b) Derivative of FIF given in Figure 2.1(b).



(c) Derivative of FIF given in Figure 2.1(c).



(d) Derivative of FIF given in Figure 2.1(d).



(e) Derivative of FIF given in Figure 2.1(e).

Figure 2.2: Derivatives of FIFs given in Figure 2.1.

Table 2.1: Parameters for positive FIFs with $u_i = 1.5$ for $i = 1, 2, \ldots, 7$.

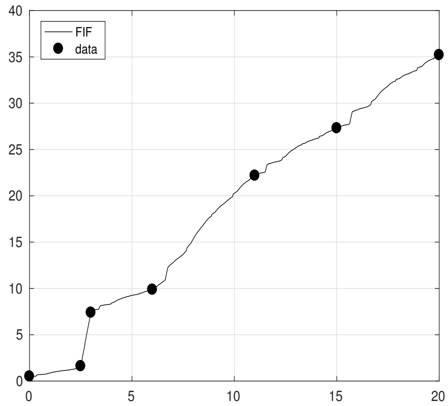
↓Parameter/Figure →	2.1(a), 2.2(a)	2.1(b), 2.2(b)	2.1(c), 2.2(c)	2.1(d), 2.2(d)	2.1(e), 2.2(e)
α_1	0.1000	0.0754	0.0754	0.0754	0
α_2	0.2300	0.0754	0.0754	0.0754	0
α_3	0.0300	0.1324	0.0132	0.1324	0
α_4	0.0600	0.0471	0.0271	0.0471	0
α_5	0.0420	0.0347	0.0347	0.0347	0
α_6	0.0500	0.0169	0.0169	0.0169	0
α_7	0.0450	0.0266	0.0266	0.0266	0
v_1	0.5000	0.5000	0.5000	0.5000	0.8000
v_2	3.5000	1.5000	1.5000	1.5000	1.5000
v_3	4.0000	153.0000	20.9733	350.0000	20.9000
v_4	5.6000	2.5000	2.5000	10.5000	1.5000
v_5	8.5000	1.0000	1.0000	1.0000	1.0000
v_6	7.9000	3.0000	3.0000	3.0000	3.0000
v_7	5.0000	0.5000	0.5000	0.5000	0.5000

Using the restricted parameters, FIF given in Figure 2.3(a) is constructed. It can be seen that, FIF (Figure 2.3(a)) is monotonically increasing. By perturbing the scaling factors α_1 and α_3, the FIF in Figure 2.3(b) is constructed. Visible changes occurred in the intervals $[x_1, x_2]$ and $[x_3, x_4]$, but changes in the other intervals are not visible.

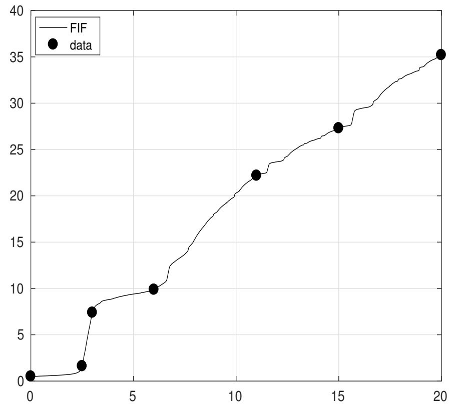
Similarly, by perturbing the scaling factors α_4, α_5 and α_6 with respect to Figure 2.3(a), Figure 2.3(c) is constructed. The FIF given in Figure 2.3(d) is plotted by changing the shape parameters v_3 and v_4. Finally by taking all the scaling factors are zero, classical rational cubic spline is constructed and is shown in Figure 2.3(e). Figure 2.4 represents derivatives of FIFs given in Figure 2.3. The values of the scaling factors and the shape parameters used to draw FIFs in Figures 2.3 and 2.4 are given in Table 2.2.

Next, we consider the convex data $\{\ (1, 10), (2, 2.5), (4, 0.625), (5, 0.4), (10, 0.1)\ \}$ as given in [124]. Using arithmetic mean method, the derivatives are approximated and are given by $d_1 = -9.6875$, $d_2 = -5.3125$, $d_3 = -0.4625$, $d_4 = -0.1975$ and $d_5 = 0$.

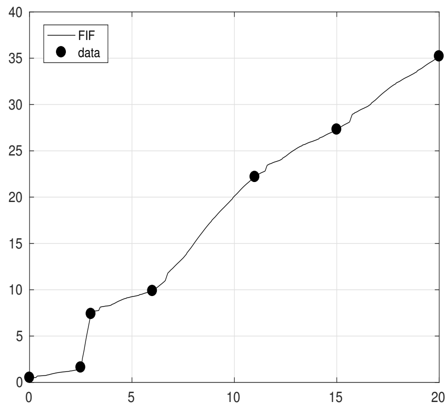
By taking random parameters, FIF (Figure 2.5(a)) is constructed and it can be noticed that FIF is non-convex. Thus, according to the restrictions given in Theorem 2.1.4, the scaling factors and the shape parameters are restricted.



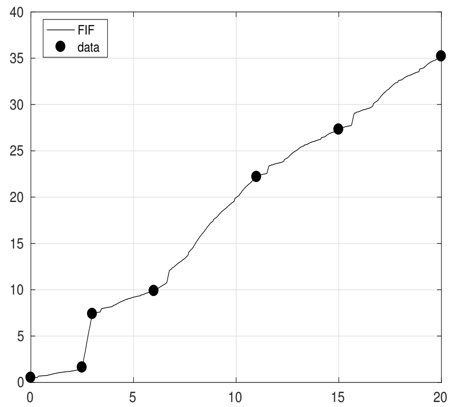
(a) Monotone rational FIF.



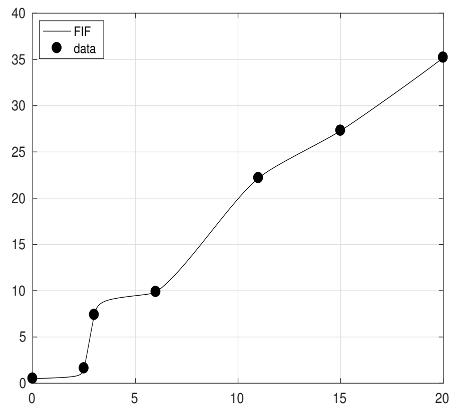
(b) Effect of α_1 and α_3.



(c) Effect of α_4, α_5 and α_6.

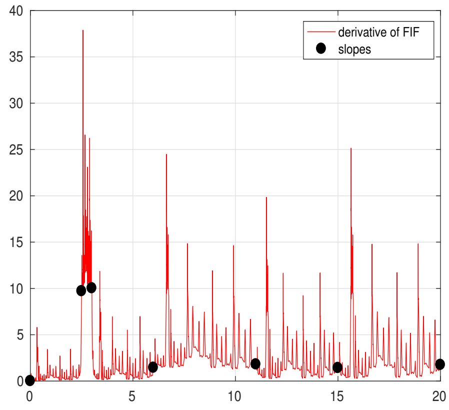


(d) Effect of v_3 and v_4.

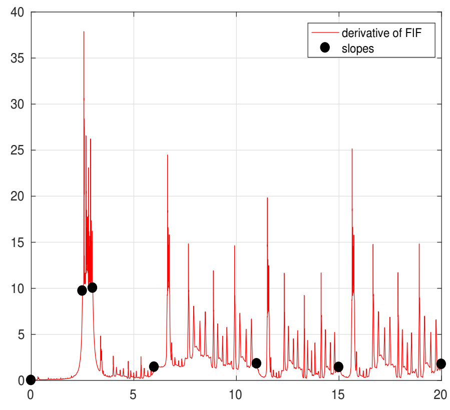


(e) Classical monotone rational cubic spline.

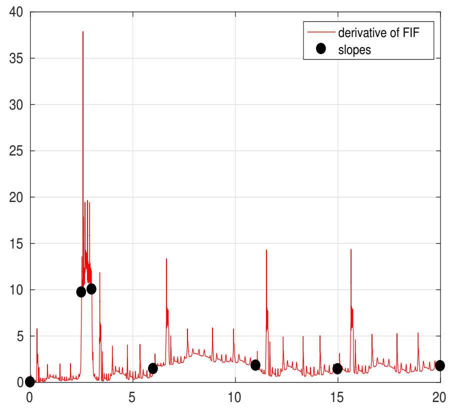
Figure 2.3: Monotonicity preserving interpolation.



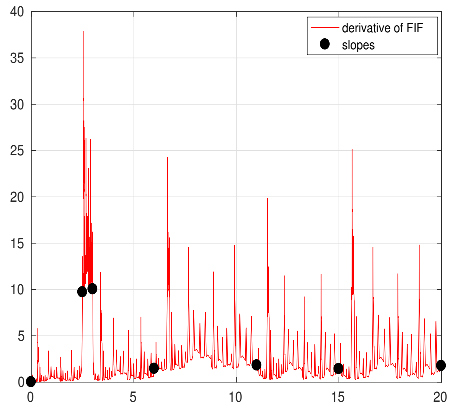
(a) Derivative of FIF given in Figure 2.3(a).



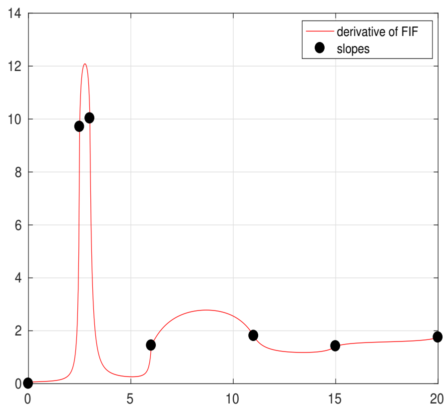
(b) Derivative of FIF given in Figure 2.3(b).



(c) Derivative of FIF given in Figure 2.3(c).



(d) Derivative of FIF given in Figure 2.3(d).



(e) Derivative of FIF given in Figure 2.3(e).

Figure 2.4: Derivatives of FIFs given in Figure 2.3.

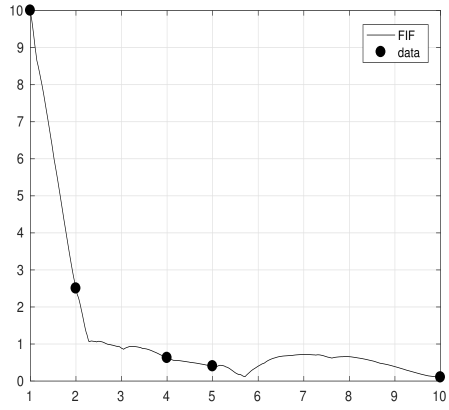
Table 2.2: Parameters for monotonic FIFs with $u_i = 1.5$ for $i = 1, 2, \ldots, 6$.

↓Parameter/Figure →	2.3(a), 2.4(a)	2.3(b), 2.4(b)	2.3(c), 2.4(c)	2.3(d), 2.4(d)	2.3(e), 2.4(e)
α_1	0.0250	0.0015	0.0250	0.0250	0
α_2	0.0240	0.0240	0.0240	0.0240	0
α_3	0.0610	0.0210	0.0610	0.0610	0
α_4	0.2000	0.2000	0.1000	0.2000	0
α_5	0.1360	0.1360	0.0940	0.1360	0
α_6	0.2160	0.2160	0.1160	0.2160	0
v_1	145.7800	35.1300	145.7800	145.7800	33.7900
v_2	3.7300	3.7300	3.7300	3.7300	3.6000
v_3	134.0000	32.6400	134.0000	365.0000	21.7900
v_4	3.2000	3.2000	2.6000	120.0000	2.0000
v_5	32.4000	32.4000	8.7000	32.4000	4.5000
v_6	31.6000	31.6000	5.3000	31.6000	4.0000

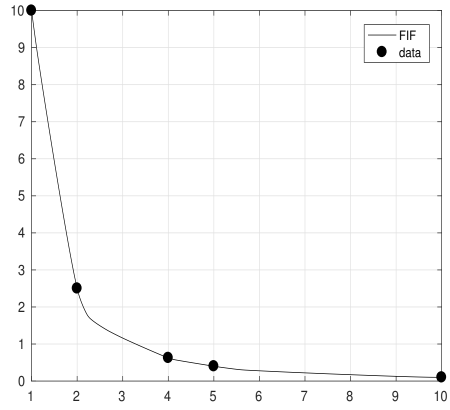
With the help of restricted parameters, convex FIF (Figure 2.5(b)) is drawn. By perturbing the scaling factor α_2, the FIF given in Figure 2.5(c) is constructed. From Figure 2.5(c), it is evident that significant change occurred in the interval $[x_2, x_3]$ and changes in other intervals are negligible.

The FIF given in Figure 2.5(d) is constructed after perturbing shape parameter v_2. In the interval $[x_2, x_3]$, curve moves upward and whereas in other intervals the changes are not visible. By taking all the scaling factors are zero, Figure 2.5(e) is drawn which is classical rational cubic spline preserving convexity. Figure 2.6 represents derivatives of FIFs given in Figure 2.5. The values of the scaling factors and the shape parameters used in drawing various FIFs in Figure 2.5 and 2.6 are given in Table 2.3.

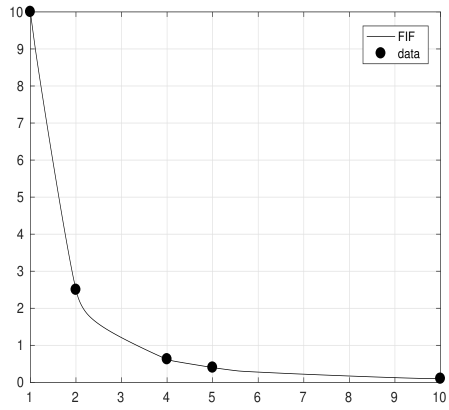
Next, let us consider the constrained interpolation data $\{ (1, 2.5), (1.25, 1.5), (2.8, 2), (3, 2.5), (3.2, 3.5), (4.2, 4.5), (4.5, 5, 5) \}$ which lies above the line $p = 0.5x + 0.28$. By taking arbitrary scaling factors and shape parameters, FIF (Figure 2.7(a)) is constructed. From Figure 2.7(a), it is clear that FIF does not lying above the line $p = 0.5x + 0.28$. Therefore, the scaling factors and the shape parameters are taken according to Theorem 2.1.5 and using these, Figure 2.7(b) is drawn.



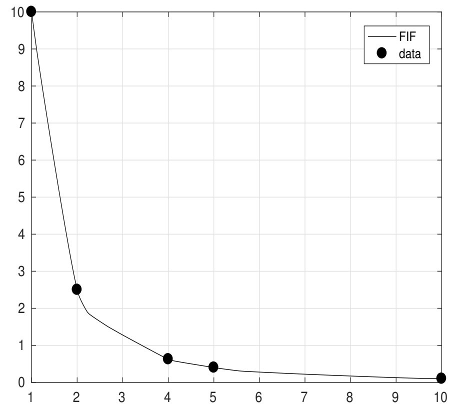
(a) Non-convex FIF.



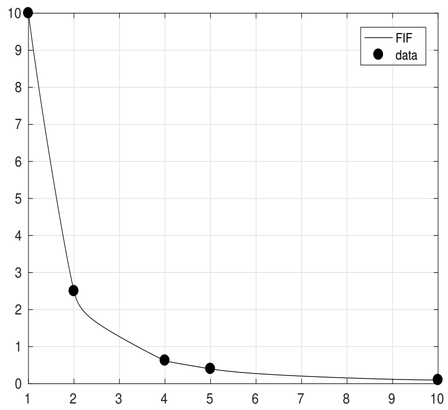
(b) Convex rational cubic FIF.



(c) Effect of α_2.

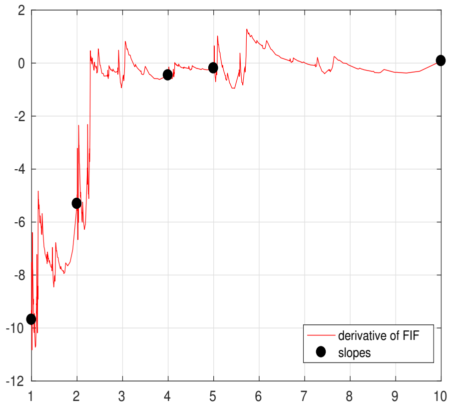


(d) Effect of v_2.

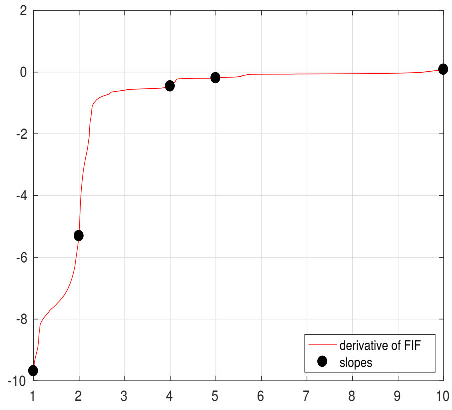


(e) Classical convex rational cubic spline.

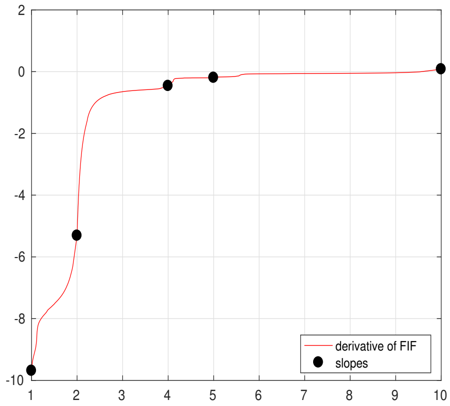
Figure 2.5: Convexity preserving interpolation.



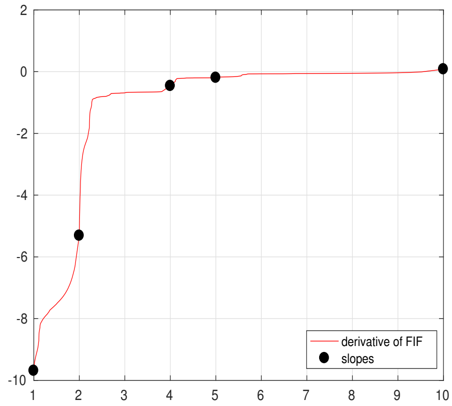
(a) Derivative of FIF given in Figure 2.5(a).



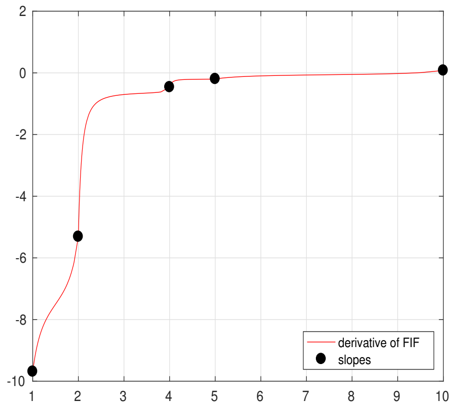
(b) Derivative of FIF given in Figure 2.5(b).



(c) Derivative of FIF given in Figure 2.5(c).



(d) Derivative of FIF given in Figure 2.5(d).



(e) Derivative of FIF given in Figure 2.5(e).

Figure 2.6: Derivatives of FIFs given in Figure 2.5.

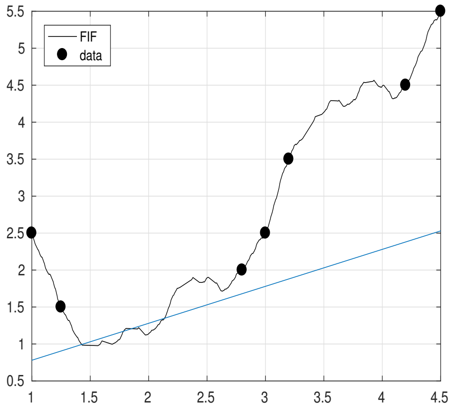
Table 2.3: Parameters for convex FIFs with $u_i = 1.5$ for $i = 1, 2, 3, 4$.

↓Parameter/Figure →	2.5(a), 2.6(a)	2.5(b), 2.6(b)	2.5(c), 2.6(c)	2.5(d), 2.6(d)	2.5(e), 2.6(e)
α_1	0.1000	0.0113	0.0113	0.0113	0
α_2	0.2000	0.0434	0.0043	0.0434	0
α_3	0.0100	0.0020	0.0020	0.0020	0
α_4	0.2000	0.0081	0.0081	0.0081	0
v_1	2.4000	3.3692	3.3692	3.3692	2.5000
v_2	14.0000	16.5544	14.9146	50.0000	18.8158
v_3	4.6000	17.1542	17.1542	17.1542	13.9545
v_4	3.0000	7.3637	7.3637	7.3637	4.4375

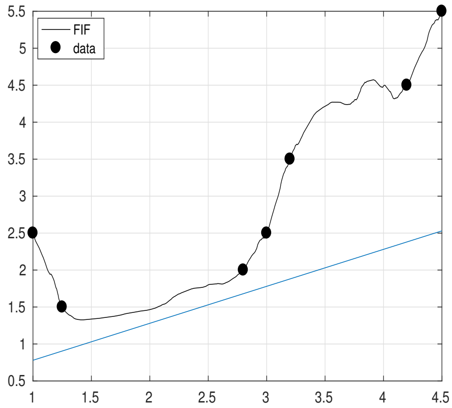
The FIF given in Figure 2.7(b) lies above the line $p = 0.5x + 0.28$. After perturbing scaling factor α_5 with respect to Figure 2.7(b), Figure 2.7(c) is constructed. In a similar fashion, after perturbing the shape parameter v_5 with respect to Figure 2.7(b), Figure 2.7(d) is drawn. Classical constrained FIF is constructed by taking all the scaling factors are zero which is shown in Figure 2.7(e). The parameters used to construct Figure 2.7 are given in Table 2.4.

Table 2.4: Parameters for constrained FIFs with $u_i = 1.5$ for $i = 1, 2, \ldots, 6$.

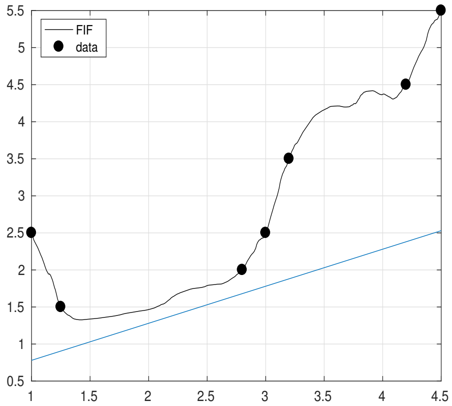
↓Parameter/Figure →	2.7(a)	2.7(b)	2.7(c)	2.7(d)	2.7(e)
α_1	0.07	0.07	0.07	0.07	0
α_2	0.4	0.09	0.09	0.09	0
α_3	0.05	0.05	0.05	0.05	0
α_4	0.05	0.05	0.05	0.05	0
α_5	0.28	0.28	0.18	0.28	0
α_6	0.08	0.08	0.08	0.08	0
v_1	3	3	3	3	3
v_2	3	46	46	46	14
v_3	3	3	3	3	3
v_4	3	3	3	3	3
v_5	3	3	3	10	3
v_6	3	3	3	3	3



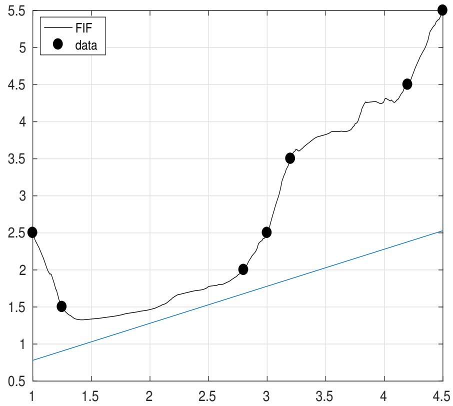
(a) Unconstrained rational FIF.



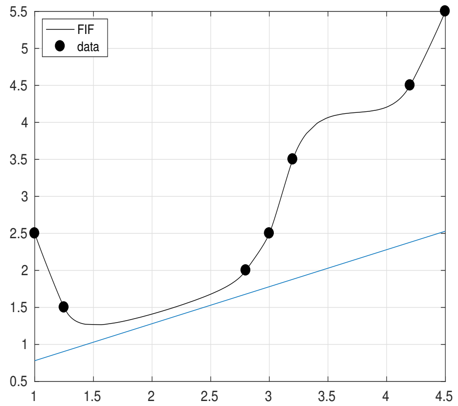
(b) Constrained rational FIF.



(c) Effect of α_5.



(d) Effect of v_5.



(e) Classical constrained rational cubic spline.

Figure 2.7: Constrained interpolation.

2.2 FIF with Three Families of Shape Parameters

In the present section, a rational cubic FIF with three families of the shape parameters is constructed based on Read-Bajraktarević operator. Let $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ be a set of data points such that $x_1 < x_2 < \cdots < x_N$. Let d_i be the derivative value at the knot point x_i. Consider the IFS (1.4) with

$$r_i(x) = \frac{p_i(x)}{q_i(x)} \equiv \frac{P_i(\theta)}{Q_i(\theta)} = \frac{A_{1,i}(1-\theta)^3 + A_{2,i}\theta(1-\theta)^2 + A_{3,i}\theta^2(1-\theta) + A_{4,i}\theta^3}{u_i(1-\theta)^2 + \theta(1-\theta)(u_i + v_i + \gamma_i) + v_i\theta^2}, \quad (2.23)$$

where $\theta = (x - x_1)/(x_N - x_1)$, $x \in [x_1, x_N]$. Here $A_{1,i}$, $A_{2,i}$, $A_{3,i}$, $A_{4,i}$ are constants such that it satisfies the required condition for the existence of $\mathcal{C}^1$-FIF. Here, u_i, v_i and γ_i are shape parameters such that $u_i > 0$, $v_i > 0$ and $\gamma_i \geq 0$. These parameters will ensure the positivity of denominator of $r_i(x)$.

Let $\mathscr{F} := \{\ \phi \in \mathcal{C}^1(I) |\ \phi(x_1) = y_1 \text{ and } \phi(x_N) = y_N\}$. Then $(\mathscr{F}, \rho)$ is a complete metric space, where ρ is the metric induced by norm $\|\phi\| = \|\phi\|_\infty + \|\phi^{(1)}\|_\infty$ on $\mathcal{C}^1(I)$. Define Read-Bajraktarević operator T on $(\mathscr{F}, \rho)$ as

$$T\phi(L_i(x)) = \alpha_i\phi(x) + r_i(x),\ x \in I,\ i = 1, 2, \ldots, N-1. \quad (2.24)$$

Let α_i such that $|\alpha_i| < a_i$ for all $i = 1, 2, \ldots, N-1$. The fixed-point Φ of T satisfies the functional equation

$$\Phi(L_i(x)) = \alpha_i\Phi(x) + r_i(x),\ x \in I,\ i = 1, 2, \ldots, N-1. \quad (2.25)$$

Here, Φ is a rational cubic FIF and derivative $\Phi^{(1)}$ is also a FIF which satisfies the functional equation

$$\Phi^{(1)}(L_i(x)) = \frac{\alpha_i\Phi^{(1)}(x) + r_i^{(1)}(x)}{a_i},\ x \in I,\ i = 1, 2, \ldots, N-1. \quad (2.26)$$

The constants $A_{1,i}$, $A_{2,i}$, $A_{3,i}$ and $A_{4,i}$ given in (2.23) are evaluated based on Hermite conditions $\Phi(x_i) = y_i$, $\Phi(x_{i+1}) = y_{i+1}$, $\Phi^{(1)}(x_i) = d_i$ and $\Phi^{(1)}(x_{i+1}) = d_{i+1}$ for $i = 1, 2, \ldots, N-1$. These conditions are equivalent [36] to the conditions on $F_i(x, y)$ for generating a $\mathcal{C}^1$-FIF given in Theorem 1.4.2. So, $\Phi^{(1)}$ is the fixed-point of the operator $T_* : \mathscr{F}_* \to \mathscr{F}_*$ defined by

$$(T_*\phi^*)(L_i(x)) = \frac{\alpha_i\phi^*(x) + r_i^{(1)}(x)}{a_i},\ x \in I,\ i = 1, 2, \ldots, N-1,$$

where $\mathscr{F}_* := \{\phi^* \in \mathcal{C}(I) | \phi^*(x_1) = d_1 \text{ and } \phi^*(x_N) = d_N\}$ is endowed with the uniform norm metric.

Substituting $x = x_1$ in (2.25) implies that

$$A_{1,i} = u_i[y_i - \alpha_i y_1].$$

Substituting $x = x_N$ in (2.25) gives

$$A_{4,i} = v_i[y_{i+1} - \alpha_i y_N].$$

The condition $\Phi^{(1)}(x_i) = d_i$ in (2.26) gives

$$A_{2,i} = u_i h_i d_i + y_i(2u_i + v_i + \gamma_i) - \alpha_i[u_i(x_N - x_1)d_1 + y_1(2u_i + v_i + \gamma_i)].$$

The condition $\Phi^{(1)}(x_{i+1}) = d_{i+1}$ in (2.26) gives

$$A_{3,i} = -v_i h_i d_{i+1} + y_{i+1}(u_i + 2v_i + \gamma_i) + \alpha_i[v_i(x_N - x_1)d_N - y_N(u_i + 2v_i + \gamma_i)].$$

Hence, the rational cubic FIF with three families of shape parameters is

$$\Phi(L_i(x)) = \alpha_i \Phi(x) + \frac{P_i(\theta)}{Q_i(\theta)}, \tag{2.27}$$

where

$$\begin{aligned} P_i(\theta) = {} & (u_i[y_i - \alpha_i y_1])(1-\theta)^3 + (v_i[y_{i+1} - \alpha_i y_N])\theta^3 + (u_i h_i d_i + y_i(2u_i + v_i + \gamma_i) \\ & - \alpha_i[u_i(x_N - x_1)d_1 + y_1(2u_i + v_i + \gamma_i)])\theta(1-\theta)^2 + (-v_i h_i d_{i+1} \\ & + y_{i+1}(u_i + 2v_i + \gamma_i) + \alpha_i[v_i(x_N - x_1)d_N - y_N(u_i + 2v_i + \gamma_i)])\theta^2(1-\theta), \\ Q_i(\theta) = {} & u_i(1-\theta)^2 + \theta(1-\theta)(u_i + v_i + \gamma_i) + v_i\theta^2, \theta = (x - x_1)/(x_N - x_1),\ x \in [x_1, x_N], \end{aligned}$$

$i = 1, 2, \ldots, N-1$.

Remark 2.2.1. *Rewrite the rational cubic FIF Φ given in (2.27) in the following form*

$$\Phi(L_i(x)) = \alpha_i \Phi(x) + \left\{\left[(1-\theta)y_i + \theta y_{i+1} + \frac{R_i(\theta)}{Q_i(\theta)}\right] - \alpha_i\left[(1-\theta)y_1 + \theta y_N + \frac{T_i(\theta)}{Q_i(\theta)}\right]\right\},$$

where

$$R_i(\theta) = h_i\theta(1-\theta)[(\Delta_i - d_{i+1})v_i\theta + (d_i - \Delta_i)u_i(1-\theta)],$$

$$T_i(\theta) = \theta(1-\theta)[\{(y_N - y_1) - (x_N - x_1)d_N\}v_i\theta +$$

$$\{(x_N - x_1)d_1 - (y_N - y_1)\}u_i(1-\theta)].$$

From this expression, it follows that if $\gamma_i \to \infty$ then Φ converges to the following affine FIF

$$\Phi(L_i(x)) = \alpha_i\Phi(x) + (y_i - \alpha_i y_1)(1-\theta) + (y_{i+1} - \alpha_i y_N)\theta.$$

Also, if $\gamma_i \to \infty$ and $\alpha_i \to 0$ then Φ converges to the straight line segment in the interval $[x_i, x_{i+1}]$, $i = 1, 2, \ldots, N-1$.

Remark 2.2.2. *If $\alpha_i = 0$ for all $i = 1, 2, \ldots, N-1$, then the rational cubic FIF given in (2.27) reduces to the classical rational cubic spline C as*

$$C(x) = \frac{U_i(\varphi)}{V_i(\varphi)}, \tag{2.28}$$

where

$$\begin{aligned} U_i(\varphi) &= u_i y_i (1-\varphi)^3 + [u_i h_i d_i + y_i(2u_i + v_i + \gamma_i)]\varphi(1-\varphi)^2 \\ &\quad + [-v_i h_i d_{i+1} + y_{i+1}(u_i + 2v_i + \gamma_i)]\varphi^2(1-\varphi) + v_i y_{i+1}\varphi^3, \\ V_i(\varphi) &= u_i(1-\varphi)^2 + \varphi(1-\varphi)(u_i + v_i + \gamma_i) + v_i\varphi^2, \ \varphi = \frac{x - x_i}{x_{i+1} - x_i}, \ x \in [x_i, x_{i+1}], \end{aligned}$$

$i = 1, 2, \ldots, N-1.$

Remark 2.2.3. *If $\alpha_i = \gamma_i = 0$, $u_i = v_i = 1$ on each subinterval $I_i = [x_i, x_{i+1}]$, $i = 1, 2, \ldots, N-1$, then the rational cubic FIF (2.27) reduces to the standard cubic Hermite spline*

$\Phi(x) = (2\varphi^3 - 3\varphi^2 + 1)y_i + (\varphi^3 - 2\varphi^2 + \varphi)h_i d_i + (-2\varphi^3 + 3\varphi^2)y_{i+1} + (\varphi^3 - \varphi^2)h_i d_{i+1}$,
where $\varphi = (x - x_i)/(x_{i+1} - x_i)$, $x \in [x_i, x_{i+1}]$.

2.2.1 Convergence analysis

In this section, an upper bound of the uniform error between an original function $S \in \mathcal{C}^2[x_1, x_N]$ and the rational cubic FIF Φ is determined. The effectiveness of the rational cubic FIF Φ in the approximation of a function S is derived with the help of classical rational cubic spline.

Theorem 2.2.1. *Let Φ as given in (2.27) and C as given in (2.28) respectively, be the rational cubic FIF and classical rational cubic spline with respect to the data* $\{(x_i, y_i),\ i = 1, 2, \ldots, N\}$ *generated from an original function* $S \in \mathcal{C}^2[x_1, x_N]$. *Let* $d_i,\ i = 1, 2, \ldots, N$ *denotes the derivative values at the knots. Then*

$$\|S - \Phi\|_\infty \leq \frac{|\boldsymbol{\alpha}|_\infty([M + h\bar{M}] + [M^* + |I|M_*])}{1 - |\boldsymbol{\alpha}|_\infty} + \|S^{(2)}\|_\infty h^2 c^*,$$

where $M = \max\{|y_i|,\ i = 1, 2, \ldots, N\}$, $\bar{M} = \max\{|d_i|,\ i = 1, 2, \ldots, N\}$, $M^* = \max\{|y_1|, |y_N|\}$, $M_* = \max\{|d_1|, |d_N|\}$, $|I| = x_N - x_1$, $h = \max\{h_i,\ i = 1, 2, \ldots, N-1\}$, $w(u_i, v_i, \gamma_i, \varphi) = \dfrac{\varphi^2(1-\varphi)^2 u_i^2 h_i^2 + \varphi^4(1-\varphi)^2 h_i^2(u_i + v_i + \gamma_i)^2}{2V_i(\varphi)[u_i + \varphi(u_i + v_i + \gamma_i)]}$ $+ \dfrac{\varphi^2(1-\varphi)^4 h_i^2(u_i + v_i + \gamma_i)^2 + \varphi^2(1-\varphi)^2 v_i^2 h_i^2}{2V_i(\varphi)[v_i + (u_i + v_i + \gamma_i)(1-\varphi)]}$, $c_i^* := \max\{w(u_i,\ v_i\ \gamma_i, \varphi) : 0 \leq \varphi \leq 1\}$, $i = 1, 2, \ldots, N-1$ *and* $c^* = \max\{c_i^* : i = 1, 2, \ldots, N-1\}$, $|\boldsymbol{\alpha}|_\infty = \max\{|\alpha_i|,\ i = 1, 2, \ldots, N-1\}$.

Proof. From (2.24), the Read-Bajraktarević operator $T_{\boldsymbol{\alpha}} : \mathscr{F} \to \mathscr{F}$ such that

$$T_{\boldsymbol{\alpha}}\phi(x) = \alpha_i \phi(L_i^{-1}(x)) + \frac{p_i(L_i^{-1}(x), \alpha_i)}{q_i(L_i^{-1}(x))},\ x \in I_i,\ i = 1, 2, \ldots, N-1,$$

where $p_i(x) \equiv P_i(\theta)$ and $q_i(x) \equiv Q_i(\theta)$ are as in (2.27). It is evident that Φ is the fixed-point of operator $T_{\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha} \neq \mathbf{0}$. Also, classical rational cubic spline C is the fixed-point of $T_{\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha} = \mathbf{0} = (0, 0, \ldots, 0) \in \mathbb{R}^{N-1}$.

Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{N-1})$ be a scale vector such that $|\alpha_i| < a_i$ for all $i = 1, 2, \ldots, N-1$ and with at least one $\alpha_i \neq 0$. For $\boldsymbol{\alpha} \neq \mathbf{0}$, $T_{\boldsymbol{\alpha}}$ is a contraction map with contraction factor $|\boldsymbol{\alpha}|_\infty$. Hence

$$\|T_{\boldsymbol{\alpha}}\Phi - T_{\boldsymbol{\alpha}}C\|_\infty \leq |\boldsymbol{\alpha}|_\infty \|\Phi - C\|_\infty. \tag{2.29}$$

Note that

$$\begin{aligned} |T_{\boldsymbol{\alpha}}C(x) - T_{\mathbf{0}}C(x)| &= \left|\alpha_i\ C \circ L_i^{-1}(x) + \frac{p_i(L_i^{-1}(x),\ \alpha_i)}{q_i(L_i^{-1}(x))} - \frac{p_i(L_i^{-1}(x),\ 0)}{q_i(L_i^{-1}(x))}\right| \\ &\leq |\alpha_i| \|C\|_\infty + \frac{|p_i(L_i^{-1}(x),\ \alpha_i) - p_i(L_i^{-1}(x),\ 0)|}{q_i(L_i^{-1}(x))}. \end{aligned} \tag{2.30}$$

Using the mean-value theorem of functions of several variables, there exist $\boldsymbol{\beta} = (\beta_1,\ \beta_2,\ \ldots,\ \beta_{N-1})$ such that $|\beta_i| < |\alpha_i|$ and

$$p_i(L_i^{-1}(x),\ \alpha_i) - p_i(L_i^{-1}(x),\ 0) = \frac{\partial p_i(L_i^{-1}(x),\ \beta_i)}{\partial \alpha_i}\alpha_i. \tag{2.31}$$

By (2.30) and (2.31),

$$|T_{\boldsymbol{\alpha}}C(x) - T_{\mathbf{0}}C(x)| \leq |\alpha_i| \left(\|C\|_\infty + \left| \frac{\partial \left(\frac{p_i(L_i^{-1}(x),\, \beta_i)}{q_i(L_i^{-1}(x))} \right)}{\partial \alpha_i} \right| \right). \tag{2.32}$$

Now, let us concentrate to find the bound for the right hand side of the Equation (2.32). The classical rational cubic spline C can be written as

$$C(x) = w_1(u_i, v_i, \gamma_i, \varphi)y_i + w_2(u_i, v_i, \gamma_i, \varphi)y_{i+1} + w_3(u_i, v_i, \gamma_i, \varphi)d_i - w_4(u_i, v_i, \gamma_i, \varphi)d_{i+1}, \tag{2.33}$$

where

$$\begin{aligned}
w_1(u_i,\ v_i,\ \gamma_i,\ \varphi) &= \frac{u_i(1-\varphi)^3 + (2u_i + v_i + \gamma_i)\varphi(1-\varphi)^2}{V_i(\varphi)},\\
w_2(u_i,\ v_i,\ \gamma_i,\ \varphi) &= \frac{(u_i + 2v_i + \gamma_i)\varphi^2(1-\varphi) + v_i\varphi^3}{V_i(\varphi)},\\
w_3(u_i,\ v_i,\ \gamma_i,\ \varphi) &= \frac{u_i h_i \varphi(1-\varphi)^2}{V_i(\varphi)},\\
w_4(u_i,\ v_i,\ \gamma_i,\ \varphi) &= \frac{v_i h_i \varphi^2(1-\varphi)}{V_i(\varphi)}.
\end{aligned}$$

It can be seen that

$$w_1(u_i,\ v_i,\ \gamma_i,\ \varphi) + w_2(u_i,\ v_i,\ \gamma_i,\ \varphi) = 1.$$

Note that

$$\begin{aligned}
w_3(u_i,\ v_i,\ \gamma_i,\ \varphi) + w_4(u_i,\ v_i,\ \gamma_i,\ \varphi) &= \frac{u_i h_i \varphi(1-\varphi)^2 + v_i h_i \varphi^2(1-\varphi)}{u_i(1-\varphi)^2 + \varphi(1-\varphi)(\gamma_i + u_i + v_i) + v_i\varphi^2}\\
&\leq \frac{u_i h_i \varphi(1-\varphi)^2 + v_i h_i \varphi^2(1-\varphi)}{u_i(1-\varphi)^2 + v_i\varphi^2}.
\end{aligned}$$

Since $u_i(1-\varphi)^2 + v_i\varphi^2 \geq \max\{u_i(1-\varphi)^2,\ v_i\varphi^2\}$, we get

$$\begin{aligned}
w_3(u_i,\ v_i,\ \gamma_i,\ \varphi) + w_4(u_i,\ v_i,\ \gamma_i,\ \varphi) &\leq \frac{u_i h_i \varphi(1-\varphi)^2}{u_i(1-\varphi)^2} + \frac{v_i h_i \varphi^2(1-\varphi)}{v_i\varphi^2}\\
&= h_i.
\end{aligned}$$

From (2.33), we obtain

$$|C(x)| \leq \max\{|y_i|, |y_{i+1}|\} + h_i \max\{|d_i|, |d_{i+1}|\}$$

$$= M_i + h_i \bar{M}_i,$$

where $M_i = \max\{|y_i|, |y_{i+1}|\}$ and $\bar{M}_i = \max\{|d_i|, |d_{i+1}|\}$.

Thus, we get

$$\|C\|_\infty \leq M + h\bar{M}. \tag{2.34}$$

Further, we have

$$\frac{\partial\left(\frac{p_i(L_i^{-1}(x),\, \alpha_i)}{q_i(L_i^{-1}(x))}\right)}{\partial \alpha_i} = -\, w_1^*(u_i, v_i, \gamma_i, \varphi) y_1 - w_2^*(u_i, v_i, \gamma_i, \varphi) y_N - w_3^*(u_i, v_i, \gamma_i, \varphi) d_1$$
$$+ w_4^*(u_i, v_i, \gamma_i, \varphi) d_N,$$

where

$$w_1^*(u_i,\ v_i,\ \gamma_i,\ \varphi) = \frac{u_i(1-\varphi)^3 + (2u_i + v_i + \gamma_i)\varphi(1-\varphi)^2}{V_i(\varphi)},$$
$$w_2^*(u_i,\ v_i,\ \gamma_i,\ \varphi) = \frac{(u_i + 2v_i + \gamma_i)\varphi^2(1-\varphi) + v_i\varphi^3}{V_i(\varphi)},$$
$$w_3^*(u_i,\ v_i,\ \gamma_i,\ \varphi) = \frac{u_i(x_N - x_1)\varphi(1-\varphi)^2}{V_i(\varphi)},$$
$$w_4^*(u_i,\ v_i,\ \gamma_i,\ \varphi) = \frac{v_i(x_N - x_1)\varphi^2(1-\varphi)}{V_i(\varphi)}.$$

Thus, we obtain that

$$\left|\frac{\partial\left(\frac{p_i(L_i^{-1}(x),\, \alpha_i)}{q_i(L_i^{-1}(x))}\right)}{\partial \alpha_i}\right| \leq |w_1^*(u_i, v_i, \gamma_i, \varphi) y_1| + |w_2^*(u_i, v_i, \gamma_i, \varphi) y_N| + |w_3^*(u_i, v_i, \gamma_i, \varphi) d_1|$$
$$+ |w_4^*(u_i, v_i, \gamma_i, \varphi) d_N|.$$

By using the similar procedure for finding bound for $\|C\|_\infty$, we get

$$\left|\frac{\partial\left(\frac{p_i(L_i^{-1}(x),\, \alpha_i)}{q_i(L_i^{-1}(x))}\right)}{\partial \alpha_i}\right| \leq M^* + |I| M_*. \tag{2.35}$$

From (2.32),(2.34) and (2.35), we obtain

$$|T_{\boldsymbol{\alpha}} C(x) - T_{\mathbf{0}} C(x)| \leq |\boldsymbol{\alpha}|_\infty([M + h\bar{M}] + [M^* + |I| M_*]),$$

and hence we get

$$\|T_{\boldsymbol{\alpha}} C - T_{\mathbf{0}} C\|_\infty \leq |\boldsymbol{\alpha}|_\infty([M + h\bar{M}] + [M^* + |I| M_*]). \tag{2.36}$$

Using (2.29) and (2.36), we get

$$\|\Phi - C\|_\infty = \|T_{\boldsymbol{\alpha}}\Phi - T_{\mathbf{0}}C\|_\infty \leq \|T_{\boldsymbol{\alpha}}\Phi - T_{\boldsymbol{\alpha}}C\|_\infty + \|T_{\boldsymbol{\alpha}}C - T_{\mathbf{0}}C\|_\infty,$$

which implies that

$$\|\Phi - C\|_\infty \leq \frac{|\boldsymbol{\alpha}|_\infty([M + h\bar{M}] + [M^* + |I|M_*])}{1 - |\boldsymbol{\alpha}|_\infty}. \tag{2.37}$$

Now, let us concentrate on the uniform error bound between the original function S and it's classical rational cubic spline C. For $x \in [x_i, x_{i+1}]$, let us consider the error function $E(S; x) = S(x) - C(x)$ as a linear functional which operates on S. Then using the Peano Kernel Theorem [104]:

$$L[S] = E(S; x) = S(x) - C(x) = \int_{x_i}^{x_{i+1}} S^{(2)}(\tau) L_x[(x-\tau)_+] d\tau,$$

where

$$L_x[(x-\tau)_+] = \begin{cases} r(\tau, x) & \text{if } x_i < \tau < x, \\ s(\tau, x) & \text{if } x < \tau < x_{i+1}, \end{cases}$$

here L_x is used to emphasize that the functional L is applied to the truncated power function $(x-\tau)_+^n$ considered as a function of x. Also

$$(x-\tau)_+^n := \begin{cases} (x-\tau)^n & \text{if } \tau < x, \\ 0 & \text{if } \tau > x, \end{cases}$$

$$r(\tau, x) = (x - \tau) - \frac{[(u_i + 2v_i + \gamma_i)(x_{i+1} - \tau) - v_i h_i]\varphi^2(1-\varphi) + v_i(x_{i+1} - \tau)\varphi^3}{u_i(1-\varphi)^2 + \varphi(1-\varphi)(\gamma_i + u_i + v_i) + v_i\varphi^2}, \tag{2.38}$$

$$s(\tau, x) = -\frac{[(u_i + 2v_i + \gamma_i)(x_{i+1} - \tau) - v_i h_i]\varphi^2(1-\varphi) + v_i(x_{i+1} - \tau)\varphi^3}{u_i(1-\varphi)^2 + \varphi(1-\varphi)(\gamma_i + u_i + v_i) + v_i\varphi^2}. \tag{2.39}$$

Therefore, we have

$$|S(x) - C(x)| \leq \|S^{(2)}\|_\infty \int_{x_i}^{x_{i+1}} |L_x[(x-\tau)_+]| d\tau. \tag{2.40}$$

The integral involved in (2.40) can be expressed as

$$\int_{x_i}^{x_{i+1}} |L_x[(x-\tau)_+]| d\tau = \int_{x_i}^{x} |r(\tau, x)| d\tau + \int_{x}^{x_{i+1}} |s(\tau, x)| d\tau. \tag{2.41}$$

It can be seen that

- Roots of $r(x,x),\ s(x,x)$ in $[0,1]$ are $\varphi=0$ and $\varphi=1$.
- Roots of $r(\tau,x)$ is $\tau^*=x-\frac{\varphi^2 h_i(u_i+v_i+\gamma_i)}{u_i+\varphi(u_i+v_i+\gamma_i)}$, and $\tau^*\in[x_i,x]$.
- Roots of $s(\tau,x)$ is $\tau_*=x_{i+1}-\frac{v_i h_i(1-\varphi)}{v_i+(u_i+v_i+\gamma_i)(1-\varphi)}$, and $\tau_*\in[x,x_{i+1}]$.

The expressions given in (2.38) and (2.39) can be simplified as

$$r(\tau,x)=\frac{[(1-\varphi)^2u_i+\varphi(1-\varphi)^2(u_i+v_i+\gamma_i)](\tau^*-\tau)}{V_i(\varphi)},$$
$$s(\tau,x)=\frac{[(u_i+v_i+\gamma_i)\varphi^2(1-\varphi)+\varphi^2 v_i](\tau-\tau_*)}{V_i(\varphi)},$$

respectively, where $V_i(\varphi)$ given in (2.28). We have

$$\begin{aligned}\int_{x_i}^{x}|r(\tau,x)|d\tau &= \int_{x_i}^{\tau^*} r(\tau,x)d\tau-\int_{\tau^*}^{x} r(\tau,x)d\tau\\ &=\frac{\varphi^2(1-\varphi)^2u_i^2h_i^2}{2V_i(\varphi)[u_i+\varphi(u_i+v_i+\gamma_i)]}+\frac{\varphi^4(1-\varphi)^2h_i^2(u_i+v_i+\gamma_i)^2}{2V_i(\varphi)[u_i+\varphi(u_i+v_i+\gamma_i)]},\\ \int_{x}^{x_{i+1}}|s(\tau,x)|d\tau &= -\int_{x}^{\tau_*} s(\tau,x)d\tau+\int_{\tau_*}^{x_{i+1}} s(\tau,x)d\tau\\ &=\frac{\varphi^2(1-\varphi)^4h_i^2(u_i+v_i+\gamma_i)^2}{2V_i(\varphi)[v_i+(u_i+v_i+\gamma_i)(1-\varphi)]}\\ &\quad+\frac{\varphi^2(1-\varphi)^2v_i^2h_i^2}{2V_i(\varphi)[v_i+(u_i+v_i+\gamma_i)(1-\varphi)]}.\end{aligned}$$

By (2.40) and (2.41), we get

$$|S(x)-C(x)|\leq c_i^*h_i^2\|S^{(2)}\|_\infty.$$

Since the above inequality is true for $x\in[x_i,x_{i+1}],\ i=1,2,\dots,N-1$, the desired error bound is given by

$$\|S-C\|_\infty\leq c^*h^2\|S^{(2)}\|_\infty. \tag{2.42}$$

Using (2.37) and (2.42) with the following inequality

$$\|S-\Phi\|_\infty\leq\|S-C\|_\infty+\|C-\Phi\|_\infty,$$

we get the required bound for $\|S-\Phi\|$. □

Remark 2.2.4. *Since $|\alpha_i|< a_i,\ i=1,2,\dots,N-1$, it can be seen that $\|S-\Phi\|_\infty=\mathcal{O}(h)$. If $|\alpha_i|< a_i^2$ then we get $\|S-\Phi\|_\infty=\mathcal{O}(h^2)$.*

2.2.2 Shape preserving interpolation

In this section, shape preserving aspects of rational cubic FIF (2.27) are investigated. Sufficient conditions on the scaling factors and the shape parameters are derived to preserve the shape properties.

Proposition 2.2.1. *[125] The polynomial $\varrho(t) = at^3 + bt^2 + ct + d > 0$ for all $t \geq 0$ if and only if $(a, b, c, d) \in T_1 \cup T_2$, where*

$T_1 = \{(a, b, c, d) : a > 0,\ b > 0,\ c > 0 \text{ and } d > 0\}$,

$T_2 = \{(a, b, c, d) : a > 0,\ d > 0,\ 4ac^3 + 4db^3 + 27a^2d^3 + 27a^2d^2 - 18abcd - b^2c^2 > 0\}$.

Positivity

Suppose $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ be a data set such that $y_i > 0$. The following theorem gives sufficient conditions on the scaling factors and the shape parameters so that rational cubic FIF (2.27) would preserve the positivity.

Theorem 2.2.2. *Let $\{(x_i, y_i) : i = 1, 2, \ldots, N-1\}$ be a data set such that $y_i > 0$. Let Φ be the corresponding rational cubic FIF defined in (2.27). Let d_i be the derivative at the knot x_i. Then the conditions on the scaling factors and the shape parameters on each interval $I_i = [x_i, x_{i+1}]$, so that Φ preserve positivity are*

$$0 \leq \alpha_i < \min\left\{a_i,\ \frac{y_i}{y_1},\ \frac{y_{i+1}}{y_N}\right\},$$

$$u_i > 0,\ v_i > 0 \text{ and } \gamma_i > \max\left\{0, \gamma_{1i}^*\ \gamma_{2i}^*\right\},$$

where

$$\gamma_{1i}^* = \frac{-u_i h_i d_i - (2u_i + v_i)y_i + \alpha_i[u_i d_1(x_N - x_1) + (2u_i + v_i)y_1]}{y_i - \alpha_i y_1},$$

$$\gamma_{2i}^* = \frac{v_i h_i d_{i+1} - (u_i + 2v_i)y_{i+1} - \alpha_i[v_i d_N(x_N - x_1) - (u_i + 2v_i)y_N]}{y_{i+1} - \alpha_i y_N}.$$

Proof. Assume that $\alpha_i \geq 0$ for all $i = 1, 2, \ldots, N-1$. Then $\Phi(L_i(x)) > 0$ if $\frac{P_i(\theta)}{Q_i(\theta)} > 0$. The parameters $u_i > 0,\ v_i > 0$ and $\gamma_i \geq 0$ for all $i = 1, 2, \ldots, N-1$, gives denominator

$Q_i(\theta) > 0$. The positivity of the function $\frac{P_i(\theta)}{Q_i(\theta)}$ depends on the numerator $P_i(\theta)$. It can be seen that

$$\begin{aligned} P_i(\theta) &= A_{1,i}(1-\theta)^3 + A_{2,i}\theta(1-\theta)^2 + A_{3,i}\theta^2(1-\theta) + A_{4,i}\theta^3 \\ &= t_{1i}\theta^3 + t_{2i}\theta^2 + t_{3i}\theta + t_{4i}, \end{aligned} \tag{2.43}$$

where

$$\begin{aligned} t_{1i} &= (u_i + v_i + \gamma_i)(y_i - \alpha_i y_1) + (-u_i - v_i - \gamma_i)(y_{i+1} - \alpha_i y_N) \\ &\quad + u_i h_i(d_i - \frac{\alpha_i}{a_i} d_1) + v_i h_i(d_{i+1} - \frac{\alpha_i}{a_i} d_N), \\ t_{2i} &= (-u_i - 2v_i - 2\gamma_i)(y_i - \alpha_i y_1) + (u_i + 2v_i + \gamma_i)(y_{i+1} - \alpha_i y_N) \\ &\quad - 2u_i h_i(d_i - \frac{\alpha_i}{a_i} d_1) - v_i h_i(d_{i+1} - \frac{\alpha_i}{a_i} d_N), \\ t_{3i} &= (-u_i + v_i + \gamma_i)(y_i - \alpha_i y_1) + u_i h_i(d_i - \frac{\alpha_i}{a_i} d_1), \\ t_{4i} &= u_i(y_i - \alpha_i y_1). \end{aligned}$$

By substituting $\theta = \frac{\upsilon}{\upsilon+1}$ in (2.43), we get $P_i(\theta) > 0$ for all $\theta \in [0,1]$ is equivalent to say $\Theta_i(\upsilon) = t^*_{1i}\upsilon^3 + t^*_{2i}\upsilon^2 + t^*_{3i}\upsilon + t^*_{4i} > 0$ for all $\upsilon \geq 0$,

where

$t^*_{1i} = t_{1i} + t_{2i} + t_{3i} + t_{4i} = v_i(y_{i+1} - \alpha_i y_N),$

$t^*_{2i} = t_{2i} + 2t_{3i} + 3t_{4i} = (u_i + 2v_i + \gamma_i)(y_{i+1} - \alpha_i y_N) - v_i h_i(d_{i+1} - \frac{\alpha_i}{a_i} d_N),$

$t^*_{3i} = t_{3i} + 3t_{4i} = (2u_i + v_i + \gamma_i)(y_i - \alpha_i y_1) + u_i h_i(d_i - \frac{\alpha_i}{a_i} d_1),$

$t^*_{4i} = t_{4i} = u_i(y_i - \alpha_i y_1).$

From proposition 2.2.1, $\Theta_i(\upsilon) > 0$ for all $\upsilon \geq 0$ if and only if $(t^*_{1i},\ t^*_{2i},\ t^*_{3i},\ t^*_{4i}) \in T_1 \cup T_2$. Conditions given in the region T_2 is omitted due to complication in calculation. Since conditions given in the region T_1 are computationally economical, the region T_1 is used to get positivity of Θ_i. We get

$$\begin{aligned} t^*_{1i} > 0 &\iff v_i(y_{i+1} - \alpha_i y_N) > 0 \iff \alpha_i < \frac{y_{i+1}}{y_N}, \\ t^*_{4i} > 0 &\iff u_i(y_i - \alpha_i y_1) > 0 \iff \alpha_i < \frac{y_i}{y_1}, \end{aligned}$$

$$t^*_{2i} > 0 \iff (u_i + 2v_i + \gamma_i)(y_{i+1} - \alpha_i y_N) - v_i h_i(d_{i+1} - \frac{\alpha_i}{a_i} d_N) > 0,$$

$$\Longleftrightarrow \gamma_i > \frac{v_i h_i d_{i+1} - (u_i + 2v_i)y_{i+1} - \alpha_i[v_i d_N(x_N - x_1) - (u_i + 2v_i)y_N]}{y_{i+1} - \alpha_i y_N},$$

$$t^*_{3i} > 0 \Longleftrightarrow (2u_i + v_i + \gamma_i)(y_i - \alpha_i y_1) + u_i h_i (d_i - \frac{\alpha_i}{a_i} d_1) > 0,$$

$$\Longleftrightarrow \gamma_i > \frac{-u_i h_i d_i - (2u_i + v_i)y_i + \alpha_i[u_i d_1(x_N - x_1) + (2u_i + v_i)y_1]}{y_i - \alpha_i y_1}.$$

Thus t^*_{1i}, t^*_{2i}, t^*_{3i} and t^*_{4i} lies in the region T_1 if the scaling factors and the shape parameters satisfy the conditions given in the statement of Theorem 2.2.2. □

Remark 2.2.5. *When* $\alpha_i = 0$, $i = 1, 2, \ldots, N-1$, *the conditions given in Theorem 2.2.2 reduces to*

$$u_i > 0,\ v_i > 0,\ \text{and}\ \gamma_i > \max\left\{0,\ \frac{-u_i h_i d_i - (2u_i + v_i)y_i}{y_i},\ \frac{v_i h_i d_{i+1} - (u_i + 2v_i)y_{i+1}}{y_{i+1}}\right\}.$$

This is the required condition for the classical rational cubic spline (2.28) to preserve the positivity.

Monotonicity

Let $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ be a monotonic data. Without loss of generality, assume that the data is monotonically increasing, i.e., $y_1 \le y_2 \le \ldots \le y_N$. Then $\Delta_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \ge 0$, $i = 1, 2, \ldots, N-1$. The aim of this section is to find the suitable parameters so that the rational cubic FIF (2.27) would preserve the monotonicity. We have

$$\Phi^{(1)}(L_i(x)) = \frac{\alpha_i \Phi^{(1)}(x)}{a_i} + \frac{\Psi_i(\theta)}{(Q_i(\theta))^2},$$

where $\Psi_i(\theta) = \sum_{k=1}^{5} B_{k,i}\theta^{k-1}(1-\theta)^{5-k}$,

$$B_{1,i} = u_i^2 d_i^*,$$

$$B_{2,i} = 2u_i\{(u_i + 2v_i + \gamma_i)\Delta_i^* - v_i d_{i+1}^*\},$$

$$B_{3,i} = B_{2,i} + B_{4,i} - (B_{1,i} + B_{5,i}) + \gamma_i(u_i + v_i + \gamma_i)\Delta_i^* - \gamma_i(u_i d_i^* + v_i d_{i+1}^*),$$

$$B_{4,i} = 2v_i\{(2u_i + v_i + \gamma_i)\Delta_i^* - u_i d_i^*\},$$

$$B_{5,i} = v_i^2 d_{i+1}^*,$$

$d_i^* = d_i - \frac{\alpha_i d_1}{a_i}$, $d_{i+1}^* = d_{i+1} - \frac{\alpha_i d_N}{a_i}$ and $\Delta_i^* = \Delta_i - \alpha_i \frac{y_N - y_1}{h_i}$.

Theorem 2.2.3. *Let $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ be a monotonically increasing data, i.e., $y_1 \le y_2 \le \ldots \le y_N$. Let Φ be the corresponding rational cubic FIF defined in (2.27). Let d_i be the derivative at the knot point x_i. Let derivatives satisfy the necessary conditions for monotonicity, namely*

$$sgn(d_i) = sgn(d_{i+1}) = sgn(\Delta_i), \quad \textit{for } \Delta_i \neq 0.$$

Then sufficient conditions on the scaling factors and the shape parameters so that Φ would preserve the monotonicity on the interval I are

$$0 \le \alpha_i < \left\{a_i, \frac{a_i d_i}{d_1}, \frac{a_i d_{i+1}}{d_N}, \frac{y_{i+1} - y_i}{y_N - y_1}\right\},$$

$$u_i > 0,\ v_i > 0 \textit{ and } \gamma_i \ge \max\left\{0, \frac{u_i d_i^* + v_i d_{i+1}^* - \Delta_i^*(u_i + v_i)}{\Delta_i^*}\right\},$$

for all $i = 1, 2, \ldots, N-1$.

Proof. Let $\{(x_i,\ y_i), i = 1, 2, \ldots, N\}$ be a monotonically increasing data. From single variable calculus, it is obvious that function Φ is monotonic if $\Phi^{(1)}(x) \ge 0$, for $x \in [x_1, x_N]$. Assume that $\alpha_i \ge 0,\ i = 1, 2, \ldots, N-1$. For each node $x_j,\ j = 1, 2, \ldots, N$

$$\Phi'(L_i(x_j)) = \frac{\alpha_i \Phi'(x_j)}{a_i} + \frac{\Psi_i(\theta_j)}{(Q_i(\theta_j))^2},\ \theta_j = \frac{x_j - x_1}{x_N - x_1}. \tag{2.44}$$

From (2.44), $\Phi'(L_i(x_j)) \ge 0$ if $\frac{\Psi_i(\theta_j)}{(Q_i(\theta_j))^2} \ge 0$. It is obvious that $(Q_i(\theta_j))^2 > 0$ for all $\theta \in [0, 1]$. Then sufficient conditions for $\Psi_i(\theta_j) \ge 0,\ \theta \in [0, 1]$ are $B_{k,i} \ge 0,\ k = 1, 2, \ldots, 5$. We have

$$B_{1,i} \ge\ 0 \Leftrightarrow \alpha_i \le \frac{a_i d_i}{d_1}.$$

We get

$$B_{5,i} \ge\ 0 \Leftrightarrow \alpha_i \le \frac{a_i d_{i+1}}{d_N}.$$

Let $0 \le \alpha_i < \{\frac{a_i d_i}{d_1}, \frac{a_i d_{i+1}}{d_N}, \frac{y_{i+1} - y_i}{y_N - y_1}\},\ i = 1, 2, \ldots, N-1.$

We obtain $B_{2,i} \ge 0$ if

$$\gamma_i \ge \frac{d_{i+1}^* v_i - \Delta_i^*(u_i + 2v_i)}{\Delta_i^*}.$$

We get $B_{4,i} \ge 0$ if

$$\gamma_i \ge \frac{d_i^* u_i - \Delta_i^*(2u_i + v_i)}{\Delta_i^*}.$$

$B_{3,i}$ can be written as

$$B_{3,i} = u_i\{(u_i + 2v_i + \gamma_i)\Delta_i^* - v_i d_{i+1}^*\} + v_i\{(2u_i + v_i + \gamma_i)\Delta_i^* - u_i d_i^*\}$$
$$+ u_i\{(u_i + v_i + \gamma_i)\Delta_i^* - u_i d_i^* - v_i d_{i+1}^*\} + v_i\{(u_i + v_i + \gamma_i)\Delta_i^* - u_i d_i^* - v_i d_{i+1}^*\}$$
$$+ \gamma_i\{(u_i + v_i + \gamma_i)\Delta_i^* - u_i d_i^* - v_i d_{i+1}^*\} + 2u_i v_i \Delta_i^*.$$

It can be seen that

$$\frac{u_i d_i^* + v_i d_{i+1}^* - \Delta_i^*(u_i + v_i)}{\Delta_i^*} > \frac{d_{i+1}^* v_i - \Delta_i^*(u_i + 2v_i)}{\Delta_i^*},$$
$$\frac{u_i d_i^* + v_i d_{i+1}^* - \Delta_i^*(u_i + v_i)}{\Delta_i^*} > \frac{d_i^* u_i - \Delta_i^*(2u_i + v_i)}{\Delta_i^*}.$$

Thus $B_{3,i} \geq 0$ if

$$\gamma_i \geq \max\left\{0, \frac{u_i d_i^* + v_i d_{i+1}^* - \Delta_i^*(u_i + v_i)}{\Delta_i^*}\right\}.$$

From the above discussion, it is evident that $\Phi^{(1)}(L_i(x_j)) \geq 0$ for all $i = 1, 2, \ldots, N-1$, $j = 1, 2, \ldots, N$, if the scaling factors and the shape parameters satisfy the conditions given in the statement of Theorem 2.2.3. $\{I; L_i(x) : i = 1, 2, \ldots, N-1\}$ is an IFS and $[x_1,\ x_N]$ is an its attractor. Since FIF is defined as recursive structure, $\Phi^{(1)}(L_i(x_j)) \geq 0$ for all $i = 1, 2, \ldots, N-1$ and for every knot x_j imply that $\Phi^{(1)}(x) \geq 0$ for all $x \in [x_1,\ x_N]$. □

Convexity

The aim of this section is to find the conditions on the scaling factors and the shape parameters for the rational cubic FIF (2.27) to preserve the convexity.

Theorem 2.2.4. *Suppose $\{(x_i, y_i) : i = 1, 2, \ldots, N\}$ is a strictly convex data. Let the derivatives satisfy $d_1 < \Delta_1 < \cdots < d_i < \Delta_i < d_{i+1} < \ldots \Delta_{N-1} < d_N$. Let Φ be the corresponding rational cubic FIF defined in (2.27). Then the sufficient conditions on the scaling factors and the shape parameters so that Φ preserve the convexity on the interval I are*

$$0 \leq \alpha_i < \min\left\{a_i^2, \frac{h_i(d_{i+1} - \Delta_i)}{d_N(x_N - x_1) - (y_N - y_1)}, \frac{h_i(\Delta_i - d_i)}{(y_N - y_1) - d_1(x_N - x_1)}\right\},$$
$$u_i > 0,\ v_i > 0 \text{ and } \gamma_i > \max\left\{\frac{v_i(d_{i+1}^* - \Delta_i^*)}{(\Delta_i^* - d_i^*)}, \frac{u_i(\Delta_i^* - d_i^*)}{(d_{i+1}^* - \Delta_i^*)}\right\}.$$

Proof. From calculus, Φ is convex if $\Phi^{(2)}(x^+)$ or $\Phi^{(2)}(x^-)$ is exist and non-negative for all $x \in (x_1, x_N)$ [96]. So informally

$$\Phi^{(2)}(L_i(x)) = \frac{\alpha_i \Phi^{(2)}(x)}{a_i^2} + \frac{\Psi_i^*(\theta)}{h_i(Q_i(\theta))^3},$$

where $\Psi_i^*(\theta) = \sum_{k=1}^{6} C_{k,i}\theta^{k-1}(1-\theta)^{6-k}$,

$$\begin{aligned}
C_{1,i} &= 2u_i^3(\Delta_i^* - d_i^*) + 2u_i^2 v_i(\Delta_i^* - d_i^*) + 2u_i^2[\gamma_i(\Delta_i^* - d_i^*) - v_i(d_{i+1}^* - \Delta_i^*)],\\
C_{2,i} &= 2C_{1,i} + 6u_i^2 v_i(\Delta_i^* - d_i^*),\\
C_{3,i} &= C_{1,i} + 12u_i^2 v_i(\Delta_i^* - d_i^*) + 6u_i v_i^2(d_{i+1}^* - \Delta_i^*),\\
C_{4,i} &= C_{6,i} + 12u_i v_i^2(d_{i+1}^* - \Delta_i^*) + 6u_i^2 v_i(\Delta_i^* - d_i^*),\\
C_{5,i} &= 2C_{6,i} + 6u_i v_i^2(d_{i+1}^* - \Delta_i^*),\\
C_{6,i} &= 2v_i^3(d_{i+1}^* - \Delta_i^*) + 2v_i^2 u_i(d_{i+1}^* - \Delta_i^*) + 2v_i^2[\gamma_i(d_{i+1}^* - \Delta_i^*) - u_i(\Delta_i^* - d_i^*)],
\end{aligned}$$

$d_i^* = d_i - \frac{\alpha_i d_1}{a_i}$, $d_{i+1}^* = d_{i+1} - \frac{\alpha_i d_N}{a_i}$, $\Delta_i^* = \Delta_i - \alpha_i \frac{y_N - y_1}{h_i}$.

It can be observed that

$$\Phi^{(2)}(x_1^+) = \frac{C_{1,1}}{h_1 u_1^3}\left[1 - \frac{\alpha_1}{a_1^2}\right]^{-1}, \tag{2.45}$$

$$\Phi^{(2)}(x_N^-) = \frac{C_{6,N-1}}{h_{N-1} v_{N-1}^3}\left[1 - \frac{\alpha_{N-1}}{a_{N-1}^2}\right]^{-1}, \tag{2.46}$$

$$\Phi^{(2)}(x_n^+) = \frac{\alpha_n}{a_n^2}\Phi^{(2)}(x_1^+) + \frac{C_{1,n}}{u_n^3 h_n}, \quad n = 2, 3, \ldots, N-1. \tag{2.47}$$

Let $0 \le \alpha_i < a_i^2$, $i = 1, 2, \ldots, N-1$. From equations (2.45), (2.46) and (2.47) it is obvious that if $C_{1,i} \ge 0$ $(i = 1, 2, \ldots, N-1)$ and $C_{6,N-1} \ge 0$, then the right-handed second derivatives at the knots x_i, $i = 1, 2, \ldots, N-1$, and the left-handed second derivative at x_N are non-negative. For a knot points x_n, $n = 1, 2, \ldots, N-1$ we have

$$\Phi^{(2)}(L_i(x_n)^+) = \frac{\alpha_i \Phi^{(2)}(x_n^+)}{a_i^2} + R_i(x_n),$$

where $R_i(x) = \frac{\Psi_i^*(\theta)}{h_i(Q_i(\theta))^3}$.

Assuming $C_{1,i} \ge 0$, $i = 1, 2, \ldots, N-1$, then $\Phi^{(2)}(L_i(x_n)^+) \ge 0$ if $R_i(x_n) \ge 0$. Note that $R_i(x_n) \ge 0$ if the coefficients $C_{j,i} \ge 0$ for $j = 1, 2, \ldots, 6$. From Three Chords

Lemma [145], it follows that for a strict convex interpolant, the end point derivatives should necessarily satisfy $d_1 < \frac{y_N - y_1}{x_N - x_1} < d_N$. Hence we get

$$\Big(\Delta_i - \alpha_i \frac{y_N - y_1}{h_i}\Big) - \Big(d_i - \frac{\alpha_i d_1}{a_i}\Big) > 0 \text{ if } \alpha_i < \frac{h_i(\Delta_i - d_i)}{(y_N - y_1) - d_1(x_N - x_1)},$$

$$\Big(d_{i+1} - \frac{\alpha_i d_N}{a_i}\Big) - \Big(\Delta_i - \alpha_i \frac{y_N - y_1}{h_i}\Big) > 0 \text{ if } \alpha_i < \frac{h_i(d_{i+1} - \Delta_i)}{d_N(x_N - x_1) - (y_N - y_1)}.$$

Let

$$0 \le \alpha_i < \min\Big\{a_i^2, \frac{h_i(\Delta_i - d_i)}{(y_N - y_1) - d_1(x_N - x_1)}, \frac{h_i(d_{i+1} - \Delta_i)}{d_N(x_N - x_1) - (y_N - y_1)}\Big\}, \tag{2.48}$$
$$i = 1, 2, \dots, N-1.$$

The assumption on scaling factors given in (2.48) and if $\gamma_i \ge \frac{v_i(d^*_{i+1} - \Delta^*_i)}{(\Delta^*_i - d^*_i)}$ ensures $C_{1,i} \ge 0$, $C_{2,i} \ge 0$ and $C_{3,i} \ge 0$. Similarly by the assumption on scaling factors given in (2.48) and if $\gamma_i \ge \frac{u_i(\Delta^*_i - d^*_i)}{(d^*_{i+1} - \Delta^*_i)}$ ensures $C_{4,i} \ge 0$, $C_{5,i} \ge 0$ and $C_{6,i} \ge 0$.

Thus the condition on the scaling factors and the shape parameters given Theorem 2.2.4 statement ensures $C_{j,i} \ge 0$ for all $j = 1, 2, \dots, 6$. This in turn ensures that the non-negative of $\Phi^{(2)}(L_i(x_n)^+)$ for $i, n = 1, 2, \dots, N-1$, and $\Phi^{(2)}(x_N^-)$. The non-negativity of $\Phi^{(2)}(x^+)$ or $\Phi^{(2)}(x^-)$ is follows from the non-negativity of $\Phi^{(2)}(L_i(x_n)^+)$ for $i, n = 1, 2, \dots, N-1$, and $\Phi^{(2)}(x_N^-)$. Hence restrictions on the scaling factors and the shape parameters given in statement of the Theorem 2.2.4 ensure the convexity of Φ. □

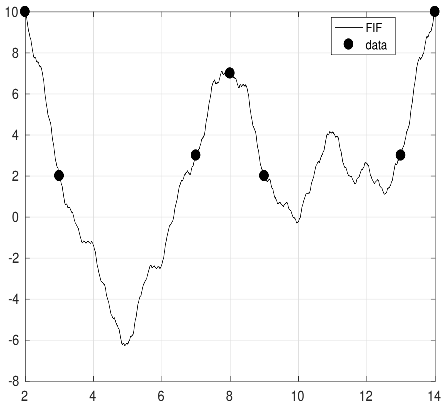
Remark 2.2.6. *Suppose $\Delta_i - d_i = 0$ or $d_{i+1} - \Delta_i = 0$, then we take $\alpha_i = 0$, $d_i = d_{i+1} = \Delta_i$. In this case, Φ become straight line $\Phi(L_i(x)) = y_i(1-\theta) + y_{i+1}\theta$ in the interval $[x_i, x_{i+1}]$.*

2.2.3 Numerical examples

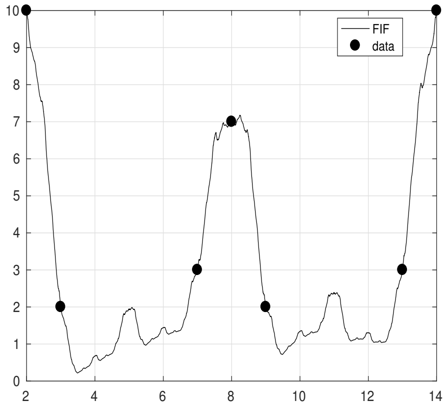
We consider the positive data $\{(2,10),(3,2),(7,3),(8,7),(9,2),(13,3),(14,10)\}$. The derivatives are calculated based on arithmetic mean method and derivatives are $d_1 = -9.6500,\ d_2 = -6.3500,\ d_3 = 3.2500,\ d_4 = -0.5000\ d_5 = -3.9500\ d_6 = 5.6500,\ d_7 = 8.3500$. To preserve positivity, scaling factors are taken according to Theorem 2.2.2 and scaling factors ranges are $\alpha_1 \in [0,\ 0.0833),\ \alpha_2 \in [0,\ 0.2000),\ \alpha_3 \in [0,\ 0.0833),\ \alpha_4 \in [0,\ 0.0833),\ \alpha_5 \in [0,\ 0.2000),\ \alpha_6 \in [0,\ 0.0833)$.

By taking random parameters, FIF (Figure 2.8(a)) is constructed and it can be observed that these FIF is non-positive. Thus, the parameters are restricted according to Theorem 2.5 and using these parameters, positive FIF (Figure 2.8(b)) is constructed. By perturbing the scaling factor α_2 with respect to Figure 2.8(b), FIF given in Figure 2.8(c) is constructed. It can be seen that, there is a significant change in the intervals $[x_2, x_3]$ and $[x_5, x_6]$ but changes in the other intervals are negligible. Also, after perturbing γ_2 with respect to Figure 2.8(b), FIF (Figure 2.8(d)) is constructed. It can be observed that, there is a change in the interval $[x_2, x_3]$ and changes in the other intervals are not visible. Next by taking all the scaling factors $\alpha_i = 0$ for all $i = 1, 2, \ldots, 6$, Figure 2.8(e) is plotted, which is classical rational cubic spline preserving positivity. Figure 2.9 represents the derivative of the FIFs that are given in Figure 2.8. The parameters are used to plot Figures 2.8 and 2.9 are given in Table 2.5.

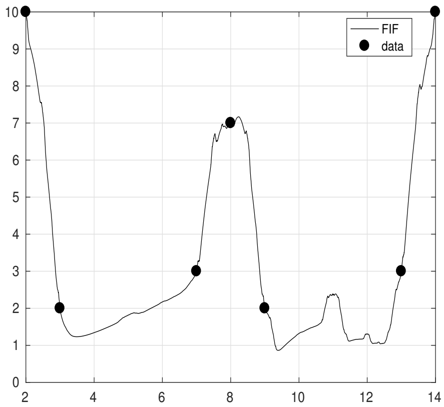
Next, we consider the monotonic increasing data $\{(7.99, 0),\ (8.09, 2.76429 \times 10^{-5}),\ (8.19, 4.37498 \times 10^{-2}),\ (8.70, 0.169183),\ (9.20, 0.469428),\ (10, 0.943740),\ (12,\ 0.998636),\ (15, 0.999919),\ (20, 0.999994)\}$. Derivatives are approximated using arithmetic mean method and derivatives are $d_1 = 0,\ d_2 = 0.2187,\ d_3{=}0.4059,\ d_4 = 0.4250,\ d_5 = 0.5976,\ d_6 = 0.4313,\ d_7 = 0.0166,\ d_8 = 0.0003\ d_9 = 0$. In order to preserve monotonicity, scaling factors are taken according to Theorem 2.2.3 and scaling factors ranges are $\alpha_1 \in [0,\ 0.28 \times 10^{-4}),\ \alpha_2 \in [0,\ 0.8326 \times 10^{-2}),\ \alpha_3 \in [0,\ 0.4246 \times 10^{-1}),\ \alpha_4 \in [0,\ 0.4163 \times 10^{-1}),\ \alpha_5 \in [0,\ 0.6661 \times 10^{-1}),\ \alpha_6 \in [0,\ 0.5489 \times 10^{-1}),\ \alpha_7 \in [0,\ 0.1283 \times 10^{-2}),\ \alpha_8 \in [0, 0.75 \times 10^{-4})$.



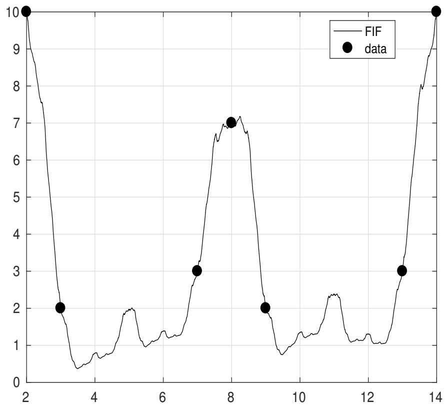
(a) Non-positive FIF.



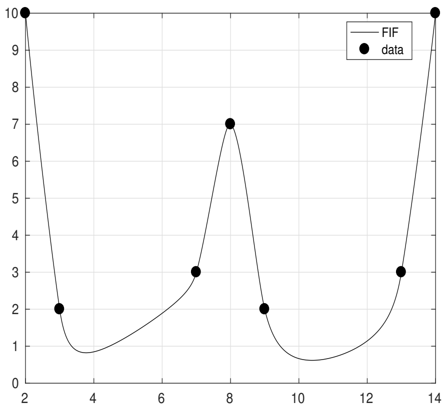
(b) Positive FIF



(c) Effect of α_2.

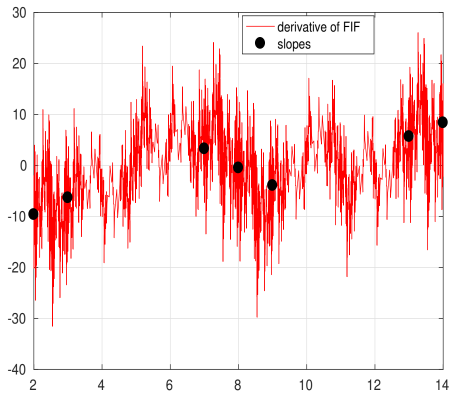


(d) Effect of γ_2.

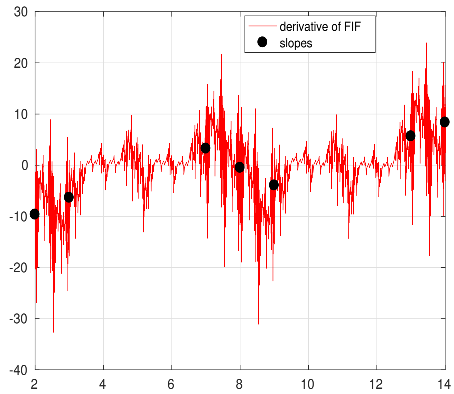


(e) Classical rational cubic spline.

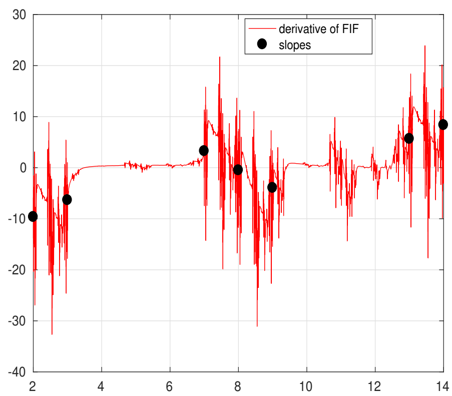
Figure 2.8: Positive preserving interpolation.



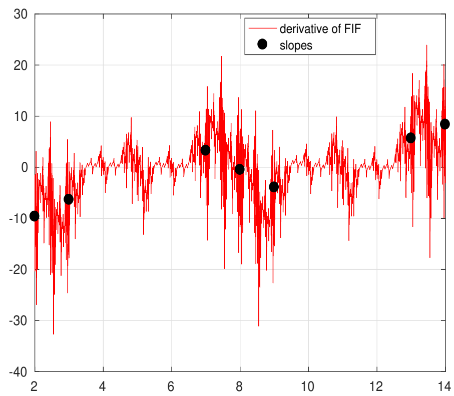
(a) Derivative of the FIF given in Figure 2.8(a).



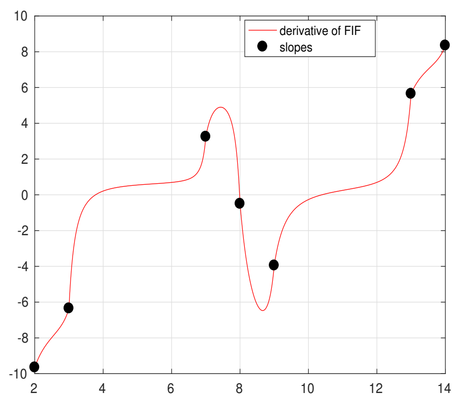
(b) Derivative of the FIF given in Figure 2.8(b).



(c) Derivative of the FIF given in Figure 2.8(c).



(d) Derivative of the FIF given in Figure 2.8(d).



(e) Derivative of the FIF given in Figure 2.8(e).

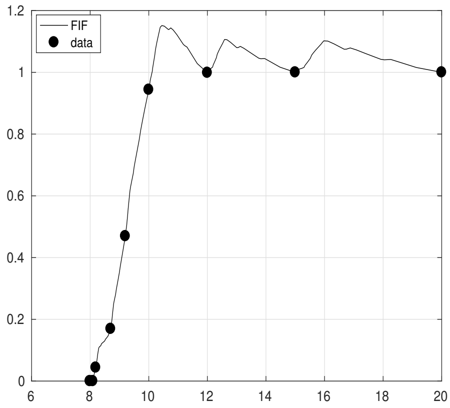
Figure 2.9: Derivatives of the FIFs given in Figure 2.8.

Table 2.5: Parameters for positive interpolation with $u_i = 1.5$ and $v_i = 1$ for $i = 1, 2, \ldots, 6$.

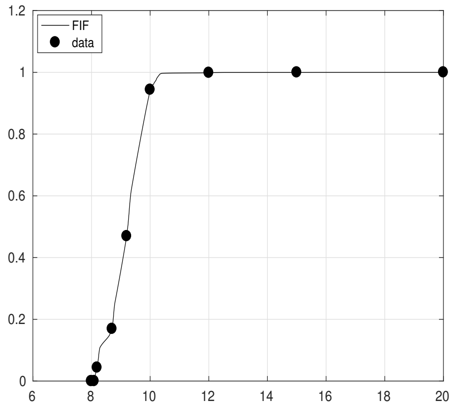
↓ Parameter/Figure →	2.8(a), 2.9(a)	2.8(b), 2.9(b)	2.8(c), 2.9(c)	2.8(d), 2.9(d)	2.8(e), 2.9(e)
α_1	0.08	0.08	0.08	0.08	0
α_2	-0.3	0.18	0.018	0.18	0
α_3	0.08	0.08	0.08	0.08	0
α_4	0.08	0.08	0.08	0.08	0
α_5	0.3	0.19	0.19	0.19	0
α_6	0.08	0.08	0.08	0.08	0
γ_1	2	1	1	1	1
γ_2	2	32	32	300	16
γ_3	2	2	2	2	2
γ_4	2	1	1	1	1
γ_5	2	1	1	1	8
γ_6	2	1	1	1	1

Taking arbitrary scaling factors and shape parameters FIF (Figure 2.10(a)) is plotted but the FIF is not preserving the monotonicity of the data. Hence the parameters are selected based on Theorem 2.2.3, using these parameters, FIF (Figure 2.10(b)) is plotted. By observing the FIF (Figure 2.10(b)), we can see that FIF satisfies monotonicity. After perturbing α_3 with respect to Figure 2.10(b), FIF given in Figure 2.10(c) is constructed. There is a significant change in $[x_3, x_4]$ but changes in other intervals are negligible. Similarly, after perturbing γ_3 with respect to Figure 2.10(b), FIF given in Figure 2.10(d) is constructed. There is a slight change in $[x_3, x_4]$ and changes in other intervals are not visible. Finally, by taking all scaling factor $\alpha_i = 0$ $i = 1, 2, \ldots, 8$, classical rational cubic spline is plotted which is given in Figure 2.10(e). Derivative of the FIFs given in Figure 2.10 are constructed and plotted in Figure 2.11. Scaling factors and shape parameters are used to compute Figures 2.10 and 2.11 are given in Table 2.6.

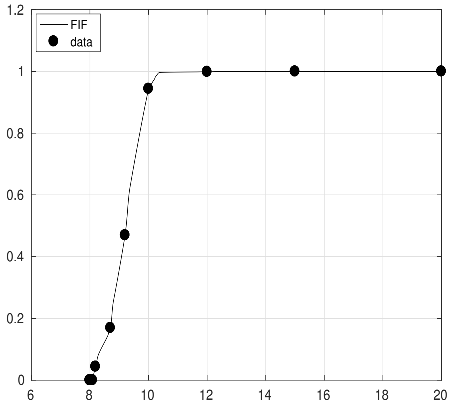
Next consider the convex data $\{(-8, 4.5), (-7, 4), (2.2, 3.55), (7, 4), (10, 4.5), (12, 5)\}$. The derivatives are estimated using arithmetic mean method and derivatives are $d_1 = -0.5442,\ d_2 = -0.4558,\ d_3 = 0.0448,\ d_4 = 0.1386,\ d_5 = 0.2167,\ d_6 = 0.2833$.



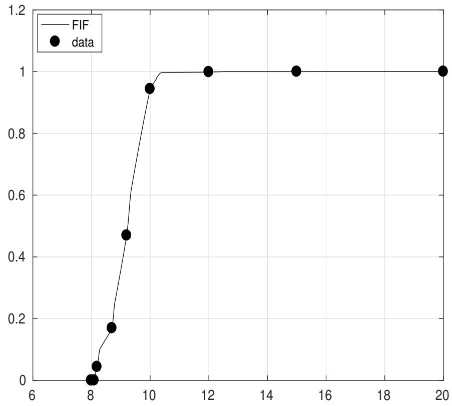
(a) Non-monotone FIF.



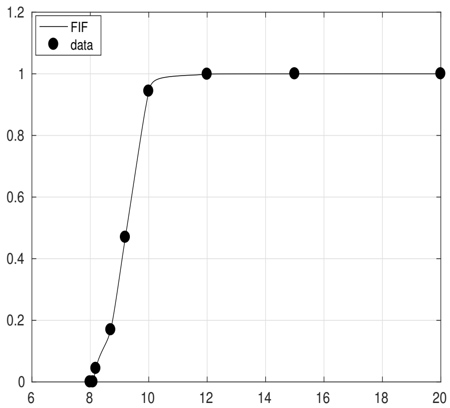
(b) Monotone FIF.



(c) Effect of α_3.

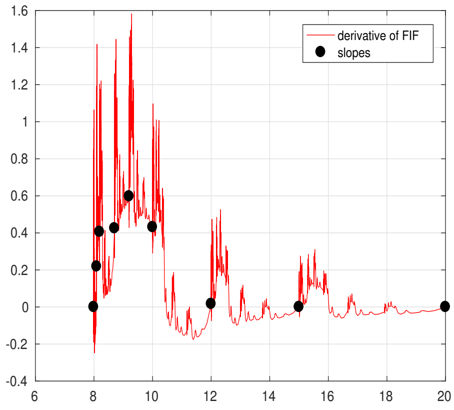


(d) Effect of γ_3.

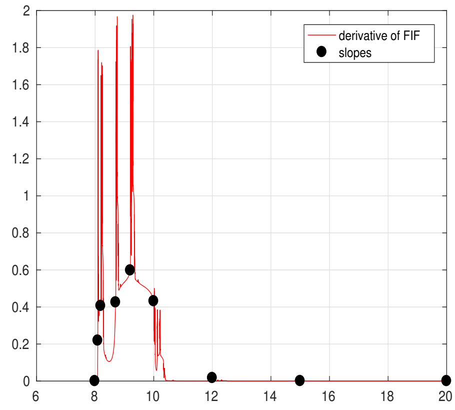


(e) Classical rational cubic spline.

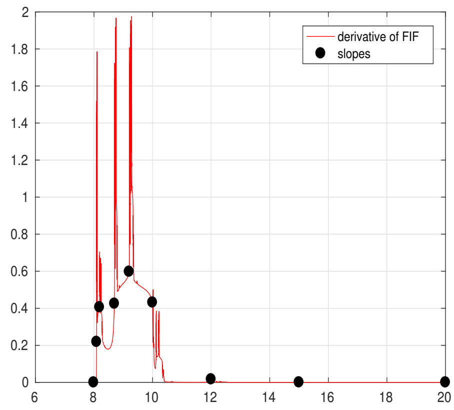
Figure 2.10: Monotonicity preserving interpolation.



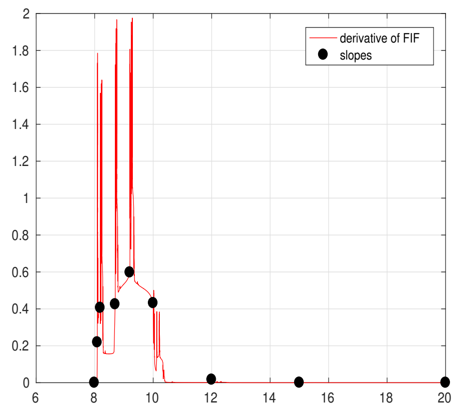
(a) Derivative of FIF given in Figure 2.10(a).



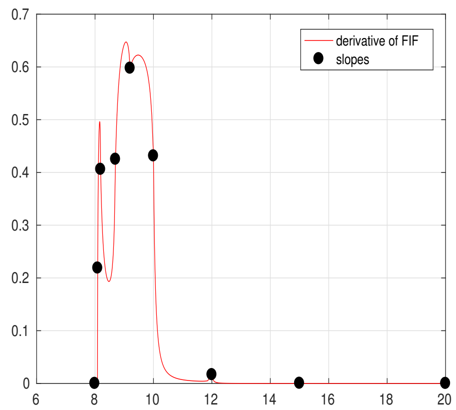
(b) Derivative of FIF given in Figure 2.10(b).



(c) Derivative of FIF given in Figure 2.10(c).



(d) Derivative of FIF given in Figure 2.10(d).



(e) Derivative of FIF given in Figure 2.10(e).

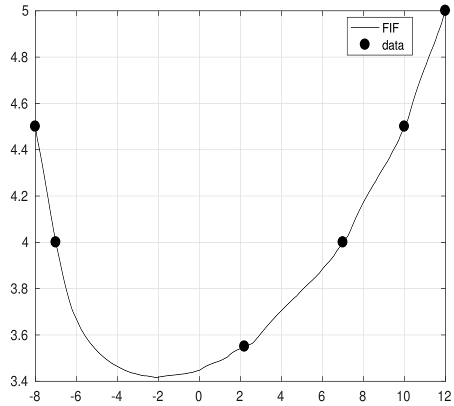
Figure 2.11: Derivatives of the FIFs given in Figure 2.10.

Table 2.6: Parameters for monotonic interpolation with $u_i = 1.5$ and $v_i = 1$ for $i = 1, 2, \ldots, 8$.

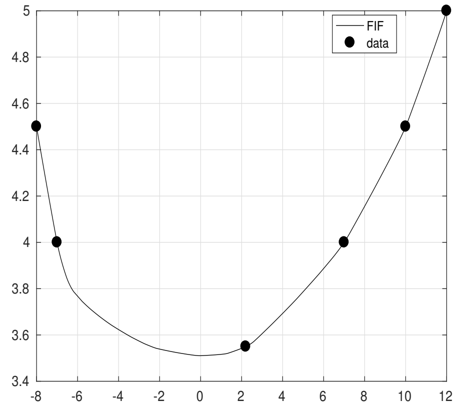
↓ Parameter/Figure →	2.10(a),2.11(a)	2.10(b),2.11(b)	2.10(c),2.11(c)	2.10(d),2.11(d)	2.10(e),2.11(e)
α_1	0.7e-2	0.2e-4	0.2e-4	0.2e-4	0
α_2	0.7e-2	0.8e-2	0.8e-2	0.8e-2	0
α_3	0.3e-1	0.4e-1	0.1e-1	0.4e-1	0
α_4	0.3e-1	0.4e-1	0.4e-1	0.4e-1	0
α_5	0.5e-1	0.6e-1	0.6e-1	0.6e-1	0
α_6	0.1	0.4e-1	0.4e-1	0.4e-1	0
α_7	0.1	0.1e-2	0.1e-2	0.1e-2	0
α_8	0.1	0.7e-4	0.7e-4	0.7e-4	0
γ_1	2	3000	3000	3000	800
γ_2	2	2	2	2	2
γ_3	2	5	5	50	3
γ_4	2	2	2	2	2
γ_5	2	2	2	2	2
γ_6	2	90	90	90	23
γ_7	2	270	270	270	58
γ_8	2	410	410	410	26

Scaling factors and shape parameters are calculated based on Theorem 2.2.4 and scaling factor ranges are $\alpha_1 \in [0, 0.0025)$, $\alpha_2 \in [0, 0.1669)$, $\alpha_3 \in [0, 0.0206)$, $\alpha_4 \in [0, 0.0074)$, $\alpha_5 \in [0, 0.0059)$. By taking arbitrary parameters, Figure 2.12(a) is plotted. It can be seen that, FIF given in Figure 2.12(a) is non-convex. Hence, scaling factors and shape parameters are taken according to Theorem 2.2.4, using these parameters, FIF given in Figure 2.12(b) is plotted which is convex.

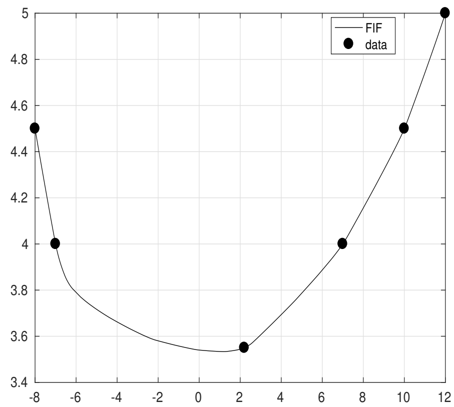
Figure 2.12(c) is plotted after perturbing the scaling factor α_2 with respect to Figure 2.12(b). Also, Figure 2.12(d) is plotted by perturbing shape parameter γ_2 with respect to Figure 2.12(b). The classical rational cubic spline given in Figure 2.12(e) is plotted using scaling factors $\alpha_i = 0$ for all $i = 1, 2, \ldots, 5$. Derivative of the FIFs given in Figure 2.12 are constructed and given in Figure 2.13. Scaling factors and shape parameters are used to compute Figures 2.12 and 2.13 are given in Table 2.7.



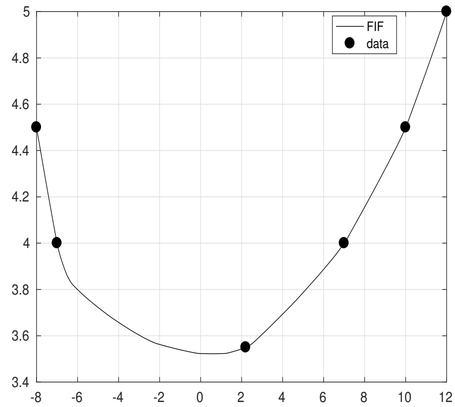
(a) Non-convex FIF.



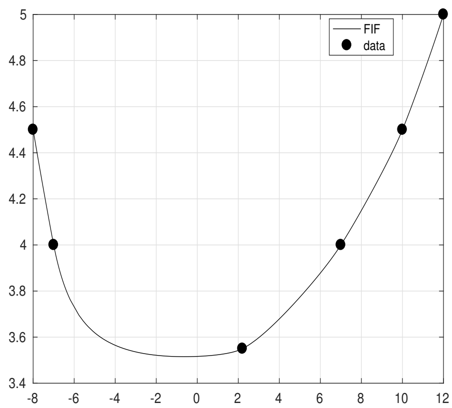
(b) Convex FIF.



(c) Effect of α_2.

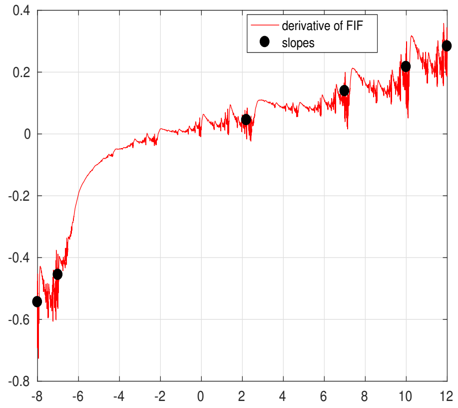


(d) Effect of γ_2.

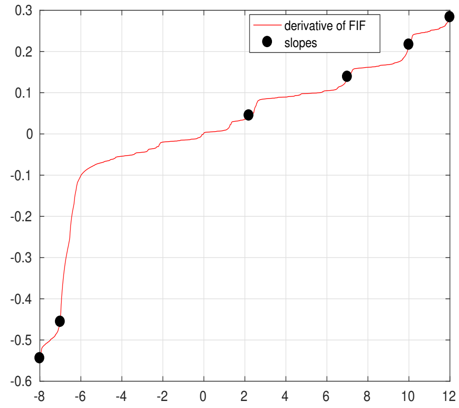


(e) Classical rational cubic spline.

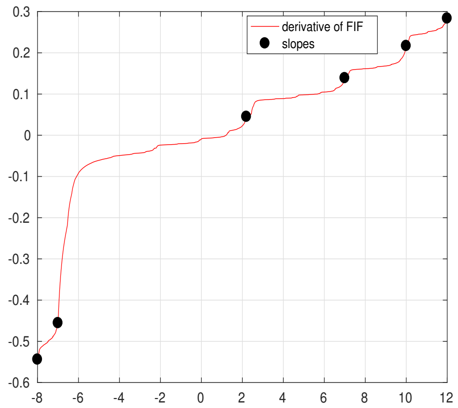
Figure 2.12: Convexity preserving interpolation.



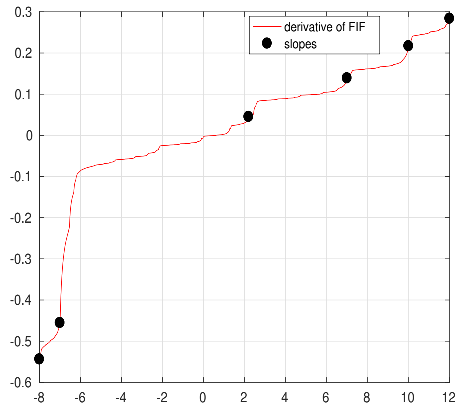
(a) Derivative of FIF given in Figure 2.12(a).



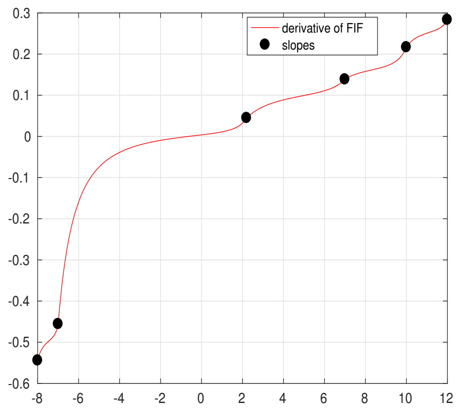
(b) Derivative of FIF given in Figure 2.12(b).



(c) Derivative of FIF given in Figure 2.12(c).



(d) Derivative of FIF given in Figure 2.12(d).



(e) Derivative of FIF given in Figure 2.12(e).

Figure 2.13: Derivatives of the FIFs given in Figure 2.12.

Table 2.7: Parameters for convex interpolation with $u_i = 1.5$ and $v_i = 1$ for $i = 1, 2, \ldots, 5$.

↓ Parameter/Figure →	2.12(a), 2.13(a)	2.12(b), 2.13(b)	2.12(c), 2.13(c)	2.12(d), 2.13(d)	2.12(e), 2.13(e)
α_1	0.05	0.002	0.002	0.002	0
α_2	0.2	0.14	0.09	0.14	0
α_3	0.1	0.02	0.02	0.02	0
α_4	0.1	0.006	0.006	0.006	0
α_5	0.09	0.005	0.005	0.005	0
γ_1	3	2	2	2	2
γ_2	4	25	25	50	8
γ_3	3	17	17	17	2
γ_4	3	8	8	8	2
γ_5	3	5	5	5	2

2.3 Conclusion

To preserve the shapes of the univariate data, rational cubic FIFs with two families of the shape parameters and three families of the shape parameters are constructed. The error analysis of the developed interpolation schemes are discussed to see the effectiveness of the interpolation schemes. Numerical experiments are implemented to see the applicability of the interpolation schemes in the field of shape preserving. The classical rational cubic splines having applications in areas such as CAGD, data analysis, image processing etc, hence our FIFs can also be applied in these areas.

Chapter 3

Shape Preserving Rational Cubic Fractal Interpolation Surfaces

In this chapter, we have developed a new FIS for preserving the shapes of the bivariate data that lies on a rectangular grid. The transfinite interpolation method [57] is used via blending functions scheme. This method defines surfaces by utilizing the curves already available. This transfinite interpolation method play vital role in shape preserving bivariate interpolation because of the fact that the shape properties of the transfinite interpolating surface follow trivially from the corresponding shape properties of the networks of boundary curves [28].

In Section 3.1, the rational cubic FIFs are constructed to approximate the original function along the grid lines of the interpolation domain. Then rational cubic FIS is constructed with the help of rational cubic FIFs and blending functions. Convergence analysis of the constructed FIS towards an original function is discussed. In Section 3.2, the scaling factors and the shape parameters involved in the FIFs are restricted suitably such that these FIFs are positive, monotone, convex whenever the given interpolation data along the grid lines are positive, monotone, convex respectively. These restrictions on the scaling factors and the shape parameters ensure the positivity, monotonicity and convexity of the constructed FIS. Also the constrained interpolation problem for the constructed FIS is also discussed. In Section 3.3, numerical examples are provided to see the applicability of the constructed interpolation scheme in the field of shape preserving interpolation.

3.1 Fractal Interpolation Surfaces

In this section, at the beginning, $\mathcal{C}^1$-rational cubic FIFs (fractal boundary curves) along the x-direction and the y-direction are constructed on a rectangular grid that contains the surface data. Then the rational cubic FIS is constructed upon a rectangular grid as a combination of the fractal boundary curves and the blending functions. In order to see the effectiveness of the constructed FIS, convergence analysis of the rational cubic FIS is reported.

3.1.1 Construction of fractal boundary curves

Let $I = [x_1, x_M]$ and $J = [y_1, y_N]$. Let $x_1 < x_2 < \cdots < x_M$, $y_1 < y_2 < \cdots < y_N$ be a partition of $D = [x_1, x_M] \times [y_1, y_N]$. Set $I_i = [x_i, x_{i+1}]$, $J_j = [y_j, y_{j+1}]$, $D_{i,j} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$, for $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N-1$. Consider the bivariate interpolation data $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M, j = 1, 2, \ldots, N\}$ on a rectangular grid D, where $z^x_{i,j}$ and $z^y_{i,j}$ are the partial derivatives of the original function with respect to x and y at the point (x_i, y_j) respectively. The aim is to find a continuously differentiable function Φ (say) on D such that $\Phi(x_i, y_j) = z_{i,j}$, $\frac{\partial \Phi(x_i,y_j)}{\partial x} = z^x_{i,j}$, $\frac{\partial \Phi(x_i,y_j)}{\partial y} = z^y_{i,j}$ for all $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$.

For each fixed $j = 1, 2, \ldots, N$, consider the data $T_j = \{(x_i, z_{i,j}, z^x_{i,j}) : i = 1, 2, \ldots, M\}$ along the j-th grid line parallel to the x-axis. In order to construct the fractal boundary curve along the grid $I \times y_j$ that interpolate the data T_j, consider the IFS $\mathcal{J}_j = \{I \times \mathbb{R}; w_{i,j}(x, z) = (L_i(x), F_{i,j}(x, z)),\ i = 1, 2, \ldots, M-1\}$, $L_i(x) = a_i x + b_i$, $F_{i,j}(x, z) = \alpha_{i,j} z + r_{i,j}(x)$ with

$$r_{i,j}(x) = \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)} = \frac{A_{1,i,j}(1-\theta)^3 + A_{2,i,j}\theta(1-\theta)^2 + A_{3,i,j}\theta^2(1-\theta) + A_{4,i,j}\theta^3}{u_{i,j}(1-\theta)^2 + (u_{i,j} + v_{i,j} + \gamma_{i,j})\theta(1-\theta) + v_{i,j}\theta^2}, \tag{3.1}$$

where $\theta = (x - x_1)/(x_M - x_1)$, $x \in [x_1, x_M]$, $i = 1, 2, \ldots, M-1$. The constants $A_{1,i,j}$, $A_{2,i,j}$, $A_{3,i,j}$ and $A_{4,i,j}$ are to be determined. Here $u_{i,j}$, $v_{i,j}$ and $\gamma_{i,j}$ are called the shape parameters. The parameters $u_{i,j}$, $v_{i,j}$ and $\gamma_{i,j}$ are assumed that $u_{i,j} > 0$, $v_{i,j} > 0$ and $\gamma_{i,j} \geq 0$.

For each fixed $j = 1, 2, \ldots, N$, we impose the following conditions on the IFS $\mathcal{J}_j = \{I \times \mathbb{R}; w_{i,j}(x, z) = (L_i(x), F_{i,j}(x, z)),\ i = 1, 2, \ldots, M-1\}$:

$$|\alpha_{i,j}| < a_i,\ L_i(x_1) = x_i,\ L_i(x_M) = x_{i+1}, F_{i,j}(x_1, z_{1,j}) = z_{i,j},$$
$$F_{i,j}(x_M, z_{M,j}) = z_{i+1,j},\ F^1_{i,j}(x_1, z^x_{1,j}) = z^x_{i,j},\ F^1_{i,j}(x_M, z^x_{M,j}) = z^x_{i+1,j},$$

where $F^1_{i,j}(x, z) = \frac{\alpha_{i,j} z + r^{(1)}_{i,j}(x)}{a_i}$. The IFS $\mathcal{J}_j = \{I \times \mathbb{R}; w_{i,j}(x, z) = (L_i(x), F_{i,j}(x, z)), i = 1, 2, \ldots, M-1\}$ satisfies the conditions for $\mathcal{C}^1$-FIF (see Theorem 1.4.2).

For each fixed $j = 1, 2, \ldots, N$, let $\mathscr{F}_j = \{\phi : I \to \mathbb{R} \mid \phi \in \mathcal{C}^1(I),\ \phi(x_1) = z_{1,j},\ \phi(x_M) = z_{M,j},\ \phi^{(1)}(x_1) = z^x_{1,j}$ and $\phi^{(1)}(x_M) = z^x_{M,j}\}$. Then $(\mathscr{F}_j, \rho)$ is a complete metric space, where ρ is the metric induced by the $\mathcal{C}^1$-norm $\|\phi\| = \|\phi\|_\infty + \|\phi^{(1)}\|_\infty$ on $\mathcal{C}^1(I)$.

Define the Read-Bajraktarević operator on $\mathscr{F}_j$ as $T : \mathscr{F}_j \to \mathscr{F}_j$ such that

$$(T\phi)(L_i(x)) = \alpha_{i,j}\phi(x) + r_{i,j}(x), \quad x \in I,\ i = 1, 2, \ldots, M-1.$$

Since T is contraction, the fixed-point B_j (say) of T satisfies the following functional equation:

$$B_j(L_i(x)) = \alpha_{i,j} B_j(x) + r_{i,j}(x), \quad x \in I,\ i = 1, 2, \ldots, M-1. \tag{3.2}$$

The FIF $B^{(1)}_j$ satisfies the following functional equation:

$$B^{(1)}_j(L_i(x)) = \frac{\alpha_{i,j} B^{(1)}_j(x) + r^{(1)}_{i,j}(x)}{a_i}, \quad x \in I,\ i = 1, 2, \ldots, M-1.$$

The constants $A_{1,i,j}$, $A_{2,i,j}$, $A_{3,i,j}$ and $A_{4,i,j}$ in $r_{i,j}(x)$ of (3.1) are evaluated using the conditions $B_j(x_i) = z_{i,j}$, $B_j(x_{i+1}) = z_{i+1,j}$, $B^{(1)}_j(x_i) = z^x_{i,j}$ and $B^{(1)}_j(x_{i+1}) = z^x_{i+1,j}$ (these conditions are equivalent to $F_{i,j}(x_1, z_{1,j}) = z_{i,j}$, $F_{i,j}(x_M, z_{M,j}) = z_{i+1,j}$, $F^1_{i,j}(x_1, z^x_{1,j}) = z^x_{i,j}$, $F^1_{i,j}(x_M, z^x_{M,j}) = z^x_{i+1,j}$ [145]).

Putting $x = x_i$ in (3.2), we have

$$A_{1,i,j} = u_{i,j}[z_{i,j} - \alpha_{i,j} z_{1,j}]. \tag{3.3}$$

Substituting $x = x_{i+1}$ in (3.2), we get

$$A_{4,i,j} = v_{i,j}[z_{i+1,j} - \alpha_{i,j} z_{M,j}]. \tag{3.4}$$

Using $B_j^{(1)}(x_i) = z_{i,j}^x$, we get

$$A_{2,i,j} = (2u_{i,j}+v_{i,j}+\gamma_{i,j})z_{i,j}+u_{i,j}h_i z_{i,j}^x-\alpha_{i,j}[(2u_{i,j}+v_{i,j}+\gamma_{i,j})z_{1,j}+u_{i,j}(x_M-x_1)z_{1,j}^x]. \tag{3.5}$$

Making use of $B_j^{(1)}(x_{i+1}) = z_{i+1,j}^x$, we obtain that

$$A_{3,i,j} = (u_{i,j}+2v_{i,j}+\gamma_{i,j})z_{i+1,j}-v_{i,j}h_i z_{i+1,j}^x-\alpha_{i,j}[(u_{i,j}+2v_{i,j}+\gamma_{i,j})z_{M,j}-v_{i,j}(x_M-x_1)z_{M,j}^x]. \tag{3.6}$$

Thus, the fractal boundary curve B_j along the grid $I \times y_j$ satisfies the following functional equation:

$$B_j(L_i(x)) = \alpha_{i,j} B_j(x) + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)}, \tag{3.7}$$

where

$$\begin{aligned}
p_{i,j}(\theta) = {} & u_{i,j}[z_{i,j} - \alpha_{i,j} z_{1,j}](1-\theta)^3 + v_{i,j}[z_{i+1,j} - \alpha_{i,j} z_{M,j}]\theta^3 + [(2u_{i,j} + v_{i,j} + \gamma_{i,j})z_{i,j} \\
& + u_{i,j} h_i z_{i,j}^x - \alpha_{i,j}((2u_{i,j} + v_{i,j} + \gamma_{i,j})z_{1,j} + u_{i,j}(x_M - x_1)z_{1,j}^x)]\theta(1-\theta)^2 \\
& + [(u_{i,j} + 2v_{i,j} + \gamma_{i,j})z_{i+1,j} - v_{i,j} h_i z_{i+1,j}^x - \alpha_{i,j}((u_{i,j} + 2v_{i,j} + \gamma_{i,j})z_{M,j} \\
& - v_{i,j}(x_M - x_1)z_{M,j}^x)]\theta^2(1-\theta), \\
q_{i,j}(\theta) = {} & u_{i,j}(1-\theta)^2 + (u_{i,j} + v_{i,j} + \gamma_{i,j})\theta(1-\theta) + v_{i,j}\theta^2, \ \theta = \frac{x - x_1}{x_M - x_1}, \ x \in [x_1, x_M], \\
& i = 1, 2, \ldots, M-1.
\end{aligned}$$

Thus, there are N different fractal boundary curves B_j's along the grid lines parallel to x-axis such that their graphs are attractors of the IFSs $\mathcal{J}_j = \{I \times \mathbb{R}; w_{i,j}(x,z) = (L_i(x), F_{i,j}(x,z)),\ i = 1, 2, \ldots, M-1\}$, $j = 1, 2, \ldots, N$. The parameters $u_{i,j}$, $v_{i,j}$, $\gamma_{i,j}$ and the scaling factors $\alpha_{i,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$ can be written in the matrix form as $\boldsymbol{u} = [u_{i,j}]_{(M-1)\times N}$, $\boldsymbol{v} = [v_{i,j}]_{(M-1)\times N}$, $\boldsymbol{\gamma} = [\gamma_{i,j}]_{(M-1)\times N}$ and $\boldsymbol{\alpha} = [\alpha_{i,j}]_{(M-1)\times N}$.

Remark 3.1.1. *For each fixed $j = 1, 2, \ldots, N$, if $\alpha_{i,j} = 0$, $i = 1, 2, \ldots, M-1$, then the fractal boundary curve B_j becomes the classical rational cubic spline as*

$$C_j(x) = \frac{P_{i,j}(\varphi)}{Q_{i,j}(\varphi)}, \tag{3.8}$$

where

$$P_{i,j}(\varphi) = u_{i,j}z_{i,j}(1-\varphi)^3 + v_{i,j}z_{i+1,j}\varphi^3 + [(2u_{i,j}+v_{i,j}+\gamma_{i,j})z_{i,j} + u_{i,j}h_i z^x_{i,j}]\varphi(1-\varphi)^2$$
$$+ [(u_{i,j}+2v_{i,j}+\gamma_{i,j})z_{i+1,j} - v_{i,j}h_i z^x_{i+1,j}]\varphi^2(1-\varphi),$$
$$Q_{i,j}(\varphi) = u_{i,j}(1-\varphi)^2 + (u_{i,j}+v_{i,j}+\gamma_{i,j})\varphi(1-\varphi) + v_{i,j}\varphi^2,\ \varphi = \frac{x-x_i}{x_{i+1}-x_i},$$
$$x \in [x_i, x_{i+1}],\ i = 1, 2, \ldots, M-1.$$

Next, for each $i = 1, 2, \ldots, M$, $T^*_i = \{(y_j, z_{i,j}, z^y_{i,j}) : j = 1, 2, \ldots, N\}$ is the interpolation data along the i-th grid line parallel to y-axis. In order to construct the fractal boundary curve along the grid $x_i \times J$, to interpolate the data T^*_i, consider the IFS $\mathcal{J}^*_i = \{J \times \mathbb{R}; w^*_{i,j}(y,z) = (L^*_j(y), F^*_{i,j}(y,z)),\ i = 1, 2, \ldots, N-1\}$, $L^*_j(y) = c_j y + d_j$, $F^*_{i,j}(y,z) = \alpha^*_{i,j}z + r^*_{i,j}(y)$ with

$$r^*_{i,j}(y) = \frac{p^*_{i,j}(\zeta)}{q^*_{i,j}(\zeta)} = \frac{A^*_{1,i,j}(1-\zeta)^3 + A^*_{2,i,j}\zeta(1-\zeta)^2 + A^*_{3,i,j}\zeta^2(1-\zeta) + A^*_{4,i,j}\zeta^3}{u^*_{i,j}(1-\zeta)^2 + (u^*_{i,j}+v^*_{i,j}+\gamma^*_{i,j})\zeta(1-\zeta) + v^*_{i,j}\zeta^2},$$

where $\zeta = (y-y_1)/(y_N-y_1)$, $y \in [y_1, y_N]$, $j = 1, 2, \ldots, N-1$. Here, $A^*_{1,i,j}$, $A^*_{2,i,j}$, $A^*_{3,i,j}$ and $A^*_{4,i,j}$ are constants to be determined. The shape parameters $u^*_{i,j}$, $v^*_{i,j}$ and $\gamma^*_{i,j}$ are assumed that $u^*_{i,j} > 0$, $v^*_{i,j} > 0$ and $\gamma^*_{i,j} \geq 0$.

For each fixed $i = 1, 2, \ldots, M$, we impose the following conditions on the IFS $\mathcal{J}^*_i = \{J \times \mathbb{R}; w^*_{i,j}(y,z) = (L^*_j(y), F^*_{i,j}(y,z)),\ j = 1, 2, \ldots, N-1\}$:

$$|\alpha^*_{i,j}| < c_j,\ L^*_j(y_1) = y_j,\ L^*_j(y_N) = y_{j+1}, F^*_{i,j}(y_1, z_{i,1}) = z_{i,j},$$
$$F^*_{i,j}(y_N, z_{i,N}) = z_{i,j+1}, F^{1*}_{i,j}(y_1, z^y_{i,1}) = z^y_{i,j},\ F^{1*}_{i,j}(y_N, z^y_{i,N}) = z^y_{i,j+1},$$

where $F^{1*}_{i,j}(y,z) = \frac{\alpha^*_{i,j}z + r^{*\,(1)}_{i,j}(y)}{c_j}$. The IFS $\mathcal{J}^*_i = \{J \times \mathbb{R}; w^*_{i,j}(y,z) = (L^*_j(y), F^*_{i,j}(y,z))$, $j = 1,\ 2,\ \ldots,\ N-1\}$ satisfies the conditions for $\mathcal{C}^1$-FIF that are given in Theorem 1.4.2.

By using the same procedure followed to construct x-direction fractal boundary curves, the fractal boundary curve B^*_i which interpolates T^*_i is constructed and is given by

$$B^*_i(L^*_j(y)) = \alpha^*_{i,j}B^*_i(y) + \frac{p^*_{i,j}(\zeta)}{q^*_{i,j}(\zeta)}, \tag{3.9}$$

where

$$p^*_{i,j}(\zeta) = u^*_{i,j}[z_{i,j} - \alpha^*_{i,j}z_{i,1}](1-\zeta)^3 + v^*_{i,j}[z_{i,j+1} - \alpha^*_{i,j}z_{i,N}]\zeta^3 + [(2u^*_{i,j} + v^*_{i,j} + \gamma^*_{i,j})z_{i,j}$$

$$+ u^*_{i,j} h^*_j z^y_{i,j} - \alpha^*_{i,j}((2u^*_{i,j} + v^*_{i,j} + \gamma^*_{i,j}) z_{i,1} + u^*_{i,j}(y_N - y_1) z^y_{i,1})]\zeta(1-\zeta)^2$$

$$+ [(u^*_{i,j} + 2v^*_{i,j} + \gamma^*_{i,j}) z_{i,j+1} - v^*_{i,j} h^*_j z^y_{i,j+1} - \alpha^*_{i,j}((u^*_{i,j} + 2v^*_{i,j} + \gamma^*_{i,j}) z_{i,N}$$

$$- v^*_{i,j}(y_N - y_1) z^y_{i,N})]\zeta^2(1-\zeta),$$

$$q^*_{i,j}(\zeta) = u^*_{i,j}(1-\zeta)^2 + (u^*_{i,j} + v^*_{i,j} + \gamma^*_{i,j})\zeta(1-\zeta) + v^*_{i,j}\zeta^2,\ \zeta = \frac{y - y_1}{y_N - y_1},\ y \in [y_1, y_N],$$

$h^*_j = y_{j+1} - y_j$, $L^*_j(y) = c_j y + d_j$ satisfying $L^*_j(y_1) = y_j$, $L^*_j(y_N) = y_{j+1}$, $\alpha^*_{i,j}$ is the scaling factor along the y-direction such that $|\alpha^*_{i,j}| < c_j$, $j = 1, 2, \ldots, N-1$.

Thus there are M different fractal boundary curves B^*_i's along the grid lines parallel to y-axis such that their graphs are the attractors of the IFSs $\mathcal{J}^*_i = \{J \times \mathbb{R}; w^*_{i,j}(y, z) = (L^*_j(y), F^*_{i,j}(y, z)),\ j = 1, 2, \ldots, N-1\}$, $i = 1, 2, \ldots, M$. The parameters $u^*_{i,j}$, $v^*_{i,j}$, $\gamma^*_{i,j}$ and the scaling factors $\alpha^*_{i,j}$, $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N-1$ can be written in the matrix form as $\boldsymbol{u}^* = [u^*_{i,j}]_{M \times (N-1)}$, $\boldsymbol{v}^* = [v^*_{i,j}]_{M \times (N-1)}$, $\boldsymbol{\gamma}^* = [\gamma^*_{i,j}]_{M \times (N-1)}$ and $\boldsymbol{\alpha}^* = [\alpha^*_{i,j}]_{M \times (N-1)}$.

Remark 3.1.2. *For each fixed i, if $\alpha^*_{i,j} = 0$ for $j = 1, 2, \ldots, N-1$, then the fractal boundary curve B^*_i becomes the classical rational cubic spline as*

$$C^*_i(y) = \frac{P^*_{i,j}(\vartheta)}{Q^*_{i,j}(\vartheta)},$$

where

$$P^*_{i,j}(\vartheta) = u^*_{i,j} z_{i,j}(1-\vartheta)^3 + v^*_{i,j} z_{i+1,j}\vartheta^3 + [(2u^*_{i,j} + v^*_{i,j} + \gamma^*_{i,j}) z_{i,j} + u^*_{i,j} h^*_j z^x_{i,j}]\vartheta(1-\vartheta)^2$$
$$+ [(u^*_{i,j} + 2v^*_{i,j} + \gamma^*_{i,j}) z_{i+1,j} - v^*_{i,j} h^*_j z^x_{i+1,j}]\vartheta^2(1-\vartheta),$$
$$Q^*_{i,j}(\vartheta) = u^*_{i,j}(1-\vartheta)^2 + (u^*_{i,j} + v^*_{i,j} + \gamma^*_{i,j})\vartheta(1-\vartheta) + v^*_{i,j}\vartheta^2,\ \vartheta = \frac{y - y_j}{y_{j+1} - y_j},$$
$$y \in [y_j, y_{j+1}],\ j = 1, 2, \ldots, N-1.$$

3.1.2 Construction of rational cubic FIS

Using the fractal boundary curves (3.7) and (3.9), the rational cubic FIS patch on each sub-rectangle $D_{i,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N-1$ is defined by

$$\Phi(x, y) = -M_1 \Psi(x, y) M_2^T,\ \text{for } (x, y) \in D_{i,j}, \tag{3.10}$$

where

$$\Psi(x,y) = \begin{bmatrix} 0 & B_j(x) & B_{j+1}(x) \\ B_i^*(y) & z_{i,j} & z_{i,j+1} \\ B_{i+1}^*(y) & z_{i+1,j} & z_{i+1,j+1} \end{bmatrix},$$

$M_1 = [-1, a_{x,0}(\theta), a_{x,1}(\theta)]$, $a_{x,0}(\theta) = (1-\theta)^2(1+2\theta)$, $a_{x,1}(\theta) = \theta^2(3-2\theta)$, $\theta = (L_i^{-1}(x) - x_1)/(x_M - x_1) = (x - x_i)/(x_{i+1} - x_i)$, $x \in [x_i, x_{i+1}]$, $M_2 = [-1, b_{y,0}(\zeta), b_{y,1}(\zeta)]$, $b_{y,0}(\zeta) = (1-\zeta)^2(1+2\zeta)$, $b_{y,1}(\zeta) = \zeta^2(3-2\zeta)$, $\zeta = (L_j^{*-1}(y) - y_1)/(y_N - y_1) = (y - y_j)/(y_{j+1} - y_j)$, $y \in [y_j, y_{j+1}]$.

The functions $a_{x,0}(\theta)$, $a_{x,1}(\theta)$, $b_{y,0}(\zeta)$ and $b_{y,1}(\zeta)$ are called the blending functions. It is easy to verify that, on each sub-rectangle $D_{i,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N-1$, the following conditions are being satisfied by Φ:

$$\Phi(x_r, y_s) = z_{r,s}, \quad \frac{\partial \Phi(x_r, y_s)}{\partial x} = z_{r,s}^x, \quad \frac{\partial \Phi(x_r, y_s)}{\partial y} = z_{r,s}^y, \quad r = i, i+1, \quad s = j, j+1.$$

It can be seen that

$$\begin{aligned} \Phi(x,y) = {} & b_{y,0}(\zeta)B_j(x) + b_{y,1}(\zeta)B_{j+1}(x) + a_{x,0}(\theta)B_i^*(y) \\ & + a_{x,1}(\theta)B_{i+1}^*(y) - a_{x,0}(\theta)b_{y,0}(\zeta)z_{i,j} - a_{x,0}(\theta)b_{y,1}(\zeta)z_{i,j+1} \\ & - a_{x,1}(\theta)b_{y,0}(\zeta)z_{i+1,j} - a_{x,1}(\theta)b_{y,1}(\zeta)z_{i+1,j+1} \end{aligned} \tag{3.11}$$

is continuous over $D_{i,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N-1$, because it is a combination of continuous functions. To prove continuity over D, it is sufficient to prove the continuity of Φ along the common boundaries of $D_{i,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N-1$. Consider the adjacent sub-rectangles $D_{i,j}$ and $D_{i+1,j}$. On $D_{i+1,j}$, Φ can be written as

$$\begin{aligned} \Phi(x,y) = {} & b_{y,0}(\zeta)B_j(x) + b_{y,1}(\zeta)B_{j+1}(x) + a_{x,0}(\theta)B_{i+1}^*(y) \\ & + a_{x,1}(\theta)B_{i+2}^*(y) - a_{x,0}(\theta)b_{y,0}(\zeta)z_{i+1,j} - a_{x,0}(\theta)b_{y,1}(\zeta)z_{i+1,j+1} \\ & - a_{x,1}(\theta)b_{y,0}(\zeta)z_{i+2,j} - a_{x,1}(\theta)b_{y,1}(\zeta)z_{i+2,j+1}, \end{aligned} \tag{3.12}$$

where $B_j(x)$, $B_{j+1}(x)$, $B_{i+1}^*(y)$ and $B_{i+2}^*(y)$ are the fractal boundary curves defined in $I \times y_j$, $I \times y_{j+1}$, $x_{i+1} \times J$ and $x_{i+2} \times J$ respectively.

Using (3.11) and (3.12), for $y^* \in [y_j, y_{j+1}]$ and $\zeta^* = (y^* - y_j)/(y_{j+1} - y_j)$,

$$\lim_{\substack{(x,y)\to(x_{i+1},y^*) \\ (x,y)\in D_{i,j}}} \Phi(x,y) = B_{i+1}^*(y^*) = B_{i+1}^*(y^*) = \lim_{\substack{(x,y)\to(x_{i+1},y^*) \\ (x,y)\in D_{i+1,j}}} \Phi(x,y).$$

Hence Φ is continuous over $D_{i,j} \cup D_{i+1,j}$ for $i = 1, 2, \dots, M-2$, $j = 1, 2, \dots, N-1$. Similarly, it is easy to see that Φ is continuous over $D_{i,j} \cup D_{i,j+1}$, $i = 1, 2, \dots, M-1$, $j = 1, 2, \dots, N-2$. Thus, Φ is continuous over the domain D. By using the similar procedure followed for obtaining continuity of Φ, it can be seen that partial derivatives $\partial\Phi/\partial x$ and $\partial\Phi/\partial y$ are continuous over D. Since both the first-order partial derivatives of Φ are continuous, $\Phi^{(1)}$ exists and continuous over the domain D. Thus the rational cubic FIS $\Phi \in \mathcal{C}^1(D)$.

Remark 3.1.3. *If* $\boldsymbol{\alpha} = [0]_{(M-1)\times N}$ *and* $\boldsymbol{\alpha}^* = [0]_{M\times(N-1)}$ *then the rational cubic FIS* Φ *defined in (3.10) becomes the classical rational cubic interpolation surface as*

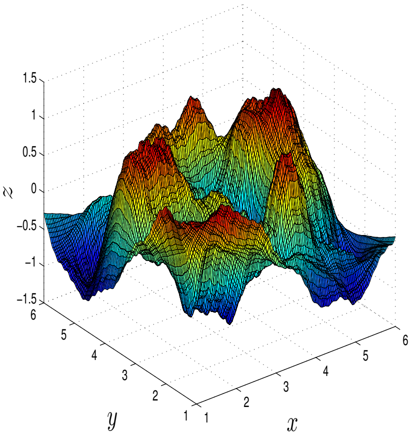
$$\begin{aligned} C(x,y) = {} & b_{y,0}(\zeta)C_j(x) + b_{y,1}(\zeta)C_{j+1}(x) + a_{x,0}(\theta)C_i^*(y) \\ & + a_{x,1}(\theta)C_{i+1}^*(y) - a_{x,0}(\theta)b_{y,0}(\zeta)z_{i,j} - a_{x,0}(\theta)b_{y,1}(\zeta)z_{i,j+1} \\ & - a_{x,1}(\theta)b_{y,0}(\zeta)z_{i+1,j} - a_{x,1}(\theta)b_{y,1}(\zeta)z_{i+1,j+1},\ (x,y) \in I_i \times J_j, \end{aligned}$$

where $C_j(x)$, $j = 1, 2, \dots, N$ *and* $C_i^*(y)$, $i = 1, 2, \dots, M$ *are the classical rational cubic splines for the data* T_j, $j = 1, 2, \dots, N$ *and* T_i^*, $i = 1, 2, \dots, M$ *respectively.*

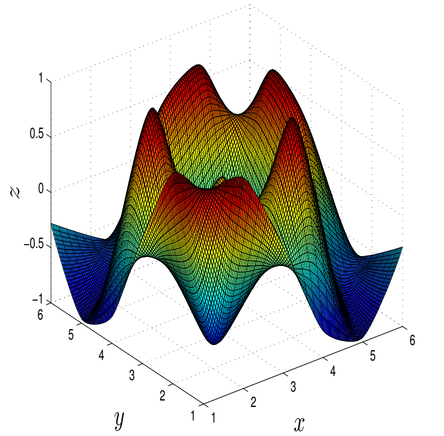
Example 3.1.1. *Let us consider the surface data as given in Table 3.1. The first-order partial derivatives are calculated using arithmetic-mean method. By taking arbitrary scaling factors and shape parameters, rational cubic FIS and classical rational cubic interpolation surface are constructed and which are shown in Figures 3.1(a) and 3.1(b), respectively. Figures 3.1(c) and 3.1(d) represent the first-order partial derivative* $\frac{\partial\Phi}{\partial x}$ *of the FISs given in Figures 3.1(a) and 3.1(b), respectively. From Figures 3.1(c) and 3.1(d), it can be seen that the proposed scheme is the best tool to approximate a function whose partial derivatives are irregular. The parameters* $\boldsymbol{u} = [1.5]_{5\times6}$, $\boldsymbol{u}^* = [1.5]_{6\times5}$, $\boldsymbol{v} = [1]_{5\times6}$, $\boldsymbol{v}^* = [1]_{6\times5}$, $\boldsymbol{\gamma} = [3]_{5\times6}$, $\boldsymbol{\gamma}^* = [3]_{6\times5}$, $\boldsymbol{\alpha} = [0.199]_{5\times6}$ *and* $\boldsymbol{\alpha}^* = [-0.199]_{6\times5}$ *are used to construct Figures 3.1(a) and 3.1(c). The parameters* $\boldsymbol{u} = [1.5]_{5\times6}$, $\boldsymbol{u}^* = [1.5]_{6\times5}$, $\boldsymbol{v} = [1]_{5\times6}$, $\boldsymbol{v}^* = [1]_{6\times5}$, $\boldsymbol{\gamma} = [3]_{5\times6}$, $\boldsymbol{\gamma}^* = [3]_{6\times5}$, $\boldsymbol{\alpha} = [0]_{5\times6}$ *and* $\boldsymbol{\alpha}^* = [0]_{6\times5}$ *are used to construct Figures 3.1(b) and 3.1(d).*

Table 3.1: Surface data

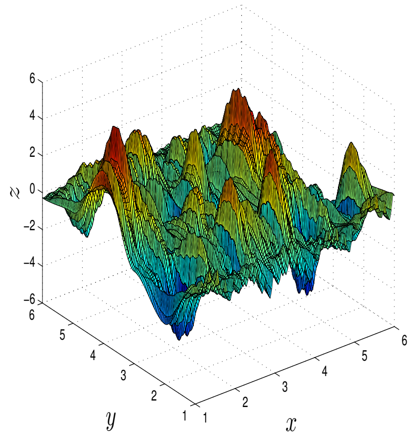
$\downarrow x/y \rightarrow$	1	2	3	4	5	6
1	0.8415	0.9093	0.1411	-0.7568	-0.9589	-0.2794
2	0.9093	-0.7568	-0.2794	0.9894	-0.5440	-0.5366
3	0.1411	-0.2794	0.4121	-0.5366	0.6503	-0.7510
4	-0.7568	0.9894	-0.5366	-0.2879	0.9129	-0.9056
5	-0.9589	-0.5440	0.6503	0.9129	-0.1324	-0.9880
6	-0.2794	-0.5366	-0.7510	-0.9056	-0.9880	-0.9918



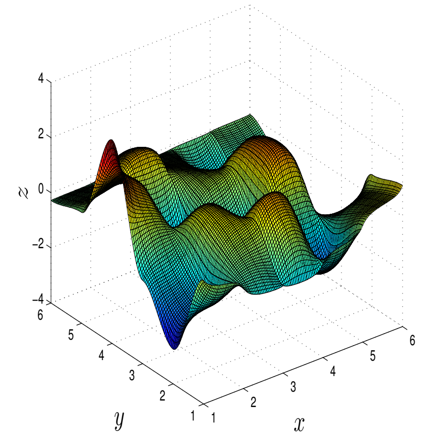
(a) FIS with arbitrary parameters.



(b) Classical rational cubic interpolation surface with arbitrary parameters.



(c) $\frac{\partial \Phi}{\partial x}$ of Figure 3.1(a).



(d) $\frac{\partial \Phi}{\partial x}$ of Figure 3.1(b).

Figure 3.1: Surface interpolation.

3.1.3 Error analysis

In this section, the uniform error bound between the rational cubic FIS Φ and the original function $S \in \mathcal{C}^4(D)$ is estimated with the help of the classical rational cubic interpolation surface C. At first, the error bound between the fractal boundary curves and the classical rational cubic splines along the grid lines, and the error bound between the classical rational cubic splines and the original function along the grid lines are established. Then with the help of error bounds that obtained along the grid lines, error bound between the rational cubic FIS Φ and the original function S is achieved.

Let us introduce the following notations: $h = \max\{h_i : i = 1, 2, \dots, M-1\}$, $h^* = \max\{h_j^* : j = 1, 2, \dots, N-1\}$, $|\boldsymbol{\alpha}|_\infty = \max\{|\boldsymbol{\alpha}_j|_\infty : j = 1, 2, \dots, N\}$, $|\boldsymbol{\alpha}_j|_\infty = \max\{|\alpha_{i,j}| : i = 1, 2, \dots, M-1\}$, $j = 1, 2, \dots, N$, $|\boldsymbol{\alpha}^*|_\infty = \max\{|\boldsymbol{\alpha}_i^*|_\infty : i = 1, 2, \dots, M\}$, $|\boldsymbol{\alpha}_i^*|_\infty = \max\{|\alpha_{i,j}^*| : j = 1,\ 2,\ \dots, N-1\}$, $i = 1, 2, \dots, M$, $\sigma_x = \max\limits_{1\le j\le N-1}(\sigma_{x,j} + \sigma_{x,j+1})$, $\sigma_{x,j} = \max\{|z_{i,j}| : i = 1, 2, \dots, M\} + h \max\{|z_{i,j}^x| : i = 1, 2, \dots, M\} + \max\{|z_{1,j}|, |z_{M,j}|\} + (x_M - x_1)\max\{|z_{1,j}^x|, |z_{M,j}^x|\}$, $\sigma_y = \max\limits_{1\le i\le M-1}(\sigma_{y,i} + \sigma_{y,i+1})$, $\sigma_{y,i} = \max\{|z_{i,j}| : j = 1, 2, \dots, N\} + h^* \max\{|z_{i,j}^y| : j = 1, 2, \dots, N\} + \max\{\ |z_{i,1}|, |z_{i,N}|\ \} + (y_N - y_1)\max\{|z_{i,1}^y|, |z_{i,N}^y|\}$.

By following the procedure given in Chapter 2, we obtain the following proposition.

Proposition 3.1.1. *Let $\{(x_i, y_j, z_{i,j}, z_{i,j}^x, z_{i,j}^y) : i = 1, 2, \dots, M,\ j = 1, 2, \dots, N\}$ be the surface data generated from the function $S \in \mathcal{C}^4(D)$. For each $j = 1, 2, \dots, N$, let B_j be the x-direction fractal boundary curve as given in (3.7) and let C_j be the classical rational cubic spline as given in (3.8). Then, for each $x \in [x_i, x_{i+1}]$, $i = 1, 2, \dots, M-1$, we have*

$$|B_j(x) - C_j(x)| \le \frac{|\boldsymbol{\alpha}_j|_\infty}{1 - |\boldsymbol{\alpha}_j|_\infty}\sigma_{x,j}.$$

The next proposition provides the error bound between classical rational cubic spline and original function.

Proposition 3.1.2. *Let $\{(x_i, y_j, z_{i,j}, z_{i,j}^x, z_{i,j}^y) : i = 1, 2, \dots, M,\ j = 1, 2, \dots, N\}$ be the surface data generated from the function $S \in \mathcal{C}^4(D)$. Let C_j be the classical rational cubic spline as in (3.8). Let S_j be the restriction of the original function along the grid* *line* $I \times y_j$. *Then for* $x \in [x_i, x_{i+1}]$, *we have*

$$|S_j(x) - C_j(x)| \leq \frac{\lambda h^2}{192\lambda^*}\left(\frac{7}{4}h^2\|S_j^{(4)}\|_\infty + 12h\|S_j^{(3)}\|_\infty + 12\|S_j^{(2)}\|_\infty\right) + \frac{\lambda h}{2\lambda^*}\max\{|S_{i,j}^x - z_{i,j}^x|,\ |S_{i+1,j}^x - z_{i+1,j}^x|\},$$

where $\lambda_{i,j} = \max\{u_{i,j},\ v_{i,j},\ \gamma_{i,j}\}$, $\lambda = \max\{\lambda_{i,j} : i = 1, 2, \ldots, M-1,\ j = 1, 2, \ldots, N\}$, $\lambda_{i,j}^* = \min\{u_{i,j},\ v_{i,j}\}$, $\lambda^* = \min\{\lambda_{i,j}^* : i = 1, 2, \ldots, M-1,\ j = 1, 2, \ldots, N\}$.

Proof. Let $x \in [x_i, x_{i+1}]$, $\varphi = \frac{x - x_i}{x_{i+1} - x_i}$. Let $S_j(x) = S(x, y_j)$ and $\mathcal{S}_{i,j}(\varphi) = S_j(x(\varphi))$ where $x(\varphi) = x_i + h_i\varphi$. We have

$$\begin{aligned} |S_j(x) - C_j(x)| &= \left|\mathcal{S}_{i,j}(\varphi) - \frac{P_{i,j}(\varphi)}{Q_{i,j}(\varphi)}\right| \\ &\leq \frac{|\mathcal{S}_{i,j}(\varphi)Q_{i,j}(\varphi) - \mathcal{P}_{i,j}(\varphi)| + |\mathcal{P}_{i,j}(\varphi) - P_{i,j}(\varphi)|}{|Q_{i,j}(\varphi)|}, \end{aligned}$$

where $\mathcal{P}_{i,j}(\varphi)$ is the two-point Hermite interpolant to $\mathcal{S}_{i,j}(\varphi)Q_{i,j}(\varphi)$. By imposing Hermite interpolation conditions with knots at $\varphi = 0$ and $\varphi = 1$, we get

$$\begin{aligned} \mathcal{P}_{i,j}(\varphi) = u_{i,j}z_{i,j}(1-\varphi)^3 + ([2u_{i,j} + v_{i,j} + \gamma_{i,j}]z_{i,j} + S_{i,j}^x h_i u_{i,j})\varphi(1-\varphi)^2 \\ + ([u_{i,j} + 2v_{i,j} + \gamma_{i,j}]z_{i+1,j} + S_{i+1,j}^x h_i v_{i,j})\varphi^2(1-\varphi) + v_{i,j}z_{i+1,j}\varphi^3. \end{aligned}$$

By the Cauchy remainder theorem, we have

$$\begin{aligned} |\mathcal{S}_{i,j}(\varphi)Q_{i,j}(\varphi) - \mathcal{P}_{i,j}(\varphi)| &\leq \frac{1}{384}\max_{0\leq\varphi\leq 1}\left|\frac{d^4}{d\varphi^4}(\mathcal{S}_{i,j}(\varphi)Q_{i,j}(\varphi))\right| \\ &= \frac{1}{384}\max_{0\leq\varphi\leq 1}|\mathcal{S}_{i,j}^{(4)}(\varphi)Q_{i,j}(\varphi) + 4\mathcal{S}_{i,j}^{(3)}(\varphi)Q_{i,j}^{(1)}(\varphi) \\ &\quad + 6\mathcal{S}_{i,j}^{(2)}(\varphi)Q_{i,j}^{(2)}(\varphi)|. \end{aligned} \tag{3.13}$$

Now, we have

$$\begin{aligned} |Q_{i,j}(\varphi)| &= Q_{i,j}(\varphi) \\ &= u_{i,j}(1-\varphi)^2 + (u_{i,j} + v_{i,j} + \gamma_{i,j})\varphi(1-\varphi) + v_{i,j}\varphi^2 \\ &\leq \lambda_{i,j}[(1-\varphi)^2 + \varphi^2] + 3\lambda_{i,j}\varphi(1-\varphi) \\ &\leq \lambda_{i,j} + \frac{3}{4}\lambda_{i,j}. \end{aligned}$$

We have $Q^{(1)}_{i,j}(\varphi) = (-u_{i,j} + v_{i,j} + \gamma_{i,j})(1-\varphi) + (-u_{i,j} + v_{i,j} - \gamma_{i,j})\varphi$ and hence we get $|Q^{(1)}_{i,j}(\varphi)| \leq 3\lambda_{i,j}$. Also, $Q^{(2)}_{i,j}(\varphi) = -2\gamma_{i,j}$ and hence $|Q^{(2)}_{i,j}(\varphi)| \leq 2\lambda_{i,j}$.

From (3.13), we can see that

$$|\mathcal{S}_{i,j}(\varphi)Q_{i,j}(\varphi) - \mathcal{P}_{i,j}(\varphi)| \leq \frac{1}{384}\Big[\frac{7}{4}h_i^4\|S_j^{(4)}\|_\infty\lambda_{i,j} + 12h_i^3\|S_j^{(3)}\|_\infty\lambda_{i,j} + 12h_i^2\|S_j^{(2)}\|_\infty\lambda_{i,j}\Big].$$

We observe that

$$\begin{aligned}|\mathcal{P}_{i,j}(\varphi) - P_{i,j}(\varphi)| &= |u_{i,j}h_i(S^x_{i,j} - z^x_{i,j})\varphi(1-\varphi)^2 - v_{i,j}h_i(S^x_{i+1,j} - z^x_{i+1,j})\varphi^2(1-\varphi)|\\ &\leq \varphi(1-\varphi)h_i\Big[u_{i,j}|S^x_{i,j} - z^x_{i,j}|(1-\varphi) + v_{i,j}|S^x_{i+1,j} - z^x_{i+1,j}|\varphi\Big]\\ &\leq \frac{\lambda_{i,j}h_i}{4}\max\{|S^x_{i,j} - z^x_{i,j}|,\ |S^x_{i+1,j} - z^x_{i+1,j}|\}.\end{aligned}$$

Further, we have

$$\begin{aligned}Q_{i,j}(\varphi) &= u_{i,j}(1-\varphi)^2 + (u_{i,j} + v_{i,j} + \gamma_{i,j})\varphi(1-\varphi) + v_{i,j}\varphi^2\\ &\geq u_{i,j}(1-\varphi)^2 + v_{i,j}\varphi^2\\ &\geq \frac{\lambda^*_{i,j}}{2} > 0.\end{aligned}$$

Hence, we get

$$\begin{aligned}|S_j(x) - C_j(x)| \leq \frac{\lambda_{i,j}h_i^2}{192\lambda^*_{i,j}}&\left(\frac{7}{4}h_i^2\|S_j^{(4)}\|_\infty + 12h_i\|S_j^{(3)}\|_\infty + 12\|S_j^{(2)}\|_\infty\right)\\ &+ \frac{\lambda_{i,j}h_i}{2\lambda^*_{i,j}}\max\{|S^x_{i,j} - z^x_{i,j}|,\ |S^x_{i+1,j} - z^x_{i+1,j}|\}.\end{aligned}$$

□

Theorem 3.1.1. *Let the surface data* $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *be generated from an original function* $S \in \mathcal{C}^4(D)$. *Let* Φ *and* C *be the corresponding rational cubic FIS and the classical rational cubic interpolation surface respectively. Then*

$$\begin{aligned}\|S - \Phi\|_\infty \leq \frac{|\boldsymbol{\alpha}|_\infty}{1 - |\boldsymbol{\alpha}|_\infty}\sigma_x + \frac{|\boldsymbol{\alpha}^*|_\infty}{1 - |\boldsymbol{\alpha}^*|_\infty}\sigma_y &+ h^*\left(\left\|\frac{\partial S}{\partial y}\right\|_\infty + \left\|\frac{\partial C}{\partial y}\right\|_\infty\right)\\ &+ \frac{\lambda h^2}{192\lambda^*}\left(\frac{7}{4}h^2\left\|\frac{\partial^4 S}{\partial x^4}\right\|_\infty + 12h\left\|\frac{\partial^3 S}{\partial x^3}\right\|_\infty + 12\left\|\frac{\partial^2 S}{\partial x^2}\right\|_\infty\right)\\ &+ \frac{\lambda h}{2\lambda^*}\max\{|S^x_{i,j} - z^x_{i,j}|,\ |S^x_{i+1,j} - z^x_{i+1,j}|\}.\end{aligned}$$

Proof. From (3.11) and Remark 3.1.3, we have

$$|\Phi(x,y) - C(x,y)| \le b_{y,0}(\zeta)|B_j(x) - C_j(x)| + b_{y,1}(\zeta)|B_{j+1}(x) - C_{j+1}(x)| + a_{x,0}(\theta)|B_i^*(y) - C_i^*(y)| + a_{x,1}(\theta)|B_{i+1}^*(y) - C_{i+1}^*(y)|. \tag{3.14}$$

From Proposition 3.1.1, we obtain

$$\begin{cases} |B_j(x) - C_j(x)| \le \dfrac{|\boldsymbol{\alpha}_j|_\infty}{1 - |\boldsymbol{\alpha}_j|_\infty}\sigma_{x,j}, \ j = 1, 2, \ldots, N, \\ |B_i^*(y) - C_i^*(y)| \le \dfrac{|\boldsymbol{\alpha}_i^*|_\infty}{1 - |\boldsymbol{\alpha}_i^*|_\infty}\sigma_{y,i}, \ i = 1, 2, \ldots, M. \end{cases} \tag{3.15}$$

Using (3.15) in (3.14), we get

$$\begin{aligned} |\Phi(x,y) - C(x,y)| &\le \frac{|\boldsymbol{\alpha}_j|_\infty}{1 - |\boldsymbol{\alpha}_j|_\infty}\sigma_{x,j} + \frac{|\boldsymbol{\alpha}_{j+1}|_\infty}{1 - |\boldsymbol{\alpha}_{j+1}|_\infty}\sigma_{x,j+1} \\ &\quad + \frac{|\boldsymbol{\alpha}_i^*|_\infty}{1 - |\boldsymbol{\alpha}_i^*|_\infty}\sigma_{y,i} + \frac{|\boldsymbol{\alpha}_{i+1}^*|_\infty}{1 - |\boldsymbol{\alpha}_{i+1}^*|_\infty}\sigma_{y,i+1} \\ &\le \frac{|\boldsymbol{\alpha}|_\infty}{1 - |\boldsymbol{\alpha}|_\infty}\sigma_x + \frac{|\boldsymbol{\alpha}^*|_\infty}{1 - |\boldsymbol{\alpha}^*|_\infty}\sigma_y. \end{aligned}$$

Since the above inequality is true for all $(x,y) \in D_{i,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N-1$, we have

$$\|\Phi - C\|_\infty \le \frac{|\boldsymbol{\alpha}|_\infty}{1 - |\boldsymbol{\alpha}|_\infty}\sigma_x + \frac{|\boldsymbol{\alpha}^*|_\infty}{1 - |\boldsymbol{\alpha}^*|_\infty}\sigma_y. \tag{3.16}$$

Expanding the original function $S \in \mathcal{C}^4(D)$ in Taylor series about the point $(x, y_j) \in D_{i,j}$, we obtain

$$S(x,y) = S(x,y_j) + (y - y_j)\frac{\partial S(x,\xi)}{\partial y}, \quad (x,\xi) \in D_{i,j}.$$

We have

$$|S(x,y) - S(x,y_j)| \le h^* \left\| \frac{\partial S}{\partial y} \right\|_\infty. \tag{3.17}$$

Similarly, expanding the classical rational cubic interpolation surface $C \in \mathcal{C}^1(D)$ in Taylor series about the point $(x, y_j) \in D_{i,j}$ gives

$$C(x,y) = C(x,y_j) + (y - y_j)\frac{\partial C(x,\eta)}{\partial y}, \quad (x,\eta) \in D_{i,j}.$$

Thus, we have

$$|C(x,y) - C(x,y_j)| \le h^* \left\| \frac{\partial C}{\partial y} \right\|_\infty. \tag{3.18}$$

From (3.17) and (3.18), we get

$$\begin{aligned}|S(x,y)-C(x,y)| &\le |S(x,y)-S(x,y_j)|+|S(x,y_j)-C(x,y_j)|\\ &\quad+|C(x,y_j)-C(x,y)|\\ &\le h^*\left(\left\|\frac{\partial S}{\partial y}\right\|_\infty+\left\|\frac{\partial C}{\partial y}\right\|_\infty\right)+|S(x,y_j)-C(x,y_j)|. \qquad (3.19)\end{aligned}$$

From (3.19) and Proposition 3.1.2, we get

$$\begin{aligned}|S(x,y)-C(x,y)| \le h^*\left(\left\|\frac{\partial S}{\partial y}\right\|_\infty+\left\|\frac{\partial C}{\partial y}\right\|_\infty\right)&+\frac{\lambda h^2}{192\lambda^*}\left(\frac{7}{4}h^2\left\|\frac{\partial^4 S}{\partial x^4}\right\|_\infty\right.\\ &\left.+12h\left\|\frac{\partial^3 S}{\partial x^3}\right\|_\infty+12\left\|\frac{\partial^2 S}{\partial x^2}\right\|_\infty\right)\\ &+\frac{\lambda h}{2\lambda^*}\max\{|S^x_{i,j}-z^x_{i,j}|,\ |S^x_{i+1,j}-z^x_{i+1,j}|\}.\end{aligned}$$

Since this inequality is true for all $(x,y)\in D_{i,j}$, $i=1,2,\ldots,M-1$, $j=1,2,\ldots,N-1$, we have

$$\begin{aligned}\|S-C\|_\infty \le h^*&\left(\left\|\frac{\partial S}{\partial y}\right\|_\infty+\left\|\frac{\partial C}{\partial y}\right\|_\infty\right)\\ &+\frac{\lambda h^2}{192\lambda^*}\left(\frac{7}{4}h^2\left\|\frac{\partial^4 S}{\partial x^4}\right\|_\infty+12h\left\|\frac{\partial^3 S}{\partial x^3}\right\|_\infty+12\left\|\frac{\partial^2 S}{\partial x^2}\right\|_\infty\right)\\ &+\frac{\lambda h}{2\lambda^*}\max\{|S^x_{i,j}-z^x_{i,j}|,\ |S^x_{i+1,j}-z^x_{i+1,j}|\}. \qquad (3.20)\end{aligned}$$

Using (3.16), (3.20) and the inequality $\|S-\Phi\|_\infty \le \|S-C\|_\infty+\|C-\Phi\|_\infty$, we obtain the required error bound for $\|S-\Phi\|_\infty$. □

3.2 Shape Preserving FIS

In this section, conditions on the scaling factors and the shape parameters are developed for Φ to preserve the positivity, monotonicity and convexity. Also, the scaling factors and the shape parameters are restricted for rational cubic FIS to lie above the plane whenever the surface data lies above the plane.

3.2.1 Positivity preserving interpolation

For $j = 1, 2, \ldots, N$, the fractal boundary curve $B_j(x)$ is positive over $[x_1, x_M]$ if $B_j(x) > 0$, for all $x \in [x_1, x_M]$. The following theorem provides the conditions for fractal boundary curves along x-direction $B_j(x)$, $j = 1, 2, \ldots, N$ to be positive.

Theorem 3.2.1. *Let $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M, j = 1, 2, \ldots, N\}$ be a positive surface data. For each $j = 1, 2, \ldots, N$, let $B_j(x)$ be the fractal boundary curve along $I \times y_j$. Then the conditions on the scaling factors and the shape parameters for $B_j(x)$ to satisfy the positivity are*

$$\alpha^1_{i,j} < \alpha_{i,j} < \alpha^2_{i,j},$$

$$u_{i,j} > 0,\ v_{i,j} > 0 \text{ and } \gamma_{i,j} > \max\{0, a^1_{i,j}, a^2_{i,j}, a^3_{i,j}, a^4_{i,j}, a^5_{i,j}, a^6_{i,j}, a^7_{i,j}, a^8_{i,j}\},$$

where

$$\alpha^1_{i,j} = \max\left\{ -a_i, \frac{z_{i,j} - k_{1,j}}{z_{1,j} - k_{2,j}}, \frac{z_{i+1,j} - k_{1,j}}{z_{M,j} - k_{2,j}}, \frac{k_{2,j} - z_{i,j}}{k_{1,j} - z_{1,j}}, \frac{k_{2,j} - z_{i+1,j}}{k_{1,j} - z_{M,j}} \right\},$$

$$\alpha^2_{i,j} = \min\left\{ a_i, \frac{z_{i,j} - k_{1,j}}{z_{1,j} - k_{1,j}}, \frac{z_{i+1,j} - k_{1,j}}{z_{M,j} - k_{1,j}}, \frac{k_{2,j} - z_{i,j}}{k_{2,j} - z_{1,j}}, \frac{k_{2,j} - z_{i+1,j}}{k_{2,j} - z_{M,j}} \right\},$$

$$a^1_{i,j} = \frac{-u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1) z^x_{1,j}]}{(z_{i,j} - k_{1,j}) - \alpha_{i,j}(z_{1,j} - k_{1,j})}, \quad a^2_{i,j} = \frac{v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1) z^x_{M,j}]}{(z_{i+1,j} - k_{1,j}) - \alpha_{i,j}(z_{M,j} - k_{1,j})},$$

$$a^3_{i,j} = \frac{u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1) z^x_{1,j}]}{(k_{2,j} - z_{i,j}) - \alpha_{i,j}(k_{2,j} - z_{1,j})}, \quad a^4_{i,j} = \frac{-v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1) z^x_{M,j}]}{(k_{2,j} - z_{i+1,j}) - \alpha_{i,j}(k_{2,j} - z_{M,j})},$$

$$a^5_{i,j} = \frac{-u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1) z^x_{1,j}]}{(z_{i,j} - k_{1,j}) - \alpha_{i,j}(z_{1,j} - k_{2,j})}, \quad a^6_{i,j} = \frac{v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1) z^x_{M,j}]}{(z_{i+1,j} - k_{1,j}) - \alpha_{i,j}(z_{M,j} - k_{2,j})},$$

$$a^7_{i,j} = \frac{u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1) z^x_{1,j}]}{(k_{2,j} - z_{i,j}) - \alpha_{i,j}(k_{1,j} - z_{1,j})}, \quad a^8_{i,j} = \frac{-v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1) z^x_{M,j}]}{(k_{2,j} - z_{i+1,j}) - \alpha_{i,j}(k_{1,j} - z_{M,j})},$$

$$0 < k_{1,j} < \min\{z_{i,j} : i = 1, 2, \ldots, M\},\ k_{2,j} > \max\{z_{i,j} : i = 1, 2, \ldots, M\},$$

for $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$.

Proof. Since the graph of $B_j(x)$ is the attractor of the IFS $\mathcal{J}_j$, $j = 1, 2, \ldots, N$, it is clear that $B_j(x) > 0$ for all $x \in [x_1, x_M]$, if $F_{i,j}(x, z) \in [k_{1,j}, k_{2,j}]$ for all $(x, z) \in [x_1, x_M] \times [k_{1,j}, k_{2,j}]$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$.

Next, the conditions on the scaling factors $\alpha_{i,j}$ and the shape parameters $u_{i,j}$, $v_{i,j}$ and $\gamma_{i,j}$ are derived such that range of $F_{i,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$ lies in

$[k_{1,j}, k_{2,j}]$, $j = 1, 2, \ldots, N$.

Case I: Let $0 \le \alpha_{i,j} < a_i$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$. Let $(x, z) \in [x_1, x_M] \times [k_{1,j}, k_{2,j}]$. By the assumption on the scaling factors, it is clear that $k_{1,j}\alpha_{i,j} \le z\alpha_{i,j} \le k_{2,j}\alpha_{i,j}$ and hence we have

$$k_{1,j}\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)} \le z\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)} \le k_{2,j}\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)}.$$

To prove $F_{i,j}(x, z) \in [k_{1,j}, k_{2,j}]$, it is enough to prove that

$$k_{1,j} \le k_{1,j}\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)}, \tag{3.21}$$

$$k_{2,j}\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)} \le k_{2,j}, \tag{3.22}$$

for $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$. Now (3.21) can be written as

$$k_{1,j}[\alpha_{i,j} - 1] + \frac{A_{1,i,j}(1-\theta)^3 + A_{2,i,j}\theta(1-\theta)^2 + A_{3,i,j}\theta^2(1-\theta) + A_{4,i,j}\theta^3}{u_{i,j}(1-\theta)^2 + (u_{i,j} + v_{i,j} + \gamma_{i,j})\theta(1-\theta) + v_{i,j}\theta^2} \ge 0, \tag{3.23}$$

where the coefficients $A_{1,i,j}$, $A_{2,i,j}$, $A_{3,i,j}$ and $A_{4,i,j}$ are as in (3.3)-(3.6). The denominator $u_{i,j}(1-\theta)^2 + (u_{i,j} + v_{i,j} + \gamma_{i,j})\theta(1-\theta) + v_{i,j}\theta^2$ can be written as $u_{i,j}(1-\theta)^3 + (2u_{i,j} + v_{i,j} + \gamma_{i,j})\theta(1-\theta)^2 + (u_{i,j} + 2v_{i,j} + \gamma_{i,j})\theta^2(1-\theta) + v_{i,j}\theta^3$. Using this degree elevated form in (3.23), cross multiplying and rearranging, it is observed that (3.23) is true, if

$$\begin{aligned}&[A_{1,i,j} + u_{i,j}k_{1,j}(\alpha_{i,j} - 1)](1-\theta)^3 + [A_{2,i,j} + (2u_{i,j} + v_{i,j} + \gamma_{i,j})k_{1,j}(\alpha_{i,j} - 1)]\theta(1-\theta)^2\\ &+[A_{3,i,j} + (u_{i,j} + 2v_{i,j} + \gamma_{i,j})k_{1,j}(\alpha_{i,j} - 1)]\theta^2(1-\theta) + [A_{4,i,j} + v_{i,j}k_{1,j}(\alpha_{i,j} - 1)]\theta^3 \ge 0.\end{aligned} \tag{3.24}$$

Now (3.24) is true if

$$A_{1,i,j} + u_{i,j}k_{1,j}(\alpha_{i,j} - 1) \ge 0, \quad A_{2,i,j} + (2u_{i,j} + v_{i,j} + \gamma_{i,j})k_{1,j}(\alpha_{i,j} - 1) \ge 0,$$

$$A_{3,i,j} + (u_{i,j} + 2v_{i,j} + \gamma_{i,j})k_{1,j}(\alpha_{i,j} - 1) \ge 0, \quad A_{4,i,j} + v_{i,j}k_{1,j}(\alpha_{i,j} - 1) \ge 0.$$

Since $u_{i,j} > 0$ and $v_{i,j} > 0$, we have

$$A_{1,i,j} + u_{i,j}k_{1,j}(\alpha_{i,j} - 1) \ge 0, \text{ if } \alpha_{i,j} \le \frac{z_{i,j} - k_{1,j}}{z_{1,j} - k_{1,j}},$$

$$A_{4,i,j} + u_{i,j}k_{1,j}(\alpha_{i,j} - 1) \ge 0, \text{ if } \alpha_{i,j} \le \frac{z_{i+1,j} - k_{1,j}}{z_{M,j} - k_{1,j}}.$$

Let $0 \leq \alpha_{i,j} < \min\left\{\frac{z_{i,j}-k_{1,j}}{z_{1,j}-k_{1,j}}, \frac{z_{i+1,j}-k_{1,j}}{z_{M,j}-k_{1,j}}\right\}$. It can be seen that

$$\begin{aligned}
&A_{2,i,j} + (2u_{i,j} + v_{i,j} + \gamma_{i,j})k_{1,j}(\alpha_{i,j} - 1) \geq 0 \Leftrightarrow \\
&(2u_{i,j} + v_{i,j} + \gamma_{i,j})[(z_{i,j} - k_{1,j}) - \alpha_{i,j}(z_{1,j} - k_{1,j})] + u_{i,j}h_i z^x_{i,j} - \alpha_{i,j}u_{i,j}(x_M - x_1)z^x_{1,j} \geq 0 \\
&\Leftrightarrow \gamma_{i,j}[(z_{i,j} - k_{1,j}) - \alpha_{i,j}(z_{1,j} - k_{1,j})] + u_{i,j}h_i z^x_{i,j} - \alpha_{i,j}u_{i,j}(x_M - x_1)z^x_{1,j} \geq 0.
\end{aligned}$$

Hence, we have

$\gamma_{i,j}[(z_{i,j} - k_{1,j}) - \alpha_{i,j}(z_{1,j} - k_{1,j})] + u_{i,j}h_i z^x_{i,j} - \alpha_{i,j}u_{i,j}(x_M - x_1)z^x_{1,j} \geq 0$ if

$$\gamma_{i,j} \geq \frac{-u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1)z^x_{1,j}]}{(z_{i,j} - k_{1,j}) - \alpha_{i,j}(z_{1,j} - k_{1,j})}.$$

Similarly, we get

$A_{3,i,j} + (u_{i,j} + 2v_{i,j} + \gamma_{i,j})k_{1,j}(\alpha_{i,j} - 1) \geq 0$ if $\gamma_{i,j} \geq \frac{v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1)z^x_{M,j}]}{(z_{i+1,j} - k_{1,j}) - \alpha_{i,j}(z_{M,j} - k_{1,j})}$.

Hence the conditions to satisfy (3.24) are

$$0 \leq \alpha_{i,j} < \min\left\{\frac{z_{i,j} - k_{1,j}}{z_{1,j} - k_{1,j}}, \frac{z_{i+1,j} - k_{1,j}}{z_{M,j} - k_{1,j}}\right\},$$

$$\gamma_{i,j} \geq \max\left\{\frac{-u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1)z^x_{1,j}]}{(z_{i,j} - k_{1,j}) - \alpha_{i,j}(z_{1,j} - k_{1,j})}, \frac{v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1)z^x_{M,j}]}{(z_{i+1,j} - k_{1,j}) - \alpha_{i,j}(z_{M,j} - k_{1,j})}\right\}.$$

Similarly, the conditions which satisfy (3.22) are

$$0 \leq \alpha_{i,j} < \min\left\{\frac{k_{2,j} - z_{i,j}}{k_{2,j} - z_{1,j}}, \frac{k_{2,j} - z_{i+1,j}}{k_{2,j} - z_{M,j}}\right\},$$

$$\gamma_{i,j} \geq \max\left\{\frac{u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1)z^x_{1,j}]}{(k_{2,j} - z_{i,j}) - \alpha_{i,j}(k_{2,j} - z_{1,j})}, \frac{-v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1)z^x_{M,j}]}{(k_{2,j} - z_{i+1,j}) - \alpha_{i,j}(k_{2,j} - z_{M,j})}\right\}.$$

Thus, when $0 \leq \alpha_{i,j} < a_i$, the conditions for $F_{i,j}(x, z)$ to lie in $[k_{1,j}, k_{2,j}]$ are

$$0 \leq \alpha_{i,j} < \min\left\{a_i, \frac{z_{i,j} - k_{1,j}}{z_{1,j} - k_{1,j}}, \frac{z_{i+1,j} - k_{1,j}}{z_{M,j} - k_{1,j}}, \frac{k_{2,j} - z_{i,j}}{k_{2,j} - z_{1,j}}, \frac{k_{2,j} - z_{i+1,j}}{k_{2,j} - z_{M,j}}\right\},$$

$$u_{i,j} > 0,\ v_{i,j} > 0 \text{ and } \gamma_{i,j} \geq \max\{a^1_{i,j}, a^2_{i,j}, a^3_{i,j}, a^4_{i,j}\}.$$

Case II: Let $-a_i < \alpha_{i,j} \leq 0$, $i = 1, 2, \ldots, M - 1$, $j = 1, 2, \ldots, N$. Let $(x, z) \in [x_1, x_M] \times [k_{1,j}, k_{2,j}]$. By the assumption on the scaling factors it is clear that $k_{2,j}\alpha_{i,j} \leq z\alpha_{i,j} \leq k_{1,j}\alpha_{i,j}$ and hence, we get

$$k_{2,j}\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)} \leq z\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)} \leq k_{1,j}\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)}.$$

To prove $F_{i,j}(x,z) \in [k_{1,j}, k_{2,j}]$, it is adequate to prove that

$$k_{1,j} \le k_{2,j}\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)}, \quad k_{1,j}\alpha_{i,j} + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)} \le k_{2,j}. \tag{3.25}$$

Using similar computations as done in the case of $0 \le \alpha_{i,j} < a_i$, it is observed that the conditions to satisfy (3.25) are

$$0 \ge \alpha_{i,j} > \max\Big\{ -a_i, \frac{z_{i,j}-k_{1,j}}{z_{1,j}-k_{2,j}}, \frac{z_{i+1,j}-k_{1,j}}{z_{M,j}-k_{2,j}}, \frac{k_{2,j}-z_{i,j}}{k_{1,j}-z_{1,j}}, \frac{k_{2,j}-z_{i+1,j}}{k_{1,j}-z_{M,j}}\Big\},$$

$$u_{i,j} > 0,\ v_{i,j} > 0 \text{ and } \gamma_{i,j} \ge \max\{a^5_{i,j}, a^6_{i,j}, a^7_{i,j}, a^8_{i,j}\}.$$

Thus, by combining all the results, we get the conditions prescribed in the statement of the Theorem 3.2.1. □

Applying the same procedure as used in Theorem 3.2.1, the conditions for $B^*_i(y)$, $i = 1, 2, \ldots, M$ to be positive are derived and presented in the following theorem.

Theorem 3.2.2. *Let $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M, j = 1, 2, \ldots, N\}$ be a positive surface data. For each $i = 1, 2, \ldots, M$, let $B^*_i(y)$ be the fractal boundary curve along $x_i \times J$. Then the conditions on the scaling factors and the shape parameters for $B^*_i(y)$ to satisfy the positivity are*

$$\alpha^3_{i,j} < \alpha^*_{i,j} < \alpha^4_{i,j},$$

$$u^*_{i,j} > 0,\ v^*_{i,j} > 0 \text{ and } \gamma^*_{i,j} > \max\{0, b^1_{i,j}, b^2_{i,j}, b^3_{i,j}, b^4_{i,j}, b^5_{i,j}, b^6_{i,j}, b^7_{i,j}, b^8_{i,j}\},$$

where

$$\alpha^3_{i,j} = \max\Big\{ -c_j, \frac{z_{i,j}-k^*_{1,i}}{z_{i,1}-k^*_{2,i}}, \frac{z_{i,j+1}-k^*_{1,i}}{z_{i,N}-k^*_{2,i}}, \frac{k^*_{2,i}-z_{i,j}}{k^*_{1,i}-z_{i,1}}, \frac{k^*_{2,i}-z_{i,j+1}}{k^*_{1,i}-z_{i,N}}\Big\},$$

$$\alpha^4_{i,j} = \min\Big\{ c_j, \frac{z_{i,j}-k^*_{1,i}}{z_{i,1}-k^*_{1,i}}, \frac{z_{i,j+1}-k^*_{1,i}}{z_{i,N}-k^*_{1,i}}, \frac{k^*_{2,i}-z_{i,j}}{k^*_{2,i}-z_{i,1}}, \frac{k^*_{2,i}-z_{i,j+1}}{k^*_{2,i}-z_{i,N}}\Big\},$$

$$b^1_{i,j} = \frac{-u^*_{i,j}[h^*_j z^y_{i,j} - \alpha^*_{i,j}(y_N-y_1)z^y_{i,1}]}{(z_{i,j}-k^*_{1,i}) - \alpha^*_{i,j}(z_{i,1}-k^*_{1,i})}, \quad b^2_{i,j} = \frac{v^*_{i,j}[h^*_j z^y_{i,j+1} - \alpha^*_{i,j}(y_N-y_1)z^y_{i,N}]}{(z_{i,j+1}-k^*_{1,i}) - \alpha^*_{i,j}(z_{i,N}-k^*_{1,i})},$$

$$b^3_{i,j} = \frac{u^*_{i,j}[h^*_j z^y_{i,j} - \alpha^*_{i,j}(y_N-y_1)z^y_{i,1}]}{(k^*_{2,i}-z_{i,j}) - \alpha^*_{i,j}(k^*_{2,i}-z_{i,1})}, \quad b^4_{i,j} = \frac{-v^*_{i,j}[h^*_j z^y_{i,j+1} - \alpha^*_{i,j}(y_N-y_1)z^y_{i,N}]}{(k^*_{2,i}-z_{i,j+1}) - \alpha^*_{i,j}(k^*_{2,i}-z_{i,N})},$$

$$b^5_{i,j} = \frac{-u^*_{i,j}[h^*_j z^y_{i,j} - \alpha^*_{i,j}(y_N-y_1)z^y_{i,1}]}{(z_{i,j}-k^*_{1,i}) - \alpha^*_{i,j}(z_{i,1}-k^*_{2,i})}, \quad b^6_{i,j} = \frac{v^*_{i,j}[h^*_j z^y_{i,j+1} - \alpha^*_{i,j}(y_N-y_1)z^y_{i,N}]}{(z_{i,j+1}-k^*_{1,i}) - \alpha^*_{i,j}(z_{i,N}-k^*_{2,i})},$$

$$b_{i,j}^7 = \frac{u_{i,j}^*[h_j^* z_{i,j}^y - \alpha_{i,j}^*(y_N - y_1) z_{i,1}^y]}{(k_{2,i}^* - z_{i,j}) - \alpha_{i,j}^*(k_{1,i}^* - z_{i,1})}, \quad b_{i,j}^8 = \frac{-v_{i,j}^*[h_j^* z_{i,j+1}^y - \alpha_{i,j}^*(y_N - y_1) z_{i,N}^y]}{(k_{2,i}^* - z_{i,j+1}) - \alpha_{i,j}^*(k_{1,i}^* - z_{i,N})},$$
$$0 < k_{1,i}^* < \min\{z_{i,j} : j = 1, 2, \ldots, N\}, \ k_{2,i}^* > \max\{z_{i,j} : j = 1, 2, \ldots, N\}.$$

$i = 1, 2, \ldots, M, \ j = 1, 2, \ldots, N-1$.

In the following theorem, the conditions for FIS to be positive is presented.

Theorem 3.2.3. *Let $\{(x_i, y_j, z_{i,j}, z_{i,j}^x, z_{i,j}^y) : i = 1, 2, \ldots, M, j = 1, 2, \ldots, N\}$ be a positive surface data. Then the rational cubic FIS Φ preserves the positivity if the scaling factors and the shape parameters satisfy the hypotheses of Theorems 3.2.1 and 3.2.2.*

Proof. The rational cubic FIS Φ (3.10) is a combination of the fractal boundary curves $B_j(x)$, $j = 1, 2, \ldots, N$ and $B_i^*(y)$, $i = 1, 2, \ldots, M$. From [28], it is known that the FIS Φ inherits all the natures of the fractal boundary curves. Fractal boundary curves $B_j(x)$ for $j = 1, 2, \ldots, N$ and $B_i^*(y)$ for $i = 1, 2, \ldots, M$ are positive whenever the scaling factors and the shape parameters satisfy the conditions given in Theorems 3.2.1 and 3.2.2 respectively. Hence the rational cubic FIS Φ is positive whenever the scaling factors and the shape parameters satisfy the conditions given in Theorems 3.2.1 and 3.2.2. □

3.2.2 Constrained interpolation

Let $\Pi := \Pi(x, y) = c\left[1 - \frac{x}{a} - \frac{y}{b}\right]$ be a given plane. Let the surface data $\{(x_i,\ y_j,\ z_{i,j},\ z_{i,j}^x,\ z_{i,j}^y) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ lies above the plane Π, i.e., $z_{i,j} > \Pi(x_i, y_j) = c\left[1 - \frac{x_i}{a} - \frac{y_j}{b}\right] := p_{i,j}$ (say), $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$. Then the conditions for Φ to lie above the plane Π are derived in this subsection. In the following theorem, the conditions for x-direction fractal boundary curves to lie above the plane Π are derived.

Theorem 3.2.4. *Let $\{(x_i, y_j, z_{i,j}, z_{i,j}^x, z_{i,j}^y) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ be the surface data lies above the plane Π. For each fixed $j = 1, 2, \ldots, N$, $B_j(x)$ is the fractal boundary curve along the grid $I \times y_j$. Then the fractal boundary curve $B_j(x)$ lies above the plane Π if the scaling factors and the shape parameters satisfy the following conditions:*

$$0 \le \alpha_{i,j} < \min\left\{a_i, \frac{z_{i,j} - p_{i,j}}{z_{1,j} - p_{1,j}}, \frac{z_{i+1,j} - p_{i+1,j}}{z_{M,j} - p_{M,j}}\right\},$$

$$u_{i,j} > 0,\ v_{i,j} > 0 \textit{ and } \gamma_{i,j} > \max\{0,\ c^1_{i,j},\ c^2_{i,j}\},$$

where

$$c^1_{i,j} = \frac{-u_{i,j}[(z_{i,j} - p_{i+1,j}) - \alpha_{i,j}(z_{1,j} - p_{M,j})] - u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1)z^x_{1,j}]}{[(z_{i,j} - p_{i,j}) - \alpha_{i,j}(z_{1,j} - p_{1,j})]},$$
$$c^2_{i,j} = \frac{-v_{i,j}[(z_{i+1,j} - p_{i,j}) - \alpha_{i,j}(z_{M,j} - p_{1,j})] + v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1)z^x_{M,j}]}{[(z_{i+1,j} - p_{i+1,j}) - \alpha_{i,j}(z_{M,j} - p_{M,j})]},$$

for $i = 1, 2, \ldots, M-1$.

Proof. For each fixed $j = 1, 2, \ldots, N$, the fractal boundary curves $B_j(x)$ lies above the plane Π if $B_j(x) > (1-\theta)p_{1,j} + \theta p_{M,j}$ for all $x \in [x_1, x_M]$, where $\theta = (x - x_1)/(x_M - x_1)$. Since the graph of $B_j(x)$ is the attractor of the IFS $\mathcal{J}_j$, to prove $B_j(x) > (1-\theta)p_{1,j} + \theta p_{M,j}$, it is enough to prove $(1-\theta)p_{i,j} + \theta p_{i+1,j} < F_{i,j}(x,z)$ for all (x,z) with $x \in [x_1, x_M]$ and $(1-\theta)p_{1,j} + \theta p_{M,j} < z$, $i = 1, 2, \ldots, M-1$.

For each fixed $j = 1, 2, \ldots, N$, let $0 \le \alpha_{i,j} < a_i$ and (x,z) such that $x \in [x_1, x_M]$, $(1-\theta)p_{1,j} + \theta p_{M,j} < z$, $i = 1, 2, \ldots, M-1$. We have

$$\alpha_{i,j}[(1-\theta)p_{1,j} + \theta p_{M,j}] + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)} < \alpha_{i,j} z + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)}.$$

Hence, to prove $(1-\theta)p_{i,j} + \theta p_{i+1,j} < F_{i,j}(x,z)$, it is enough to prove the following inequality:

$$(1-\theta)p_{i,j} + \theta p_{i+1,j} < \alpha_{i,j}[(1-\theta)p_{1,j} + \theta p_{M,j}] + \frac{p_{i,j}(\theta)}{q_{i,j}(\theta)}. \tag{3.26}$$

Now (3.26) can be written as

$$\frac{q_{i,j}(\theta)[(\alpha_{i,j}p_{1,j} - p_{i,j})(1-\theta) + (\alpha_{i,j}p_{M,j} - p_{i+1,j})\theta] + p_{i,j}(\theta)}{q_{i,j}(\theta)} > 0. \tag{3.27}$$

Since $q_{i,j}(\theta) > 0$, to prove (3.27), it is enough to prove that

$$q_{i,j}(\theta)[(\alpha_{i,j}p_{1,j} - p_{i,j})(1-\theta) + (\alpha_{i,j}p_{M,j} - p_{i+1,j})\theta] + p_{i,j}(\theta) > 0.$$

Now $q_{i,j}(\theta)[(\alpha_{i,j}p_{1,j} - p_{i,j})(1-\theta) + (\alpha_{i,j}p_{M,j} - p_{i+1,j})\theta] + p_{i,j}(\theta) = B_{1,i,j}(1-\theta)^3 + B_{2,i,j}\theta(1-\theta)^2 + B_{3,i,j}\theta^2(1-\theta) + B_{4,i,j}\theta^3$, where

$$B_{1,i,j} = u_{i,j}[(z_{i,j} - p_{i,j}) - \alpha_{i,j}(z_{1,j} - p_{1,j})],$$

$$
\begin{aligned}
B_{2,i,j} &= (u_{i,j} + v_{i,j} + \gamma_{i,j})[(z_{i,j} - p_{i,j}) - \alpha_{i,j}(z_{1,j} - p_{1,j})] + u_{i,j}[(z_{i,j} - p_{i+1,j}) \\
&\quad - \alpha_{i,j}(z_{1,j} - p_{M,j})] + u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1) z^x_{1,j}], \\
B_{3,i,j} &= (u_{i,j} + v_{i,j} + \gamma_{i,j})[(z_{i+1,j} - p_{i+1,j}) - \alpha_{i,j}(z_{M,j} - p_{M,j})] \\
&\quad + v_{i,j}[(z_{i+1,j} - p_{i,j}) - \alpha_{i,j}(z_{M,j} - p_{1,j})] \\
&\quad - v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1) z^x_{M,j}], \\
B_{4,i,j} &= v_{i,j}[(z_{i+1,j} - p_{i+1,j}) - \alpha_{i,j}(z_{M,j} - p_{M,j})].
\end{aligned}
$$

It can be seen that $B_{1,i,j}(1-\theta)^3 + B_{2,i,j}\theta(1-\theta)^2 + B_{3,i,j}\theta^2(1-\theta) + B_{4,i,j}\theta^3 > 0$ if the coefficients $B_{k,i,j} > 0$, $k = 1, 2, 3, 4$. We get

$$
B_{1,i,j} > 0 \text{ if } \alpha_{i,j} < \frac{z_{i,j} - p_{i,j}}{z_{1,j} - p_{1,j}} \quad \text{and} \quad B_{4,i,j} > 0 \text{ if } \alpha_{i,j} < \frac{z_{i+1,j} - p_{i+1,j}}{z_{M,j} - p_{M,j}}.
$$

Let $0 \le \alpha_{i,j} < \min\left\{\frac{z_{i,j}-p_{i,j}}{z_{1,j}-p_{1,j}}, \frac{z_{i+1,j}-p_{i+1,j}}{z_{M,j}-p_{M,j}}\right\}$. The coefficient $B_{2,i,j} > 0$ if

$$
\gamma_{i,j} > \frac{-u_{i,j}[(z_{i,j} - p_{i+1,j}) - \alpha_{i,j}(z_{1,j} - p_{M,j})] - u_{i,j}[h_i z^x_{i,j} - \alpha_{i,j}(x_M - x_1) z^x_{1,j}]}{[(z_{i,j} - p_{i,j}) - \alpha_{i,j}(z_{1,j} - p_{1,j})]}.
$$

The coefficient $B_{3,i,j} > 0$ if

$$
\gamma_{i,j} > \frac{-v_{i,j}[(z_{i+1,j} - p_{i,j}) - \alpha_{i,j}(z_{M,j} - p_{1,j})] + v_{i,j}[h_i z^x_{i+1,j} - \alpha_{i,j}(x_M - x_1) z^x_{M,j}]}{[(z_{i+1,j} - p_{i+1,j}) - \alpha_{i,j}(z_{M,j} - p_{M,j})]}.
$$

Hence, the conditions on the scaling factors and the shape parameters to satisfy (3.26) are

$$
\begin{gathered}
0 \le \alpha_{i,j} < \min\left\{a_i, \frac{z_{i,j} - p_{i,j}}{z_{1,j} - p_{1,j}}, \frac{z_{i+1,j} - p_{i+1,j}}{z_{M,j} - p_{M,j}}\right\}, \\
u_{i,j} > 0,\ v_{i,j} > 0 \text{ and } \gamma_{i,j} > \max\{c^1_{i,j}, c^2_{i,j}\},\ i = 1, 2, \ldots, M-1.
\end{gathered}
$$

□

Using the procedure followed in Theorem 3.2.4, the conditions for y-direction fractal boundary curves to lie above the plane are obtained and given in the following theorem.

Theorem 3.2.5. *Let* $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *be the surface data lies above the plane* Π. *For each fixed* $i = 1, 2, \ldots, M$, $B^*_i(y)$ *is the fractal* *boundary curve along the grid* $x_i \times J$. *Then the fractal boundary curve* $B_i^*(y)$ *lies above the plane if the scaling factors and the shape parameters satisfy the following restrictions:*

$$0 \leq \alpha_{i,j}^* < \min\left\{c_j, \frac{z_{i,j} - p_{i,j}}{z_{i,1} - p_{i,1}}, \frac{z_{i,j+1} - p_{i,j+1}}{z_{i,N} - p_{i,N}}\right\},$$

$$u_{i,j}^* > 0,\ v_{i,j}^* > 0 \text{ and } \gamma_{i,j}^* > \max\{0,\ d_{i,j}^1,\ d_{i,j}^2\},$$

where

$$d_{i,j}^1 = \frac{-u_{i,j}^*[(z_{i,j} - p_{i,j+1}) - \alpha_{i,j}^*(z_{i,1} - p_{i,N})] - u_{i,j}^*[h_j^* z_{i,j}^y - \alpha_{i,j}^*(y_N - y_1) z_{i,1}^y]}{[(z_{i,j} - p_{i,j}) - \alpha_{i,j}^*(z_{i,1} - p_{i,1})]},$$

$$d_{i,j}^2 = \frac{-v_{i,j}^*[(z_{i,j+1} - p_{i,j}) - \alpha_{i,j}^*(z_{j,N} - p_{i,1})] + v_{i,j}^*[h_j^* z_{i,j+1}^y - \alpha_{i,j}^*(y_N - y_1) z_{i,N}^y]}{[(z_{i,j+1} - p_{i,j+1}) - \alpha_{i,j}^*(z_{i,N} - p_{i,N})]},$$

for $j = 1, 2, \ldots, N-1$.

In the following theorem, the conditions for Φ are obtained so that Φ would lie above the plane Π whenever the surface data lies above the plane Π.

Theorem 3.2.6. *Let* $\Pi := \Pi(x, y) = c\Big[1 - \frac{x}{a} - \frac{y}{b}\Big]$ *be a given plane. Suppose the surface data* $\{(x_i, y_j, z_{i,j},\ z_{i,j}^x,\ z_{i,j}^y) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *lies above the plane* Π. *Then,* Φ *to lie above the plane* Π *if scaling factors and shape parameters satisfy the conditions given in Theorems 3.2.4 and 3.2.5.*

Proof. From [28], it can be seen that Φ inherits all the nature of the fractal boundary curves. Hence, the conditions given in Theorems 3.2.4 and 3.2.5 ensure for Φ to lie above the plane Π. □

3.2.3 Monotonicity preserving interpolation

Let $\{(x_i, y_j, z_{i,j}, z_{i,j}^x, z_{i,j}^y) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ be a monotonically increasing data set, i.e., $z_{i,j} \leq z_{i+1,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$, $z_{i,j} \leq z_{i,j+1}$, $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N-1$ with $z_{i,j}^x \geq 0$, $z_{i,j}^y \geq 0$, for $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$.

The following theorem gives the conditions for fractal boundary curves $B_j(x)$, $j = 1, 2, \ldots, N$ along the x-direction to be monotone.

Theorem 3.2.7. *Let* $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M, \ j = 1, 2, \ldots, N\}$ *be a monotonically increasing data set with* $z^x_{i,j} \geq 0, \ z^y_{i,j} \geq 0$, *for* $i = 1, 2, \ldots, M, \ j = 1, 2, \ldots, N$. *For each* $j = 1, 2, \ldots, N$, *let* $B_j(x)$ *be the fractal boundary curve along* $I \times y_j$. *Then the conditions on the scaling factors and the shape parameters for* $B_j(x)$ *to be monotonically increasing are*

$$\alpha^5_{i,j} < \alpha_{i,j} < \alpha^6_{i,j},$$

$$u_{i,j} > 0, \ v_{i,j} > 0 \ \textit{and}$$

$$\gamma_{i,j} > \max\{0, e^1_{i,j}, e^2_{i,j}, e^3_{i,j}, e^4_{i,j}, e^5_{i,j}, e^6_{i,j}, e^7_{i,j}, e^8_{i,j}, e^9_{i,j}, e^{10}_{i,j}, e^{11}_{i,j}, e^{12}_{i,j}\},$$

where

$$\alpha^5_{i,j} = \max\Big\{ -a_i, \frac{a_i(z^x_{i,j} - l_{1,j})}{(z^x_{1,j} - l_{2,j})}, \frac{a_i(z^x_{i+1,j} - l_{1,j})}{(z^x_{M,j} - l_{2,j})}, \frac{a_i(\Delta_{i,j} - l_{1,j})}{(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{2,j})},$$
$$\frac{a_i(l_{2,j} - z^x_{i,j})}{(l_{1,j} - z^x_{1,j})}, \frac{a_i(l_{2,j} - z^x_{i+1,j})}{(l_{1,j} - z^x_{M,j})}, \frac{a_i(l_{2,j} - \Delta_{i,j})}{(l_{1,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})}\Big\},$$
$$\alpha^6_{i,j} = \min\Big\{a_i, \frac{a_i(z^x_{i,j} - l_{1,j})}{(z^x_{1,j} - l_{1,j})}, \frac{a_i(z^x_{i+1,j} - l_{1,j})}{(z^x_{M,j} - l_{1,j})}, \frac{a_i(\Delta_{i,j} - l_{1,j})}{(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{1,j})},$$
$$\frac{a_i(l_{2,j} - z^x_{i,j})}{(l_{2,j} - z^x_{1,j})}, \frac{a_i(l_{2,j} - z^x_{i+1,j})}{(l_{2,j} - z^x_{M,j})}, \frac{a_i(l_{2,j} - \Delta_{i,j})}{(l_{2,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})}\Big\},$$

$$e^1_{i,j} = \frac{-a_i v_{i,j}(\Delta^*_{i,j} - z^{x*}_{i+1,j})}{a_i(\Delta_{i,j} - l_{1,j}) - \alpha_{i,j}(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{1,j})}, \ e^2_{i,j} = \frac{-a_i[(u_{i,j} + v_{i,j})\Delta^*_{i,j} - u_{i,j} z^{x*}_{i,j} - v_{i,j} z^{x*}_{i+1,j}]}{a_i(\Delta_{i,j} - l_{1,j}) - \alpha_{i,j}(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{1,j})},$$
$$e^3_{i,j} = \frac{-a_i u_{i,j}(\Delta^*_{i,j} - z^{x*}_{i,j})}{a_i(\Delta_{i,j} - l_{1,j}) - \alpha_{i,j}(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{1,j})}, \ e^4_{i,j} = \frac{a_i v_{i,j}(\Delta^*_{i,j} - z^{x*}_{i+1,j})}{a_i(l_{2,j} - \Delta_{i,j}) - \alpha_{i,j}(l_{2,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})},$$
$$e^5_{i,j} = \frac{a_i[(u_{i,j} + v_{i,j})\Delta^*_{i,j} - u_{i,j} z^{x*}_{i,j} - v_{i,j} z^{x*}_{i+1,j}]}{a_i(l_{2,j} - \Delta_{i,j}) - \alpha_{i,j}(l_{2,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})}, \ e^6_{i,j} = \frac{a_i u_{i,j}(\Delta^*_{i,j} - z^{x*}_{i,j})}{a_i(l_{2,j} - \Delta_{i,j}) - \alpha_{i,j}(l_{2,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})},$$
$$e^7_{i,j} = \frac{-a_i v_{i,j}(\Delta^*_{i,j} - z^{x*}_{i+1,j})}{a_i(\Delta_{i,j} - l_{1,j}) - \alpha_{i,j}(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{2,j})}, \ e^8_{i,j} = \frac{-a_i[(u_{i,j} + v_{i,j})\Delta^*_{i,j} - u_{i,j} z^{x*}_{i,j} - v_{i,j} z^{x*}_{i+1,j}]}{a_i(\Delta_{i,j} - l_{1,j}) - \alpha_{i,j}(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{2,j})},$$
$$e^9_{i,j} = \frac{-a_i u_{i,j}(\Delta^*_{i,j} - z^{x*}_{i,j})}{a_i(\Delta_{i,j} - l_{1,j}) - \alpha_{i,j}(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{2,j})}, \ e^{10}_{i,j} = \frac{a_i v_{i,j}(\Delta^*_{i,j} - z^{x*}_{i+1,j})}{a_i(l_{2,j} - \Delta_{i,j}) - \alpha_{i,j}(l_{1,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})},$$
$$e^{11}_{i,j} = \frac{a_i[(u_{i,j} + v_{i,j})\Delta^*_{i,j} - u_{i,j} z^{x*}_{i,j} - v_{i,j} z^{x*}_{i+1,j}]}{a_i(l_{2,j} - \Delta_{i,j}) - \alpha_{i,j}(l_{1,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})}, \ e^{12}_{i,j} = \frac{a_i u_{i,j}(\Delta^*_{i,j} - z^{x*}_{i,j})}{a_i(l_{2,j} - \Delta_{i,j}) - \alpha_{i,j}(l_{1,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})},$$

$$0 \leq l_{1,j} < \min\{z^x_{1,j}, z^x_{i+1,j}, \Delta_{i,j}, \frac{z_{M,j} - z_{1,j}}{x_M - x_1}, \ i = 1, 2, \ldots, M-1\},$$

$$l_{2,j} > \max\{z^x_{1,j}, z^x_{i+1,j}, \Delta_{i,j}, \frac{z_{M,j} - z_{1,j}}{x_M - x_1}, \ i = 1, 2, \dots, M-1\},$$
$$z^{x*}_{i,j} = z^x_{i,j} - \frac{\alpha_{i,j} z^x_{1,j}}{a_i}, \ z^{x*}_{i+1,j} = z^x_{i+1,j} - \frac{\alpha_{i,j} z^x_{M,j}}{a_i},$$
$$\Delta^*_{i,j} = \Delta_{i,j} - \alpha_{i,j} \frac{z_{M,j} - z_{1,j}}{h_i} \quad \textit{and} \quad \Delta_{i,j} = \frac{z_{i+1,j} - z_{i,j}}{h_i}.$$

Proof. For $j = 1, 2, \dots, N$, the fractal boundary curve $B_j(x)$ is monotonically increasing over $[x_1, x_M]$ if $B^{(1)}_j(x) \geq 0$ for all $x \in [x_1, x_M]$. It is clear that $B^{(1)}_j(x)$ is also a fractal function and it is the attractor of the IFS $\mathcal{J}^d_j = \{I \times \mathbb{R} : w^d_{i,j}(x, z^d) = (L_i(x), F^1_{i,j}(x, z^d)), \ i = 1, 2, \dots, M-1\}$,

$$F^1_{i,j}(x, z^d) = \frac{\alpha_{i,j} z^d}{a_i} + \frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2},$$

$$\psi_{i,j}(\theta) = \sum_{k=1}^{5} C_{k,i,j} \theta^{k-1} (1-\theta)^{5-k},$$

$$\begin{aligned}
C_{1,i,j} &= u^2_{i,j} z^{x*}_{i,j},\\
C_{2,i,j} &= 2u_{i,j}\{(u_{i,j} + 2v_{i,j} + \gamma_{i,j})\Delta^*_{i,j} - v_{i,j} z^{x*}_{i+1,j}\},\\
C_{3,i,j} &= A_{2,i,j} + A_{4,i,j} - (A_{1,i,j} + A_{5,i,j}) + \gamma_{i,j}(u_{i,j} + v_{i,j} + \gamma_{i,j})\Delta^*_{i,j}\\
&\quad - \gamma_{i,j}(u_{i,j} z^{x*}_{i,j} + v_{i,j} z^{x*}_{i+1,j}),\\
C_{4,i,j} &= 2v_{i,j}\{(2u_{i,j} + v_{i,j} + \gamma_{i,j})\Delta^*_{i,j} - u_{i,j} z^{x*}_{i,j}\},\\
C_{5,i,j} &= v^2_{i,j} z^{x*}_{i+1,j},
\end{aligned}$$

$$z^{x*}_{i,j} = z^x_{i,j} - \frac{\alpha_{i,j} z^x_{1,j}}{a_i}, \ z^{x*}_{i+1,j} = z^x_{i+1,j} - \frac{\alpha_{i,j} z^x_{M,j}}{a_i} \text{ and } \Delta^*_{i,j} = \Delta_{i,j} - \alpha_{i,j} \frac{z_{M,j} - z_{1,j}}{h_i}.$$

It is clear that $B^{(1)}_j(x) \geq 0$ for all $x \in [x_1, x_M]$ if $F^1_{i,j}(x, z^d) \in [l_{1,j}, l_{2,j}]$ for all $(x, z^d) \in [x_1, x_M] \times [l_{1,j}, l_{2,j}]$, $i = 1, 2, \dots, M-1$, $j = 1, 2, \dots, N$.

Next the conditions on the scaling factors $\alpha_{i,j}$ and the shape parameters $u_{i,j}$, $v_{i,j}$ and $\gamma_{i,j}$ are derived so that the range of $F^1_{i,j}$, $i = 1, 2, \dots, M-1$ lies in $[l_{1,j}, l_{2,j}]$, for all $j = 1, 2, \dots, N$.

Case I: Let $0 \leq \alpha_{i,j} < a_i$, $i = 1, 2, \dots, M-1$, $j = 1, 2, \dots, N$. Let $(x, z^d) \in [x_1, x_M] \times [l_{1,j}, l_{2,j}]$. By the assumption on the scaling factors, it is clear that

$$l_{1,j}\alpha_{i,j} \leq z^d \alpha_{i,j} \leq l_{2,j}\alpha_{i,j}$$

and hence we obtain

$$\frac{l_{1,j}\alpha_{i,j}}{a_i}+\frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2}\le\frac{z^d\alpha_{i,j}}{a_i}+\frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2}\le\frac{l_{2,j}\alpha_{i,j}}{a_i}+\frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2}.$$

To prove $F^1_{i,j}(x,z^d)\in[l_{1,j},l_{2,j}]$, it is enough to prove

$$l_{1,j}\le\frac{l_{1,j}\alpha_{i,j}}{a_i}+\frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2}, \tag{3.28}$$

$$\frac{l_{2,j}\alpha_{i,j}}{a_i}+\frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2}\le l_{2,j}, \tag{3.29}$$

for $i=1,2,\dots,M-1$, $j=1,2,\dots,N$. Now (3.28) can be written as

$$\frac{l_{1,j}[\alpha_{i,j}-a_i](q_{i,j}(\theta))^2+a_i\psi_{i,j}(\theta)}{a_i(q_{i,j}(\theta))^2}\ge 0. \tag{3.30}$$

Since $q_{i,j}(\theta)=u_{i,j}(1-\theta)^2+(u_{i,j}+v_{i,j}+\gamma_{i,j})\theta(1-\theta)+v_{i,j}\theta^2>0$, the denominator of (3.30) is positive. It is enough to prove $l_{1,j}[\alpha_{i,j}-a_i](q_{i,j}(\theta))^2+a_i\psi_{i,j}(\theta)\ge 0$. We have $l_{1,j}[\alpha_{i,j}-a_i](q_{i,j}(\theta))^2+a_i\psi_{i,j}(\theta)=C^*_{1,i,j}(1-\theta)^4+C^*_{2,i,j}\theta(1-\theta)^3+C^*_{3,i,j}\theta^2(1-\theta)^2+C^*_{4,i,j}\theta^3(1-\theta)+C^*_{5,i,j}\theta^4$, where

$$\begin{aligned}
C^*_{1,i,j}&=l_{1,j}[\alpha_{i,j}-a_i]u^2_{i,j}+a_iC_{1,i,j},\\
C^*_{2,i,j}&=(2u_{i,j}(u_{i,j}+v_{i,j}+\gamma_{i,j}))(l_{1,j}[\alpha_{i,j}-a_i])+a_iC_{2,i,j},\\
C^*_{3,i,j}&=(2u_{i,j}v_{i,j}+(u_{i,j}+v_{i,j}+\gamma_{i,j})^2)(l_{1,j}[\alpha_{i,j}-a_i])+a_iC_{3,i,j},\\
C^*_{4,i,j}&=(2v_{i,j}(u_{i,j}+v_{i,j}+\gamma_{i,j}))(l_{1,j}[\alpha_{i,j}-a_i])+a_iC_{4,i,j},\\
C^*_{5,i,j}&=l_{1,j}[\alpha_{i,j}-a_i]v^2_{i,j}+a_iC_{5,i,j}.
\end{aligned}$$

It is clear that $l_{1,j}[\alpha_{i,j}-a_i](q_{i,j}(\theta))^2+a_i\psi_{i,j}(\theta)\ge 0$ if the $C^*_{k,i,j}$, $k=1,2,\dots,5$ are nonnegative. One can see that

$$C^*_{1,i,j}\ge 0 \text{ if } \alpha_{i,j}\le\frac{a_i(z^x_{i,j}-l_{1,j})}{(z^x_{1,j}-l_{1,j})} \text{ and } C^*_{5,i,j}\ge 0 \text{ if } \alpha_{i,j}\le\frac{a_i(z^x_{i+1,j}-l_{1,j})}{(z^x_{M,j}-l_{1,j})}.$$

Let $\alpha_{i,j}<\min\left\{\frac{a_i(z^x_{i,j}-l_{1,j})}{(z^x_{1,j}-l_{1,j})},\frac{a_i(z^x_{i+1,j}-l_{1,j})}{(z^x_{M,j}-l_{1,j})},\frac{a_i(\Delta_{i,j}-l_{1,j})}{(\frac{z_{M,j}-z_{1,j}}{x_M-x_1}-l_{1,j})}\right\}$.
We see that $C^*_{2,i,j}\ge 0$ if $(u_{i,j}+v_{i,j}+\gamma_{i,j})[l_{1,j}\alpha_{i,j}-a_il_{1,j}+a_i\Delta^*_{i,j}]+a_iv_{i,j}[\Delta^*_{i,j}-z^{x*}_{i+1,j}]\ge 0$. Since $(u_{i,j}+v_{i,j})[l_{1,j}\alpha_{i,j}-a_il_{1,j}+a_i\Delta^*_{i,j}]\ge 0$, the coefficient $C^*_{2,i,j}\ge 0$ if

$$\gamma_{i,j}>\frac{-a_iv_{i,j}(\Delta^*_{i,j}-z^{x*}_{i+1,j})}{a_i(\Delta_{i,j}-l_{1,j})-\alpha_{i,j}(\frac{z_{M,j}-z_{1,j}}{x_M-x_1}-l_{1,j})}.$$

Now $C_{3,i,j}$ can be re-written as

$C_{3,i,j} = u_{i,j}\{(u_{i,j}+v_{i,j}+\gamma_{i,j})\Delta^*_{i,j}+v_{i,j}\Delta^*_{i,j}-v_{i,j}z^{x*}_{i+1,j}\}+v_{i,j}\{(u_{i,j}+v_{i,j}+\gamma_{i,j})\Delta^*_{i,j}+u_{i,j}\Delta^*_{i,j}-u_{i,j}z^{x*}_{i,j}\}+(u_{i,j}+v_{i,j}+\gamma_{i,j})\{(u_{i,j}+v_{i,j}+\gamma_{i,j})\Delta^*_{i,j}-u_{i,j}z^{x*}_{i,j}-v_{i,j}z^{x*}_{i+1,j}\}+2u_{i,j}v_{i,j}\Delta^*_{i,j}$. Hence

$$
\begin{aligned}
C^*_{3,i,j} = {} & [u_{i,j}(u_{i,j}+v_{i,j}+\gamma_{i,j}) + v_{i,j}(u_{i,j}+v_{i,j}+\gamma_{i,j}) \\
& + \gamma_{i,j}(u_{i,j}+v_{i,j}+\gamma_{i,j})][l_{1,j}\alpha_{i,j} - l_{1,j}a_i] \\
& + a_i u_{i,j}\{(u_{i,j}+v_{i,j}+\gamma_{i,j})\Delta^*_{i,j} + v_{i,j}\Delta^*_{i,j} - v_{i,j}z^{x*}_{i+1,j}\} \\
& + a_i v_{i,j}\{(u_{i,j}+v_{i,j}+\gamma_{i,j})\Delta^*_{i,j} + u_{i,j}\Delta^*_{i,j} - u_{i,j}z^{x*}_{i,j}\} \\
& + a_i(u_{i,j}+v_{i,j}+\gamma_{i,j})\{(u_{i,j}+v_{i,j}+\gamma_{i,j})\Delta^*_{i,j} - u_{i,j}z^{x*}_{i,j} - v_{i,j}z^{x*}_{i+1,j}\} \\
& + 2u_{i,j}v_{i,j}(l_{1,j}\alpha_{i,j} - l_{1,j}a_i + a_i\Delta^*_{i,j}).
\end{aligned}
$$

Further we can simplify as

$$
\begin{aligned}
C^*_{3,i,j} = {} & u_{i,j}\{(u_{i,j}+v_{i,j}+\gamma_{i,j})(l_{1,j}\alpha_{i,j} - l_{1,j}a_i + a_i\Delta^*_{i,j}) + a_i v_{i,j}\Delta^*_{i,j} - a_i v_{i,j}z^{x*}_{i+1,j}\} \\
& + v_{i,j}\{(u_{i,j}+v_{i,j}+\gamma_{i,j})(l_{1,j}\alpha_{i,j} - l_{1,j}a_i + a_i\Delta^*_{i,j}) + a_i u_{i,j}\Delta^*_{i,j} - a_i u_{i,j}z^{x*}_{i,j}\} \\
& + (u_{i,j}+v_{i,j}+\gamma_{i,j})\{\gamma_{i,j}(l_{1,j}\alpha_{i,j} - l_{1,j}a_i + a_i\Delta^*_{i,j}) + a_i(u_{i,j}+v_{i,j})\Delta^*_{i,j} - a_i u_{i,j}z^{x*}_{i,j} \\
& - a_i v_{i,j}z^{x*}_{i+1,j}\} + 2u_{i,j}v_{i,j}(l_{1,j}\alpha_{i,j} - l_{1,j}a_i + a_i\Delta^*_{i,j}).
\end{aligned}
$$

We obtain $C^*_{3,i,j} \geq 0$ if $\gamma_{i,j} \geq \max\{e^1_{i,j},\ e^2_{i,j},\ e^3_{i,j}\}$.

Similarly, one can see that

$$C^*_{4,i,j} \geq 0 \text{ if } \gamma_{i,j} \geq \frac{-a_i u_{i,j}(\Delta^*_{i,j} - z^{x*}_{i,j})}{a_i(\Delta_{i,j} - l_{1,j}) - \alpha_{i,j}(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{1,j})}.$$

Hence the conditions on the scaling factors and the shape parameters to satisfy (3.28) are

$$
\begin{gathered}
0 \leq \alpha_{i,j} < \min\left\{\frac{a_i(z^x_{i,j} - l_{1,j})}{(z^x_{1,j} - l_{1,j})}, \frac{a_i(z^x_{i+1,j} - l_{1,j})}{(z^x_{M,j} - l_{1,j})}, \frac{a_i(\Delta_{i,j} - l_{1,j})}{(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{1,j})}\right\}, \\
u_{i,j} > 0,\ v_{i,j} > 0 \text{ and } \gamma_{i,j} \geq \max\{e^1_{i,j}, e^2_{i,j}, e^3_{i,j}\}.
\end{gathered}
$$

Similarly, the conditions to satisfy (3.29) are

$$
\begin{gathered}
0 \leq \alpha_{i,j} < \min\left\{\frac{a_i(l_{2,j} - z^x_{i,j})}{(l_{2,j} - z^x_{1,j})}, \frac{a_i(l_{2,j} - z^x_{i+1,j})}{(l_{2,j} - z^x_{M,j})}, \frac{a_i(l_{2,j} - \Delta_{i,j})}{(l_{2,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})}\right\}, \\
u_{i,j} > 0,\ v_{i,j} > 0 \text{ and } \gamma_{i,j} \geq \max\{e^4_{i,j}, e^4_{i,j}, e^6_{i,j}\}.
\end{gathered}
$$

Hence, when $0 \leq \alpha_{i,j} < a_i$, the conditions on the scaling factors and the shape parameters for $F^1_{i,j}(x, z^d)$ to lie in $[l_{1,j}, l_{2,j}]$ are

$$0 \leq \alpha_{i,j} < \min\Big\{a_i, \frac{a_i(z^x_{i,j} - l_{1,j})}{(z^x_{1,j} - l_{1,j})}, \frac{a_i(z^x_{i+1,j} - l_{1,j})}{(z^x_{M,j} - l_{1,j})}, \frac{a_i(\Delta_{i,j} - l_{1,j})}{(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{1,j})},$$

$$\frac{a_i(l_{2,j} - z^x_{i,j})}{(l_{2,j} - z^x_{1,j})}, \frac{a_i(l_{2,j} - z^x_{i+1,j})}{(l_{2,j} - z^x_{M,j})}, \frac{a_i(l_{2,j} - \Delta_{i,j})}{(l_{2,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})}\Big\},$$

$$u_{i,j} > 0,\ v_{i,j} > 0 \text{ and } \gamma_{i,j} \geq \max\{e^1_{i,j}, e^2_{i,j}, e^3_{i,j}, e^4_{i,j}, e^4_{i,j}, e^6_{i,j}\}.$$

Case II: Let $-a_i < \alpha_{i,j} \leq 0$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$.

Let $(x, z^d) \in [x_1, x_M] \times [l_{1,j}, l_{2,j}]$. By the assumption on the scaling factors it is clear that $l_{2,j}\alpha_{i,j} \leq z^d\alpha_{i,j} \leq l_{1,j}\alpha_{i,j}$ and hence

$$\frac{l_{2,j}\alpha_{i,j}}{a_i} + \frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2} \leq \frac{z^d\alpha_{i,j}}{a_i} + \frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2} \leq \frac{l_{1,j}\alpha_{i,j}}{a_i} + \frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2}.$$

To prove $F^1_{i,j}(x, z^d) \in [l_{1,j}, l_{2,j}]$, it is sufficient to prove that

$$l_{2,j} \leq \frac{l_{1,j}\alpha_{i,j}}{a_i} + \frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2}, \quad \frac{l_{2,j}\alpha_{i,j}}{a_i} + \frac{\psi_{i,j}(\theta)}{(q_{i,j}(\theta))^2} \leq l_{1,j}, \tag{3.31}$$

for $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$. Using the similar computations as in the case of $0 \leq \alpha_{i,j} < a_i$, it is observed that the conditions to satisfy (3.31) are

$$0 \geq \alpha_{i,j} > \max\Big\{-a_i, \frac{a_i(z^x_{i,j} - l_{1,j})}{(z^x_{1,j} - l_{2,j})}, \frac{a_i(z^x_{i+1,j} - l_{1,j})}{(z^x_{M,j} - l_{2,j})}, \frac{a_i(\Delta_{i,j} - l_{1,j})}{(\frac{z_{M,j}-z_{1,j}}{x_M - x_1} - l_{2,j})},$$

$$\frac{a_i(l_{2,j} - z^x_{i,j})}{(l_{1,j} - z^x_{1,j})}, \frac{a_i(l_{2,j} - z^x_{i+1,j})}{(l_{1,j} - z^x_{M,j})}, \frac{a_i(l_{2,j} - \Delta_{i,j})}{(l_{1,j} - \frac{z_{M,j}-z_{1,j}}{x_M - x_1})}\Big\},$$

$$u_{i,j} > 0,\ v_{i,j} > 0 \text{ and } \gamma_{i,j} \geq \max\{e^7_{i,j}, e^8_{i,j}, e^9_{i,j}, e^{10}_{i,j}, e^{11}_{i,j}, e^{12}_{i,j}\}.$$

This completes the proof of the theorem. □

Applying the same procedure as used in Theorem 3.2.7, the conditions for $B^*_i(y)$, $i = 1, 2, \ldots, M$ to be monotone are derived and presented in the following theorem.

Theorem 3.2.8. *Let* $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *be a monotonically increasing data set, i.e.,* $z_{i,j} \leq z_{i+1,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$, $z_{i,j} \leq z_{i,j+1}$, $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N-1$ *with* $z^x_{i,j} \geq 0$, $z^y_{i,j} \geq 0$, *for* $i =$

$1,2,\ldots,M$, $j=1,2,\ldots,N$. For each $i=1,2,\ldots,M$, let $B_i^*(y)$ be the fractal boundary curve along $x_i\times J$. Then the conditions on the scaling factors and the shape parameters for $B_i^*(y)$ to be monotonically increasing are

$$\alpha_{i,j}^7<\alpha_{i,j}^*<\alpha_{i,j}^8,$$

$$u_{i,j}^*>0,\ v_{i,j}^*>0 \text{ and } \gamma_{i,j}^*>\max\{0,f_{i,j}^1,f_{i,j}^2,f_{i,j}^3,f_{i,j}^4,f_{i,j}^5,f_{i,j}^6,f_{i,j}^7,f_{i,j}^8,f_{i,j}^9,f_{i,j}^{10},f_{i,j}^{11},f_{i,j}^{12}\}$$

$$\alpha_{i,j}^7=\max\Big\{-c_j,\frac{c_j(z_{i,j}^y-l_{1,i}^*)}{(z_{i,1}^y-l_{2,i}^*)},\frac{c_j(z_{i,j+1}^y-l_{1,i}^*)}{(z_{i,N}^y-l_{2,i}^*)},\frac{c_j(\Delta_{i,j}^1-l_{1,i}^*)}{(\frac{z_{i,N}-z_{i,1}}{y_N-y_1}-l_{2,i}^*)},$$
$$\frac{c_j(l_{2,i}^*-z_{i,j}^y)}{(l_{1,i}^*-z_{i,1}^y)},\frac{c_j(l_{2,i}^*-z_{i,j+1}^y)}{(l_{1,i}^*-z_{i,N}^y)},\frac{c_j(l_{2,i}^*-\Delta_{i,j}^1)}{(l_{1,i}^*-\frac{z_{i,N}-z_{i,1}}{y_N-y_1})}\Big\},$$

$$\alpha_{i,j}^8=\min\Big\{c_j,\frac{c_j(z_{i,j}^y-l_{1,i}^*)}{(z_{i,1}^y-l_{1,i}^*)},\frac{c_j(z_{i,j+1}^y-l_{1,i}^*)}{(z_{i,N}^y-l_{1,i}^*)},\frac{c_j(\Delta_{i,j}^1-l_{1,i}^*)}{(\frac{z_{i,N}-z_{i,1}}{y_N-y_1}-l_{1,i}^*)},$$
$$\frac{c_j(l_{2,i}^*-z_{i,j}^y)}{(l_{2,i}^*-z_{i,1}^y)},\frac{c_j(l_{2,i}^*-z_{i,j+1}^y)}{(l_{2,i}^*-z_{i,N}^y)},\frac{c_j(l_{2,i}^*-\Delta_{i,j}^1)}{(l_{2,i}^*-\frac{z_{i,N}-z_{i,1}}{y_N-y_1})}\Big\},$$

$$f_{i,j}^1=\frac{-c_jv_{i,j}^*(\Delta_{i,j}^{1*}-z_{i,j+1}^{y*})}{c_j(\Delta_{i,j}^1-l_{1,i}^*)-\alpha_{i,j}^*(\frac{z_{i,N}-z_{i,1}}{y_N-y_1}-l_{1,i}^*)},\ f_{i,j}^2=\frac{-c_j[(u_{i,j}^*+v_{i,j}^*)\Delta_{i,j}^{1*}-u_{i,j}^*z_{i,j}^{x*}-v_{i,j}^*z_{i+1,j}^{x*}]}{c_j(\Delta_{i,j}^1-l_{1,i}^*)-\alpha_{i,j}^*(\frac{z_{i,N}-z_{i,1}}{y_N-y_1}-l_{1,i}^*)},$$

$$f_{i,j}^3=\frac{-c_ju_{i,j}^*(\Delta_{i,j}^{1*}-z_{i,j}^{y*})}{c_j(\Delta_{i,j}^1-l_{1,i}^*)-\alpha_{i,j}^*(\frac{z_{i,N}-z_{i,1}}{y_N-y_1}-l_{1,i}^*)},\ f_{i,j}^4=\frac{c_jv_{i,j}^*(\Delta_{i,j}^{1*}-z_{i,j+1}^{y*})}{c_j(l_{2,i}^*-\Delta_{i,j}^1)-\alpha_{i,j}^*(l_{2,i}^*-\frac{z_{i,N}-z_{i,1}}{y_N-y_1})},$$

$$f_{i,j}^5=\frac{c_j[(u_{i,j}^*+v_{i,j}^*)\Delta_{i,j}^{1*}-u_{i,j}^*z_{i,j}^{x*}-v_{i,j}^*z_{i+1,j}^{x*}]}{c_j(l_{2,i}^*-\Delta_{i,j}^1)-\alpha_{i,j}^*(l_{2,i}^*-\frac{z_{i,N}-z_{i,1}}{y_N-y_1})},\ f_{i,j}^6=\frac{c_ju_{i,j}^*(\Delta_{i,j}^{1*}-z_{i,j}^{y*})}{c_j(l_{2,i}^*-\Delta_{i,j}^1)-\alpha_{i,j}^*(l_{2,i}^*-\frac{z_{i,N}-z_{i,1}}{y_N-y_1})},$$

$$f_{i,j}^7=\frac{-c_jv_{i,j}^*(\Delta_{i,j}^{1*}-z_{i,j+1}^{y*})}{c_j(\Delta_{i,j}^1-l_{1,i}^*)-\alpha_{i,j}^*(\frac{z_{i,N}-z_{i,1}}{y_N-y_1}-l_{2,i}^*)},\ f_{i,j}^8=\frac{-c_j[(u_{i,j}^*+v_{i,j}^*)\Delta_{i,j}^{1*}-u_{i,j}^*z_{i,j}^{x*}-v_{i,j}^*z_{i+1,j}^{x*}]}{c_j(\Delta_{i,j}^1-l_{1,i}^*)-\alpha_{i,j}^*(\frac{z_{i,N}-z_{i,1}}{y_N-y_1}-l_{2,i}^*)},$$

$$f_{i,j}^9=\frac{-c_ju_{i,j}^*(\Delta_{i,j}^{1*}-z_{i,j}^{y*})}{c_j(\Delta_{i,j}^1-l_{1,i}^*)-\alpha_{i,j}^*(\frac{z_{i,N}-z_{i,1}}{y_N-y_1}-l_{2,i}^*)},\ f_{i,j}^{10}=\frac{c_jv_{i,j}^*(\Delta_{i,j}^{1*}-z_{i,j+1}^{y*})}{c_j(l_{2,i}^*-\Delta_{i,j}^1)-\alpha_{i,j}^*(l_{1,i}^*-\frac{z_{i,N}-z_{i,1}}{y_N-y_1})},$$

$$f_{i,j}^{11}=\frac{c_j[(u_{i,j}^*+v_{i,j}^*)\Delta_{i,j}^{1*}-u_{i,j}^*z_{i,j}^{x*}-v_{i,j}^*z_{i+1,j}^{x*}]}{c_j(l_{2,i}^*-\Delta_{i,j}^1)-\alpha_{i,j}^*(l_{1,i}^*-\frac{z_{i,N}-z_{i,1}}{y_N-y_1})},\ f_{i,j}^{12}=\frac{c_ju_{i,j}^*(\Delta_{i,j}^{1*}-z_{i,j}^{y*})}{c_j(l_{2,i}^*-\Delta_{i,j}^1)-\alpha_{i,j}^*(l_{1,i}^*-\frac{z_{i,N}-z_{i,1}}{y_N-y_1})},$$

$$0\le l_{1,i}^*<\min\{z_{i,1}^y,z_{i,j+1}^y,\Delta_{i,j}^1,\frac{z_{i,N}-z_{i,1}}{y_N-y_1},\ j=1,2,\ldots,N-1\},$$

$$l_{2,i}^*>\max\{z_{i,1}^y,z_{i,j+1}^y,\Delta_{i,j}^1,\frac{z_{i,N}-z_{i,1}}{y_N-y_1},\ j=1,2,\ldots,N-1\},$$

$$z_{i,j}^{y*}=z_{i,j}^y-\frac{\alpha_{i,j}^*z_{i,1}^y}{c_j},\ z_{i,j+1}^{y*}=z_{i,j+1}^y-\frac{\alpha_{i,j}^*z_{i,N}^y}{c_j},$$

$$\Delta_{i,j}^{1*}=\Delta_{i,j}^1-\alpha_{i,j}^*\frac{z_{i,N}-z_{i,1}}{h_j^*}\quad \text{and}\quad \Delta_{i,j}^1=\frac{z_{i,j+1}-z_{i,j}}{h_j^*}.$$

In the following theorem, we obtained the conditions for Φ to be monotonically increasing.

Theorem 3.2.9. *Let* $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *be a monotonic increasing data set, i.e.,* $z_{i,j} \le z_{i+1,j}$, $i = 1, 2, \ldots, M-1$, $j = 1, 2, \ldots, N$, $z_{i,j} \le z_{i,j+1}$, $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N-1$ *with* $z^x_{i,j} \ge 0$, $z^y_{i,j} \ge 0$, *for* $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$. *Then the rational cubic FIS* Φ *is monotonically increasing if the scaling factors and the shape parameters satisfy the conditions given in Theorems 3.2.7 and 3.2.8.*

Proof. Fractal boundary curves $B_j(x)$ and $B^*_i(y)$ for $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$ are monotonically increasing whenever the scaling factors and the shape parameters satisfy the conditions given in Theorems 3.2.7 and 3.2.8 respectively. Hence Φ is monotonically increasing whenever the scaling factors and the shape parameters satisfy the conditions given in Theorems 3.2.7 and 3.2.8. □

3.2.4 Convexity preserving interpolation

Let $\{(x_i, y_j, z_{i,j}, z^x_{i,j}, z^y_{i,j}) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ be a strict convex surface data with

$$z^x_{1,j} < \Delta_{1,j} < z^x_{2,j} < \cdots < z^x_{M-1,j} < \Delta_{M-1,j} < z^x_{M,j},\ j = 1, 2, \ldots, N, \tag{3.32}$$

$$z^y_{i,1} < \Delta^1_{i,1} < z^y_{i,2} < \cdots < z^y_{i,N-1} < \Delta^1_{i,N-1} < z^y_{i,N},\ i = 1, 2, \ldots, M. \tag{3.33}$$

In this section, the convexity of rational cubic FIS Φ is discussed. To preserve the convexity, suitable restrictions are made on the scaling factors and the shape parameters. The following theorem provides the conditions for x-direction fractal boundary curves to be convex.

Theorem 3.2.10. *For each fixed* $j = 1, 2, \ldots, N$, *let* $T_j = \{(x_i,\ z_{i,j},\ z^x_{i,j}) : i = 1,\ 2,\ \ldots,\ M\}$ *be a univariate strict convex data in the grid line* $I \times y_j$ *satisfying (3.32). Let* $B_j(x)$ *be the corresponding fractal boundary curve. Then the conditions on the scaling factors and the shape parameters so that* $B_j(x)$ *satisfies the convexity condition are*

$$0 \leq \alpha_{i,j} < \min\left\{a_i^2, \tfrac{h_i(\Delta_{i,j}-z^x_{i,j})}{(z_{M,j}-z_{1,j})-z^x_{1,j}(x_M-x_1)}, \tfrac{h_i(z^x_{i+1,j}-\Delta_{i,j})}{z^x_{M,j}(x_M-x_1)-(z_{M,j}-z_{1,j})}, \tfrac{a_i(z^x_{i+1,j}-z^x_{i,j})}{(z^x_{M,j}-z^x_{1,j})}\right\},$$

$$u_{i,j} > 0,\ v_{i,j} > 0 \ and\ \gamma_{i,j} \geq \max\left\{0, \frac{v_{i,j}(z^{x*}_{i+1,j} - \Delta^*_{i,j})}{(\Delta^*_{i,j} - z^{x*}_{i,j})}, \frac{u_{i,j}(\Delta^*_{i,j} - z^{x*}_{i,j})}{(z^{x*}_{i+1,j} - \Delta^*_{i,j})}\right\},$$

$i = 1, 2, \ldots, M-1$.

Proof. Since $B_j(x)$ belongs to the class $\mathcal{C}^1(I)$, to prove the convexity, it is sufficient to prove $B_j^{(2)}(x^+)$ or $B_j^{(2)}(x^-)$ exists and is non-negative for each point in (x_1, x_M) [96]. Let $0 \leq \alpha_{i,j} < a_i^2$, $i = 1, 2, \ldots, M-1$. Informally, we have

$$B_j^{(2)}(L_i(x)) = \frac{\alpha_{i,j} B_j^{(2)}(x)}{a_i^2} + \frac{\Psi_{i,j}(\theta)}{h_i(q_{i,j}(\theta))^3},$$

where $\Psi_{i,j}(\theta) = \sum_{k=1}^{6} D_{k,i,j}\theta^{k-1}(1-\theta)^{6-k}$,

$$\begin{aligned}
D_{1,i,j} &= 2u^3_{i,j}(\Delta^{x*}_{i,j} - z^{x*}_{i,j}) + 2u^2_{i,j}v_{i,j}(\Delta^{x*}_{i,j} - z^{x*}_{i,j}) + 2u^2_{i,j}[\gamma_i(\Delta^{x*}_{i,j} - z^{x*}_{i,j}) \\
&\quad - v_{i,j}(z^{x*}_{i+1,j} - \Delta^{x*}_{i,j})],\\
D_{2,i,j} &= 2D_{1,i,j} + 6u^2_{i,j}v_{i,j}(\Delta^{x*}_{i,j} - z^{x*}_{i,j}),\\
D_{3,i,j} &= D_{1,i,j} + 12u^2_{i,j}v_{i,j}(\Delta^{x*}_{i,j} - z^{x*}_{i,j}) + 6u_{i,j}v^2_{i,j}(z^{x*}_{i+1,j} - \Delta^{x*}_{i,j}),\\
D_{4,i,j} &= D_{6,i,j} + 12u_{i,j}v^2_{i,j}(z^{x*}_{i+1,j} - \Delta^{x*}_{i,j}) + 6u^2_{i,j}v_{i,j}(\Delta^{x*}_{i,j} - z^{x*}_{i,j}),\\
D_{5,i,j} &= 2D_{6,i,j} + 6u_{i,j}v^2_{i,j}(z^{x*}_{i+1,j} - \Delta^{x*}_{i,j}),\\
D_{6,i,j} &= 2v^3_{i,j}(z^{x*}_{i+1,j} - \Delta^{x*}_{i,j}) + 2v^2_{i,j}u_{i,j}(z^{x*}_{i+1,j} - \Delta^{x*}_{i,j}) + 2v^2_{i,j}[\gamma_i(z^{x*}_{i+1,j} - \Delta^{x*}_{i,j}) \\
&\quad - u_{i,j}(\Delta^{x*}_{i,j} - z^{x*}_{i,j})].
\end{aligned}$$

Then,

$$B_j^{(2)}(L_i(x^+)) = \frac{\alpha_{i,j} B_j^{(2)}(x^+)}{a_i^2} + \frac{\Psi_{i,j}(\theta)}{h_i(q_{i,j}(\theta))^3}. \tag{3.34}$$

Set $i = 1$ and $x = x_1$ in (3.34), we get

$$B_j^{(2)}(x_1^+) = \frac{D_{1,1,j}}{h_1 u^3_{1,j}}\left[1 - \frac{\alpha_{1,j}}{a_1^2}\right]^{-1}. \tag{3.35}$$

Set $i = M-1$ and $x = x_M$ in (3.34), we obtain

$$B_j^{(2)}(x_M^-) = \frac{D_{6,M-1,j}}{h_{M-1}v^3_{M-1,j}}\left[1 - \frac{\alpha_{M-1,j}}{a^2_{M-1}}\right]^{-1}. \tag{3.36}$$

Set $x = x_1$ gives in (3.34), one can see that

$$B_j^{(2)}(x_i^+) = \frac{\alpha_{i,j} B_j^{(2)}(x_1^+)}{a_i^2} + \frac{D_{1,i,j}}{h_i u_{i,j}^3}, \ i = 2, 3, \dots, M-1. \tag{3.37}$$

From (3.35), (3.36) and (3.37) it is clear that the right handed second derivatives at the knots x_i, $i = 1, 2, \dots, M-1$ and the left handed second derivative at the knot x_M are non-negative if the coefficients $D_{1,i,j}$, $i = 1, 2, \dots, M-1$ and $D_{6,M-1,j}$ are non-negative respectively.

Since $B_j^{(2)}(x^+)$ is determined by iteratively, to prove $B_j^{(2)}(x^+)$ is non-negative, it is enough to prove

$$B_j^{(2)}(L_i(x_n^+)) = \frac{\alpha_{i,j} B_j^{(2)}(x_n^+)}{a_i^2} + R_{i,j}(x_n^+),$$

is non-negative for $n = 1, 2, \dots, M-1$, where $R_{i,j}(x) = \frac{\Psi_{i,j}(\theta)}{h_i(q_{i,j}(\theta))^3}$.

Assume $D_{1,i,j}$, for $i = 1, 2, \dots, M-1$ are non-negative. Then $B_j^{(2)}(L_i(x_n^+))$ is non-negative if $R_{i,j}(x_n) \geq 0$. $R_{i,j}(x_n) \geq 0$ if the coefficients $D_{m,i,j} \geq 0$ if $m = 1, 2, \dots, 6$, $i = 1, 2, \dots, M-1$. Using the Three Chords Lemma for convex functions [145], the end point derivatives should satisfy $z_{1,j}^x < \frac{z_{M,j} - z_{1,j}}{x_M - x_1} < z_{M,j}^x$. It can be seen that

$$(\Delta_{i,j}^* - z_{i,j}^{x*}) > 0 \text{ if } \alpha_{i,j} < \frac{h_i(\Delta_{i,j} - z_{i,j}^x)}{(z_{M,j} - z_{1,j}) - z_{1,j}^x(x_M - x_1)},$$

$$(z_{i+1,j}^{x*} - \Delta_{i,j}^*) > 0 \text{ if } \alpha_{i,j} < \frac{h_i(z_{i+1,j}^x - \Delta_{i,j})}{z_{M,j}^x(x_M - x_1) - (z_{M,j} - z_{1,j})},$$

$$(z_{i+1,j}^{x*} - z_{i,j}^{x*}) > 0 \text{ if } \alpha_{i,j} < \frac{a_i(z_{i+1,j}^x - z_{i,j}^x)}{(z_{M,j}^x - z_{1,j}^x)}.$$

Assume that $0 \leq \alpha_{i,j} < \min\left\{ \frac{h_i(\Delta_{i,j} - z_{i,j}^x)}{(z_{M,j} - z_{1,j}) - z_{1,j}^x(x_M - x_1)}, \frac{h_i(z_{i+1,j}^x - \Delta_{i,j})}{z_{M,j}^x(x_M - x_1) - (z_{M,j} - z_{1,j})}, \frac{a_i(z_{i+1,j}^x - z_{i,j}^x)}{(z_{M,j}^x - z_{1,j}^x)} \right\}$. From the assumptions on the scaling factors and $\gamma_{i,j} \geq \frac{v_{i,j}(z_{i+1,j}^{x*} - \Delta_{i,j}^*)}{(\Delta_{i,j}^* - z_{i,j}^{x*})}$ ensure that $D_{1,i,j} \geq 0$, $D_{2,i,j} \geq 0$ and $D_{3,i,j} \geq 0$. Similarly, from the assumptions on the scaling factors and $\gamma_{i,j} \geq \frac{u_{i,j}(\Delta_{i,j}^* - z_{i,j}^{x*})}{(z_{i+1,j}^{x*} - \Delta_{i,j}^*)}$ ensure that $D_{4,i,j} \geq 0$, $D_{5,i,j} \geq 0$ and $D_{6,i,j} \geq 0$.

Hence, the assumption on the scaling factors and the shape parameters given in the statement of the Theorem 3.2.10 ensure $B_j^{(2)}(L_i(x_n^+)) \geq 0$, for $i = 1, 2, \dots, M-1$, $n = 1, 2, \dots, M-1$ and $B_j^{(2)}(x_M^-) \geq 0$. □

Using the procedure followed in Theorem 3.2.10, in the following theorem, we obtained the conditions for y-direction fractal boundary curves to be convex.

Theorem 3.2.11. *For each fixed* $i = 1, 2, \ldots, M$*, let* $T_i^* = \{(y_j,\ z_{i,j},\ z_{i,j}^y) : j = 1, 2, \ldots, N\}$ *be a univariate strict convex data in the grid line* $x_i \times J$ *satisfying (3.33). Let* $B_i^*(y)$ *be the corresponding fractal boundary curve. Then the conditions on the scaling factors and the shape parameters so that* $B_i^*(y)$ *satisfies convexity are*

$$0 \leq \alpha_{i,j}^* < \min\left\{c_j^2, \frac{h_j^*(\Delta_{i,j}^1 - z_{i,j}^y)}{(z_{i,N} - z_{i,1}) - z_{i,1}^y(y_N - y_1)}, \frac{h_j^*(z_{i,j+1}^y - \Delta_{i,j}^1)}{z_{i,N}^y(y_N - y_1) - (z_{i,N} - z_{i,1})}, \frac{c_j(z_{i,j+1}^y - z_{i,j}^y)}{(z_{i,N}^y - z_{i,1}^y)}\right\},$$

$$u_{i,j}^* > 0,\ v_{i,j}^* > 0 \text{ and } \gamma_{i,j}^* \geq \max\left\{0, \frac{v_{i,j}^*(z_{i,j+1}^{y*} - \Delta_{i,j}^{1*})}{(\Delta_{i,j}^{1*} - z_{i,j}^{y*})}, \frac{u_{i,j}^*(\Delta_{i,j}^{1*} - z_{i,j}^{y*})}{(z_{i,j+1}^{y*} - \Delta_{i,j}^{1*})}\right\},$$

$j = 1, 2, \ldots, N-1.$

The following theorem provides the conditions for Φ to be convex.

Theorem 3.2.12. *Let* $\{(x_i, y_j, z_{i,j}, z_{i,j}^x, z_{i,j}^y) : i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N\}$ *be a strict convex surface data satisfying (3.32) and (3.33). Then, the conditions given in Theorems 3.2.10 and 3.2.11 ensure the convexity of the rational cubic FIS* Φ*.*

Proof. Since Φ inherits all the natures of the fractal boundary curves [28], the conditions given in Theorems 3.2.10 and 3.2.11 ensure the convexity of the rational cubic FIS Φ. □

3.3 Numerical Examples

In this section, numerical illustrations are given to ensure the validity of the results that obtained in Section 3.2. We consider the positive data that is given in Table 3.2. First-order partial derivatives $z^x_{i,j}$ and $z^y_{i,j}$ are computed using arithmetic mean method.

Table 3.2: Positive data

$\downarrow x/y \rightarrow$	1	2	3	4
1	22	38	167	189
2	4	6	41	52
3	181	31	14	17
4	312	298	297	371

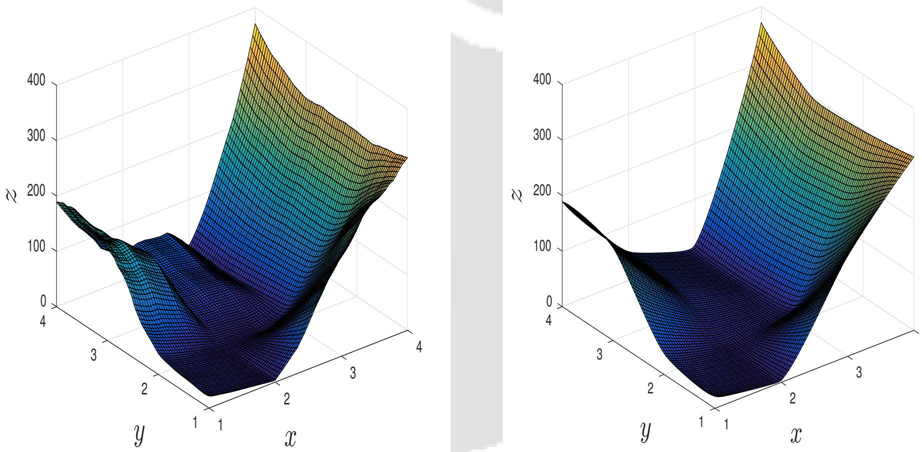
The scaling parameters and the shape parameters are taken according to the Theorems 3.2.1 and 3.2.2, and making use of these, Figure 3.2(a) is constructed. By observing Figure 3.2(a), it is clear that Figure 3.2(a) is positive. By taking all the scaling factors are zero, Figure 3.2(b) is constructed. The values $k_{1,j} = k^*_{1,i} = 0.5$, $k_{2,j} = k^*_{2,i} = 500$ for $i = 1, 2, 3$, $j = 1, 2, 3$ are used to calculate the scaling factors. The parameters that are used to construct Figure 3.2 are given in Tables 3.3 and 3.4.

Table 3.3: Scaling matrices for positive interpolation

Scaling matrices in x-direction	Figs	Scaling matrices in y-direction	Figs
$\boldsymbol{\alpha} = \begin{bmatrix} 0.008 & 0.15 & -0.27 \\ -0.01 & 0.09 & 0.26 \\ -0.18 & -0.05 & 0.08 \\ 0.12 & -0.1 & 0.08 \end{bmatrix}^T$	3.2(a)	$\boldsymbol{\alpha}^* = \begin{bmatrix} -0.03 & 0.3 & 0.29 \\ 0.08 & -0.01 & 0.28 \\ -0.06 & -0.02 & 0.06 \\ 0.3 & -0.28 & 0.29 \end{bmatrix}$	3.2(a)
$\boldsymbol{\alpha} = [0]_{3\times 4}$	3.2(b)	$\boldsymbol{\alpha}^* = [0]_{4\times 3}$	3.2(b)

Table 3.4: Shape parameter matrices for positive interpolation with $\boldsymbol{u} = [1.5]_{3\times4}$, $\boldsymbol{u}^* = [1.5]_{4\times3}$ $\boldsymbol{v} = [1]_{3\times4}$ and $\boldsymbol{v}^* = [1]_{4\times3}$.

Shape parameters matrices in x-direction	Figs	Shape parameters matrices in y-direction	Figs
$\boldsymbol{\gamma} = \begin{bmatrix} 78 & 3 & 3 \\ 3 & 13 & 2 \\ 42 & 60 & 3 \\ 3 & 92 & 3 \end{bmatrix}^T$	3.2(a)	$\boldsymbol{\gamma}^* = \begin{bmatrix} 11 & 3 & 3 \\ 15 & 3 & 3 \\ 4 & 9 & 3 \\ 3 & 3 & 3 \end{bmatrix}$	3.2(a)
$\boldsymbol{\gamma} = \begin{bmatrix} 24 & 3 & 3 \\ 4 & 6 & 3 \\ 3 & 12 & 3 \\ 3 & 12 & 3 \end{bmatrix}^T$	3.2(b)	$\boldsymbol{\gamma}^* = \begin{bmatrix} 4 & 3 & 3 \\ 8 & 3 & 3 \\ 4 & 6 & 3 \\ 3 & 3 & 3 \end{bmatrix}$	3.2(b)



(a) Positive FIS.

(b) Positive classical rational cubic interpolation surface.

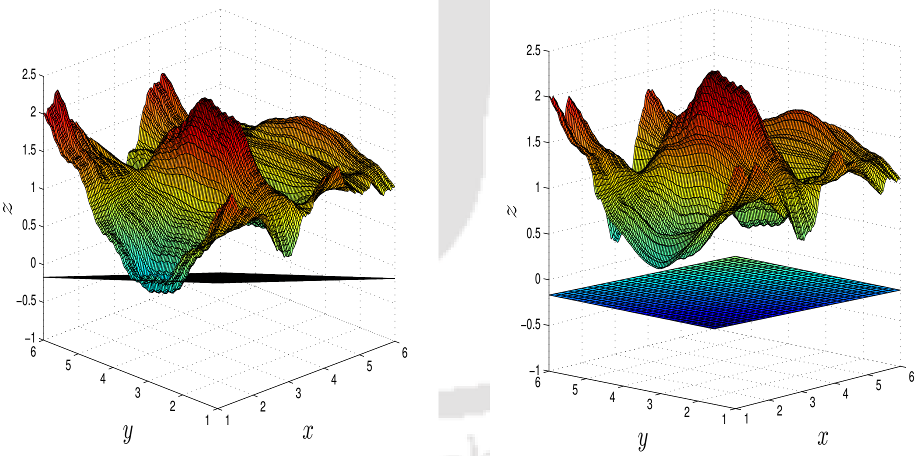
Figure 3.2: Positive surface interpolation.

We consider the constrained data (Table 3.5) which generated from the function $z(x, y) = \sin(x^{1.4})\cos y + 1.2$. This surface data lies above the plane $\Pi(x, y) = \left[1 - \frac{x}{6} - \frac{y}{6}\right]$. The first-order partial derivatives are calculated using the arithmetic mean method. Using arbitrary scaling factors and shape parameters, the rational cubic FIS is constructed which is shown in Figure 3.3(a). It is easy to see that FIS does not lie

Table 3.5: Constrained data

$\downarrow x/y \rightarrow$	1	2	3	4	5	6
1	1.6546	0.8498	0.3670	0.6500	1.4387	2.0080
2	1.4603	0.9995	0.7231	0.8851	1.3366	1.6625
3	0.6606	1.6155	2.1884	1.8526	0.9168	0.2414
4	1.5403	0.9379	0.5766	0.7884	1.3786	1.8047
5	1.1496	1.2388	1.2924	1.2610	1.1735	1.1104
6	1.0505	1.3151	1.4739	1.3808	1.1215	0.9343

completely above the plane Π. To overcome this, scaling factors and shape parameters are taken according to restrictions given in Propositions 3.2.4 and 3.2.5, using these parameters, rational cubic FIS has been constructed which is shown in Figure 3.3(b). This FIS is completely lying above the plane Π. Parameters are used to construct Figure 3.3 are given in Tables 3.6 and 3.7.



(a) Unconstrained FIS. (b) Constrained FIS.

Figure 3.3: Constrained surface interpolation.

Next, we consider the monotonically increasing data as given in the Table 3.8. First-order partial derivatives $z^x_{i,j}$ and $z^y_{i,j}$ are computed using arithmetic mean method. If the partial derivatives $z^x_{i,j}$ or $z^y_{i,j}$ at the point (x_i, y_j) are negative, then the partial derivatives

Table 3.6: Scaling matrices for constrained interpolation.

Scaling matrices in x-direction	Figs	Scaling matrices in y-direction	Figs
$\boldsymbol{\alpha} = [-0.199]_{5\times 6}$	3.3(a)	$\boldsymbol{\alpha}^* = [0.199]_{6\times 5}$	3.3(a)
$\boldsymbol{\alpha} = [0.18]_{5\times 6}$	3.3(b)	$\boldsymbol{\alpha}^* = \begin{bmatrix} 0.16 & 0.01 & 0.03 & 0.18 & 0.18 \\ 0.18 & 0.18 & 0.18 & 0.18 & 0.18 \\ 0.18 & 0.18 & 0.18 & 0.18 & 0.18 \\ 0.18 & 0.18 & 0.18 & 0.18 & 0.18 \\ 0.18 & 0.18 & 0.18 & 0.18 & 0.18 \\ 0.18 & 0.18 & 0.18 & 0.18 & 0.18 \end{bmatrix}$	3.3(b)

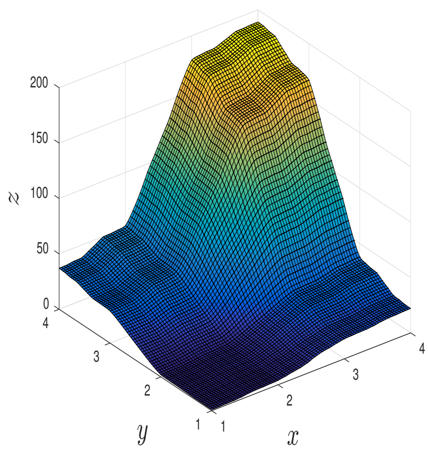
Table 3.7: Shape parameters for constrained interpolation with $\boldsymbol{u} = [1.5]_{5\times 6}$, $\boldsymbol{u}^* = [1.5]_{6\times 5}$, $\boldsymbol{v} = [1]_{5\times 6}$ and $\boldsymbol{v}^* = [1]_{6\times 5}$.

Shape matrices in x-direction	Figs	Shape matrices in y-direction	Figs
$\boldsymbol{\gamma} = [3]_{5\times 6}$	3.3(a)	$\boldsymbol{\gamma}^* = [3]_{6\times 5}$	3.3(a)
$\boldsymbol{\gamma} = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 & 3 \\ 6 & 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 & 3 \end{bmatrix}^T$	3.3(b)	$\boldsymbol{\gamma}^* = [4]_{6\times 5}$	3.3(b)

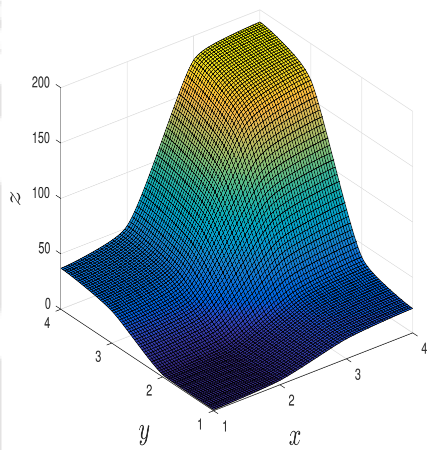
$z^x_{i,j}$ or $z^y_{i,j}$ are assumed to be zero at the point (x_i, y_j). In this case, the values of $l_{1,j}$, $j = 1, 2, \dots, N$ and $l^*_{1,i}$, $i = 1, 2, \dots, M$ becomes zero. Hence to calculate the scaling factors, the fractions whose denominator is zero are omitted. The scaling factors and the shape parameters are calculated according to Theorems 3.2.7 and 3.2.8, making use of these, monotone FIS is constructed which is shown in Figure 3.4(a). By taking all the scaling factors as zero, the classical monotone interpolation surface is constructed and it is given in Figure 3.4(b). The values $l_{1,j} = l^*_{1,i} = 0$, $l_{2,j} = l^*_{2,i} = 200$ for $i = 1, 2, 3$, $j = 1, 2, 3$ are used to calculate the scaling factors. The parameters that are used to construct Figure 3.4 are given in Tables 3.9 and 3.10.

Table 3.8: Monotonic data

↓ x/y →	1	2	3	4
1	1	2	29	37
2	4	6	41	52
3	18	31	149	178
4	22	38	167	189



(a) Monotonic FIS.



(b) Monotonic classical rational cubic interpolation surface.

Figure 3.4: Monotone surface interpolation.

Table 3.9: Scaling matrices for monotone interpolation

Scaling matrices in x-direction	Figs	Scaling matrices in y-direction	Figs
$\boldsymbol{\alpha} = \begin{bmatrix} 0.13 & 0.3 & 0.17 \\ 0.1 & -0.02 & 0.18 \\ 0.07 & -0.09 & 0.11 \\ 0.09 & -0.09 & 0.06 \end{bmatrix}^T$	3.4(a)	$\boldsymbol{\alpha}^* = \begin{bmatrix} 0.02 & -0.02 & 0.2 \\ 0.03 & -0.03 & 0.2 \\ 0.07 & -0.09 & 0.16 \\ 0.08 & -0.11 & 0.11 \end{bmatrix}$	3.4(a)
$\boldsymbol{\alpha} = [0]_{3\times 4}$	3.4(b)	$\boldsymbol{\alpha}^* = [0]_{4\times 3}$	3.4(b)

Table 3.10: Shape parameter matrices for monotone interpolation with $\boldsymbol{u} = [1.5]_{3\times4}$, $\boldsymbol{u}^* = [1.5]_{4\times3}$ $\boldsymbol{v} = [1]_{3\times4}$ and $\boldsymbol{v}^* = [1]_{4\times3}$.

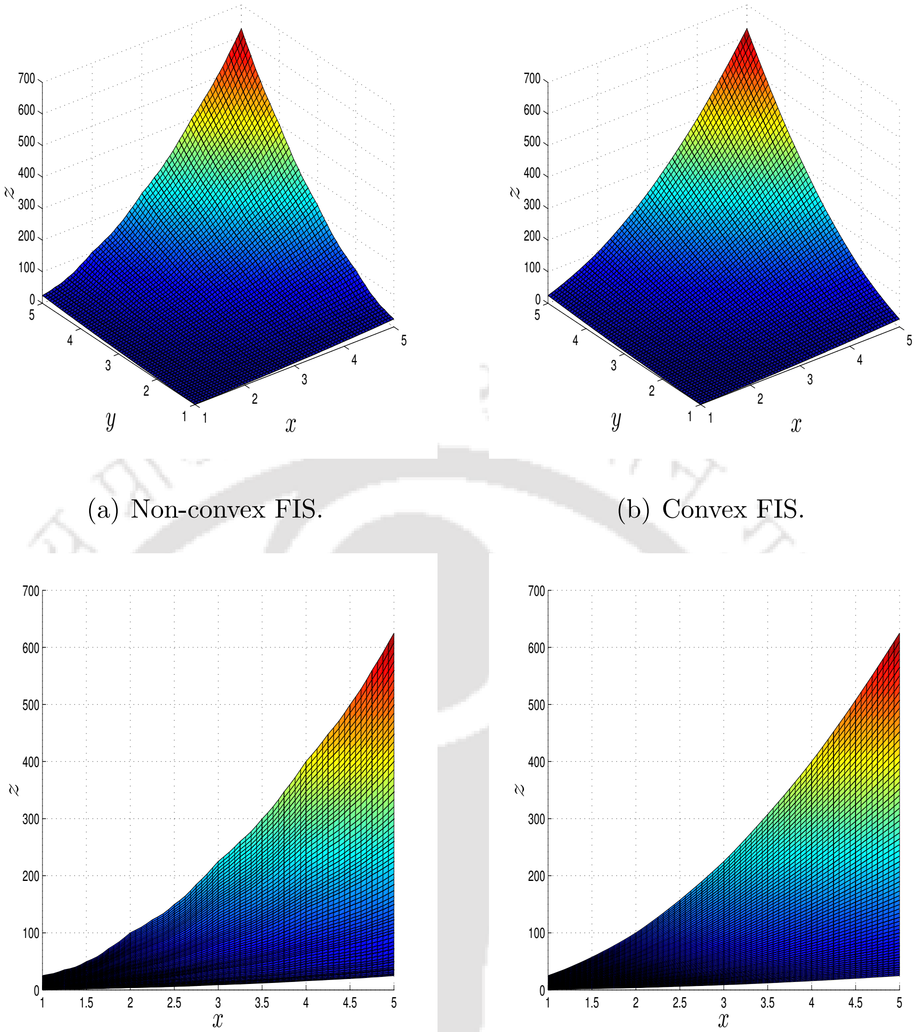
Shape parameters matrices in x-direction	Figs	Shape parameters matrices in y-direction	Figs
$\boldsymbol{\gamma} = \begin{bmatrix} 31 & 3 & 31 \\ 37 & 3 & 46 \\ 25 & 3 & 34 \\ 54 & 4 & 55 \end{bmatrix}^T$	3.4(a)	$\boldsymbol{\gamma}^* = \begin{bmatrix} 51 & 3 & 33 \\ 34 & 3 & 25 \\ 37 & 4 & 32 \\ 28 & 5 & 31 \end{bmatrix}$	3.4(a)
$\boldsymbol{\gamma} = \begin{bmatrix} 3 & 3 & 3 \\ 4 & 3 & 4 \\ 5 & 3 & 5 \\ 5 & 3 & 9 \end{bmatrix}^T$	3.4(b)	$\boldsymbol{\gamma}^* = \begin{bmatrix} 15 & 3 & 3 \\ 10 & 3 & 3 \\ 6 & 3 & 4 \\ 5 & 3 & 55 \end{bmatrix}$	3.4(b)

Next, we consider the convex data which is given in Table 3.11. The first-order partial derivatives are calculated using the arithmetic mean method. By taking arbitrary

Table 3.11: Convex data.

$\downarrow x/y \rightarrow$	1	2	3	4	5
1	1	4	9	16	25
2	4	16	36	64	100
3	9	36	81	144	225
4	16	64	144	256	400
5	25	100	225	400	625

scaling factors, FIS is constructed and it is shown in Figure 3.5(a). By observing Figure 3.5(a), it can be seen that the FIS is not convex. Hence, the scaling factors and the shape parameters are restricted according to restrictions given in Theorems 3.2.10 and 3.2.11. Using these restricted scaling factors and shape parameters, convex FIS has been constructed and is shown in Figure 3.5(b). Figures 3.5(c) and 3.5(d) represents the xz-view of the Figures 3.5(a) and 3.5(b) respectively. The parameters are used to construct Figure 3.5 are given given Tables 3.12 and 3.13.



(a) Non-convex FIS. (b) Convex FIS.

(c) xz-view of Figure 3.5(a). (d) xz-view of Figure 3.5(b).

Figure 3.5: Convex surface interpolation.

Table 3.12: Scaling matrices for convex interpolation

Scaling matrices in x-direction	Figs	Scaling matrices in y-direction	Figs
$\boldsymbol{\alpha} = [0.24]_{4\times 5}$	3.5(a),3.5(c)	$\boldsymbol{\alpha}^* = [0.24]_{5\times 4}$	3.5(a),3.5(c)
$\boldsymbol{\alpha} = [0.04]_{4\times 5}$	3.5(b),3.5(d)	$\boldsymbol{\alpha}^* = [0.04]_{5\times 4}$	3.5(b),3.5(d)

Table 3.13: Shape matrices for convex interpolation with $\boldsymbol{u} = [1.5]_{4\times 5}$, $\boldsymbol{u}^* = [1.5]_{5\times 4}$, $\boldsymbol{v} = [1]_{4\times 5}$ and $\boldsymbol{v}^* = [1]_{5\times 4}$.

Shape matrices in x-direction	Figs	Shape matrices in y-direction	Figs
$\boldsymbol{\gamma} = [3]_{4\times 5}$	3.5(a),3.5(c)	$\boldsymbol{\gamma}^* = [3]_{5\times 4}$	3.5(a),3.5(c)
$\boldsymbol{\gamma} = [5]_{4\times 5}$	3.5(b),3.5(d)	$\boldsymbol{\gamma}^* = [5]_{5\times 4}$	3.5(b),3.5(d)

3.4 Conclusion

In this chapter, a new type of FIS is constructed over a rectangular grid. The constructed FIS not only interpolate the surface data but also preserves the shape features of the surface data. Since shapes of the interpolant can be adjusted by using different choices of the scaling factors, developed scheme offers a large amount of flexibility for simulation or modeling of objects with smooth geometric shapes. The constructed interpolation scheme may play important role in computer graphics, engineering design, and data visualization problems.

Chapter 4

Fractal Cubic Spline Solutions for Self-Adjoint Boundary-Value Problems

Singular perturbation problems commonly occur in many branches of science and engineering, for instance fluid mechanics, elasticity, quantum mechanics, chemical reactor theory, convection-diffusion processes, optimal control, etc. The solution of singular perturbation problems exhibits a multi-scale character, that is, there are thin layers where the solution varies rapidly, while away from the layers the solution behaves regularly and varies slowly. So the numerical treatment of singular perturbation problems is computationally difficult. Surla and Stojanović [132] generated a difference scheme via spline in tension to get the numerical solutions of the self-adjoint singularly perturbed BVPs. Bawa and Natesan [13] established a numerical method using cubic spline method for the numerical solutions of self-adjoint singularly perturbed BVPs with mixed boundary conditions. Also, Bawa and Natesan [14] used adaptive spline to solve the self-adjoint BVPs with the Dirichlet boundary conditions.

In this chapter, we consider the self-adjoint singularly perturbed BVPs with the Dirichlet and Neumann boundary conditions. These BVPs exhibit boundary layers at both ends of the interval. We take the continuity conditions of the fractal cubic spline as the discretized equations to solve the BVPs. The discretization equations given by the continuity conditions are second-order and hence the resulting fractal cubic spline method is second-order. The present chapter is organized as follows: In Section 4.1, singularly perturbed BVPs with the Dirichlet boundary conditions are solved using fractal cubic spline method. Convergence analysis of the proposed method is established. Numerical results are given to show the practical applicability of the proposed method and numerical results are compared with cubic spline method results. In Section 4.2, numerical solutions of singularly perturbed BVPs with the Neumann boundary conditions are obtained using the fractal cubic spline method. Error analysis of the derived method is discussed. The obtained numerical results are tabulated and compared with the numerical results of cubic spline method.

4.1 Dirichlet Boundary-Value Problem

Consider the Dirichlet BVP

$$\begin{cases} -\varepsilon u''(x) + q(x)u(x) = f(x), \; x \in (0,1), \\ u(0) = \eta_0, \quad u(1) = \eta_1, \end{cases} \tag{4.1}$$

where $0 < \varepsilon \leq 1$, q and f are sufficiently smooth functions in $[0,1]$ and $q(x) > 0$ in $[0,1]$.

In order to develop a numerical method for the numerical solution of differential equation (4.1), let us consider a uniform mesh on the interval $I = [0,1]$ such that $0 = x_0 < x_1 < \cdots < x_N = 1$, where $x_i = ih$, $i = 0,1,\ldots,N$ and $h = 1/N$. Let $U_0, U_1, \ldots, U_N$ be the approximations of $u(x)$ and $M_0, M_1, \ldots, M_N$ be the approximations of $u''(x)$ at $x_0, x_1, \ldots, x_N$. Let $q_i = q(x_i)$, $f_i = f(x_i)$, $i = 0,1,\ldots,N$.

Consider the IFS $\Big\{I \times \mathbb{R}; w_i(x,y) = (L_i(x), F_i(x,y)) : i = 1,2,\ldots,N\Big\}$, where $L_i : I \to [x_{i-1}, x_i]$ is defined by $L_i(x) = hx + x_{i-1}$, $x \in I$, $F_i : I \times \mathbb{R} \to \mathbb{R}$ is defined by $F_i(x,y) = h^2\alpha y + h^2 r_i(x)$, $(x,y) \in I \times \mathbb{R}$ with $r_i(x) = A_i(x-x_0)^3 + B_i(x-x_0)^2 + C_i(x-x_0) + D_i$ and α is the scaling factor such that $|\alpha| < 1$. It can be observed that $L_i(x_0) = x_{i-1}$ and $L_i(x_N) = x_i$, $i = 1,2,\ldots,N$.

We assume the following conditions on the IFS:

$$F_i(x_0, U_0) = U_{i-1},\ F_i(x_N, U_N) = U_i, \quad i = 1, 2, \dots, N,$$

$$F_{i,1}(x_N, U_{N,1}) = F_{i+1,1}(x_0, U_{0,1}),\ i = 1, 2, \dots, N-1,$$

$$F_{i,2}(x_0, M_0) = M_{i-1},\ F_{i,2}(x_N, M_N) = M_i, \quad i = 1, 2, \dots, N,$$

where $F_{i,k}(x, y) = \frac{h^2\alpha y + h^2 r_i^{(k)}(x)}{h^k}$, $k = 1, 2$, $U_{0,1} = \frac{h^2 r_1^{(1)}(x_0)}{h - h^2\alpha}$ and $U_{N,1} = \frac{h^2 r_N^{(1)}(x_N)}{h - h^2\alpha}$.

It can be seen that, the conditions prescribed on the IFS

$$\Big\{I \times \mathbb{R}; w_i(x, y) = (L_i(x), F_i(x, y)) : i = 1, 2, \dots, N\Big\}$$

satisfy the conditions for constructing two times continuously differentiable FIFs (see [11, 30]).

Let $\mathscr{F} = \{\phi \in \mathcal{C}^2(I, \mathbb{R}) \mid \phi(x_0) = U_0,\ \phi(x_N) = U_N,\ \phi^{(2)}(x_0) = M_0,\ \phi^{(2)}(x_N) = M_N\}$. Then $(\mathscr{F}, \rho)$ is a complete metric space where ρ is the metric induced by the norm $\|\phi\| = \|\phi\|_\infty + \|\phi^{(1)}\|_\infty + \|\phi^{(2)}\|_\infty$. Define the Read-Bajraktarević operator T on $(\mathscr{F}, \rho)$ as

$$T\phi(L_i(x)) = F_i(x, \phi(x)) = h^2\alpha\phi(x) + h^2 r_i(x),\ x \in I,\ i = 1, 2, \dots, N.$$

The map T is a contraction map and hence the fixed-point Φ (say) of T satisfies the functional equation

$$\Phi(L_i(x)) = F_i(x, \Phi(x)) = h^2\alpha\Phi(x) + h^2 r_i(x),\ x \in I,\ i = 1, 2, \dots, N. \tag{4.2}$$

Since $\Phi^{(2)} \in \mathcal{C}^2(I, \mathbb{R})$, $\Phi^{(2)}$ satisfies the functional equation

$$\Phi^{(2)}(L_i(x)) = \alpha\Phi^{(2)}(x) + r_i^{(2)}(x),\ x \in I,\ i = 1, 2, \dots, N. \tag{4.3}$$

The conditions $F_i(x_0, U_0) = U_{i-1}$, $F_i(x_N, U_N) = U_i$, $F_{i,2}(x_0, M_0) = M_{i-1}$, $F_{i,2}(x_N, M_N) = M_i$ can be reformulated as $\Phi(x_{i-1}) = U_{i-1}$, $\Phi(x_i) = U_i$, $\Phi^{(2)}(x_{i-1}) = M_{i-1}$, $\Phi^{(2)}(x_i) = M_i$ respectively [36]. For the fractal cubic spline Φ, the constants A_i, B_i, C_i and D_i are evaluated based on the conditions $\Phi(x_{i-1}) = U_{i-1}$, $\Phi(x_i) = U_i$, $\Phi^{(2)}(x_{i-1}) = M_{i-1}$, $\Phi^{(2)}(x_i) = M_i$. Substituting $x = x_0$ in (4.3), we get

$$B_i = \frac{1}{2}\Big[M_{i-1} - \alpha M_0\Big],$$

and substituting $x = x_N$ in (4.3), we obtain

$$A_i = \frac{1}{6}\Big[M_i - \alpha M_N\Big] - \frac{1}{6}\Big[M_{i-1} - \alpha M_0\Big].$$

Substituting $x = x_0$ in (4.2), we obtain

$$D_i = \frac{1}{h^2}\Big[U_{i-1} - h^2\alpha U_0\Big],$$

and substituting $x = x_N$ in (4.2), we get

$$C_i = \frac{1}{h^2}\Big[U_i - \alpha h^2 U_N\Big] - \frac{1}{h^2}\Big[U_{i-1} - \alpha h^2 U_0\Big] - \frac{1}{6}\Big[M_i - \alpha M_N\Big] - \frac{1}{3}\Big[M_{i-1} - \alpha M_0\Big].$$

The FIF $\Phi^{(1)}$ satisfies the following functional equation:

$$\Phi^{(1)}(L_i(x)) = \alpha h \Phi^{(1)}(x) + h r_i^{(1)}(x),\ x \in I,\ i = 1, 2, \dots, N.$$

The condition $F_{i,1}(x_N, U_{N,1}) = F_{i+1,1}(x_0, U_{0,1}),\ i = 1, 2, \dots, N-1$, can be reformulated as

$$\Phi^{(1)}(L_i(x_N)) = \Phi^{(1)}(L_{i+1}(x_0)),\ i = 1, 2, \dots, N-1,$$

which coincide with the standard continuity condition

$$\Phi^{(1)}(x_i^-) = \Phi^{(1)}(x_i^+),\ i = 1, 2, \dots, N-1. \tag{4.4}$$

On $[x_{i-1}, x_i]$, we have

$$\Phi^{(1)}(L_i(x)) = h\alpha\Phi^{(1)}(x) + 3A_i h(x - x_0)^2 + 2B_i h(x - x_0) + C_i h,$$

on $[x_i, x_{i+1}]$, we get

$$\Phi^{(1)}(L_{i+1}(x)) = h\alpha\Phi^{(1)}(x) + 3A_{i+1} h(x - x_0)^2 + 2B_{i+1} h(x - x_0) + C_{i+1} h.$$

Making use of (4.4), for each node x_i, $i = 1, 2, \dots, N-1$, we get the continuity condition for $\Phi^{(1)}$ as

$$\alpha\Phi^{(1)}(x_N) + 3A_i + 2B_i + C_i = \alpha\Phi^{(1)}(x_0) + C_{i+1},$$

which is simplified as

$$\begin{aligned}&-3\alpha\Phi^{(1)}(x_0)-\frac{3\alpha}{2}M_0+\frac{M_{i-1}}{2}+2M_i+\frac{M_{i+1}}{2}-\frac{3\alpha}{2}M_N+3\alpha\Phi^{(1)}(x_N)\\&=\frac{3}{h^2}[U_{i+1}-2U_i+U_{i-1}].\end{aligned}\tag{4.5}$$

The differential equation given in (4.1) is discretized at x_i as $-\varepsilon M_i+q_iU_i=f_i$, $i=0,1,\ldots,N$. We take the continuity condition given in (4.5), we substitute

$$M_i=\frac{q_iU_i-f_i}{\varepsilon},\Phi^{(1)}(x_0)=\frac{-3U_0+4U_1-U_2}{2h},\Phi^{(1)}(x_N)=\frac{3U_N-4U_{N-1}+U_{N-2}}{2h},$$

and hence, we get the following system:

$$\begin{aligned}&-\left[\frac{6\alpha\varepsilon}{h}\right]U_1+\left[\frac{3\alpha\varepsilon}{2h}\right]U_2+\left[\frac{q_{i-1}}{2}-\frac{3\varepsilon}{h^2}\right]U_{i-1}+\left[2q_i+\frac{6\varepsilon}{h^2}\right]U_i+\left[\frac{q_{i+1}}{2}\right.\\&\left.-\frac{3\varepsilon}{h^2}\right]U_{i+1}+\left[\frac{3\alpha\varepsilon}{2h}\right]U_{N-2}-\left[\frac{6\alpha\varepsilon}{h}\right]U_{N-1}=-\frac{3\alpha f_0}{2}+\frac{f_{i-1}}{2}+2f_i+\frac{f_{i+1}}{2}\\&-\frac{3\alpha f_N}{2}-\left[-\frac{3\alpha}{2}q_0+\frac{9\alpha\varepsilon}{2h}\right]\eta_0-\left[-\frac{3\alpha}{2}q_N+\frac{9\alpha\varepsilon}{2h}\right]\eta_1,\ i=1,2,\ldots,N-1,\end{aligned}\tag{4.6}$$

where $U_0=\eta_0$ and $U_N=\eta_1$ are the discretizations for the Dirichlet boundary conditions. The system given in (4.6) gives the approximations $U_1,U_2,\ldots,U_{N-1}$ of the solution $u(x)$ at $x_1,x_2,\ \ldots,\ x_{N-1}$.

Remark 4.1.1. *If $\alpha=0$, then the system given in (4.6) will become the system corresponding to the cubic spline as*

$$\left[\frac{q_{i-1}}{2}-\frac{3\varepsilon}{h^2}\right]U_{i-1}+\left[2q_i+\frac{6\varepsilon}{h^2}\right]U_i+\left[\frac{q_{i+1}}{2}-\frac{3\varepsilon}{h^2}\right]U_{i+1}=\frac{f_{i-1}}{2}+2f_i+\frac{f_{i+1}}{2},$$

$i=1,2,\ldots,N-1.$

4.1.1 Convergence analysis

In this section, convergence analysis of the proposed numerical scheme is investigated. Let $|\alpha|< h^2$. Assume that the unknown function $u(x)$ is sufficiently smooth. First we derive the truncation error $T_i(h)$, $i=1,2,\ldots,N-1$ associated with the difference equations given in (4.6) as follows:

By replacing U_i by $u(x_i)$ in (4.6), we get

$$T_i(h)=[r_i^0u(x_0)+r_i^1u(x_1)+r_i^2u(x_2)+r_i^-u(x_{i-1})+r_i^cu(x_i)+r_i^+u(x_{i+1})$$

$$+ r_i^{N-2}u(x_{N-2}) + r_i^{N-1}u(x_{N-1}) + r_i^N u(x_N)] - [q_i^0 f_0 + q_i^- f_{i-1}$$
$$+ q_i^c f_i + q_i^+ f_{i+1} + q_i^N f_N], \tag{4.7}$$

where

$$r_i^0 = -\frac{3\alpha}{2}q_0 + \frac{9\alpha\varepsilon}{2h},\ r_i^1 = -\frac{6\alpha\varepsilon}{h},\ r_i^2 = \frac{3\alpha\varepsilon}{2h},\ r_i^- = \frac{q_{i-1}}{2} - \frac{3\varepsilon}{h^2},\ r_i^c = 2q_i + \frac{6\varepsilon}{h^2},$$
$$r_i^+ = \frac{q_{i+1}}{2} - \frac{3\varepsilon}{h^2},\ r_i^{N-2} = \frac{3\alpha\varepsilon}{2h},\ r_i^{N-1} = -\frac{6\alpha\varepsilon}{h},\ r_i^N = -\frac{3\alpha}{2}q_N + \frac{9\alpha\varepsilon}{2h},$$
$$q_i^0 = -\frac{3\alpha}{2},\ q_i^- = \frac{1}{2},\ q_i^c = 2,\ q_i^+ = \frac{1}{2},\ q_i^N = -\frac{3\alpha}{2}.$$

Using differential equation (4.1) in (4.7), we get

$$T_i(h) = [r_i^0 u(x_0) + r_i^1 u(x_1) + r_i^2 u(x_2) + r_i^- u(x_{i-1}) + r_i^c u(x_i) + r_i^+ u(x_{i+1})$$
$$+ r_i^{N-2}u(x_{N-2}) + r_i^{N-1}u(x_{N-1}) + r_i^N u(x_N)] - [q_i^0(-\varepsilon u''(x_0) + q_0 u(x_0))$$
$$+ q_i^-(-\varepsilon u''(x_{i-1}) + q_{i-1}u(x_{i-1})) + q_i^c(-\varepsilon u''(x_i) + q_i u(x_i))$$
$$+ q_i^+(-\varepsilon u''(x_{i+1}) + q_{i+1}u(x_{i+1})) + q_i^N(-\varepsilon u''(x_N) + q_N u(x_N))]. \tag{4.8}$$

After simplifying (4.8), we obtain

$$T_i(h) = -3\alpha\varepsilon\Big[\frac{-3u(x_0) + 4u(x_1) - u(x_2)}{2h}\Big] + 3\alpha\varepsilon\Big[\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h}\Big]$$
$$- \frac{3\alpha\varepsilon}{2}u''(x_0) - \frac{3\alpha\varepsilon}{2}u''(x_N) - \frac{3\varepsilon}{h^2}u(x_{i-1}) + \frac{6\varepsilon}{h^2}u(x_i) - \frac{3\varepsilon}{h^2}u(x_{i+1})$$
$$+ \frac{\varepsilon}{2}u''(x_{i-1}) + 2\varepsilon u''(x_i) + \frac{\varepsilon}{2}u''(x_{i+1}). \tag{4.9}$$

Using the Taylor expansions for $u(x_{i-1})$, $u(x_{i+1})$, $u''(x_{i-1})$, $u''(x_{i+1})$ about the point x_i

$$u(x_{i-1}) = u(x_i) - hu'(x_i) + \frac{h^2}{2!}u''(x_i) - \frac{h^3}{3!}u'''(x_i) + \frac{h^4}{4!}u^{(iv)}(x_i) - \cdots,$$
$$u(x_{i+1}) = u(x_i) + hu'(x_i) + \frac{h^2}{2!}u''(x_i) + \frac{h^3}{3!}u'''(x_i) + \frac{h^4}{4!}u^{(iv)}(x_i) + \cdots,$$
$$u''(x_{i-1}) = u''(x_i) - hu'''(x_i) + \frac{h^2}{2!}u^{(iv)}(x_i) - \cdots,$$
$$u''(x_{i+1}) = u''(x_i) + hu'''(x_i) + \frac{h^2}{2!}u^{(iv)}(x_i) + \cdots,$$

also

$$\frac{-3u(x_0) + 4u(x_1) - u(x_2)}{2h} = u'(x_0) - \frac{h^2}{3}u'''(x_0) - \cdots,$$

$$\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h} = u'(x_N) - \frac{h^2}{3}u'''(x_N) + \cdots$$

in (4.9), we get

$$\begin{aligned} T_i(h) = &- 3\alpha\varepsilon\Big[u'(x_0) - \frac{h^2}{3}u'''(x_0) + \mathcal{O}(h^3)\Big] + 3\alpha\varepsilon\Big[u'(x_N) - \frac{h^2}{3}u'''(x_N) + \mathcal{O}(h^3)\Big] \\ &- \frac{3\alpha\varepsilon}{2}u''(x_0) - \frac{3\alpha\varepsilon}{2}u''(x_N) + \frac{\varepsilon h^2}{4}u^{(iv)}(x_i) + \mathcal{O}(h^3). \end{aligned} \tag{4.10}$$

When $h \to 0$, we get $|T_i(h)| \leq K_i h^2$, where K_i is a constant. Thus $T_i(h) = \mathcal{O}(h^2)$ as $h \to 0$ for $i = 1, 2, \ldots, N-1$.

The system (4.6) can be written as $AU = d$, where

$$A = \begin{bmatrix}
A_{1,1} & A_{1,2} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{1,N-2} & A_{1,N-1} \\
A_{2,1} & A_{2,2} & A_{2,3} & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{2,N-2} & A_{2,N-1} \\
A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{3,N-2} & A_{3,N-1} \\
A_{4,1} & A_{4,2} & A_{4,3} & A_{4,4} & A_{4,5} & 0 & \ldots & 0 & 0 & 0 & 0 & A_{4,N-2} & A_{4,N-1} \\
A_{5,1} & A_{5,2} & 0 & A_{5,4} & A_{5,5} & A_{5,6} & \ldots & 0 & 0 & 0 & 0 & A_{5,N-2} & A_{5,N-1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
A_{N-5,1} & A_{N-5,2} & 0 & 0 & 0 & 0 & \ldots & A_{N-5,N-6} & A_{N-5,N-5} & A_{N-5,N-4} & 0 & A_{N-5,N-2} & A_{N-5,N-1} \\
A_{N-4,1} & A_{N-4,2} & 0 & 0 & 0 & 0 & \ldots & 0 & A_{N-4,N-5} & A_{N-4,N-4} & A_{N-4,N-3} & A_{N-4,N-2} & A_{N-4,N-1} \\
A_{N-3,1} & A_{N-3,2} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & A_{N-3,N-4} & A_{N-3,N-3} & A_{N-3,N-2} & A_{N-3,N-1} \\
A_{N-2,1} & A_{N-2,2} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & A_{N-2,N-3} & A_{N-2,N-2} & A_{N-2,N-1} \\
A_{N-1,1} & A_{N-1,2} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{N-1,N-2} & A_{N-1,N-1}
\end{bmatrix}$$

$A_{i,j}$ is the coefficient of U_j, $U = (U_1, U_2, \ldots, U_{N-1})^T$, $d = (d_1, d_2, \ldots, d_{N-1})^T$, d_i, $i = 1, 2, \ldots, N-1$ are the right hand side of the system.

In the following proposition, suitable conditions on h and α are derived to prove the diagonal dominance of the matrix A. Let $q_* = \min\{q(x) : x \in [0,1]\}$, $q^* = \max\{q(x) : x \in [0,1]\}$.

Proposition 4.1.1. *Let A be the matrix corresponding to the system (4.6). Let h satisfies $h^2 q^* < 6\varepsilon$. Then, the following conditions on the scaling factor ensure the matrix A to be strictly diagonally dominant:*

$$0 \leq \alpha < \min\Big\{h^2, \ \frac{6\varepsilon - h^2 q^*}{3h\varepsilon}, \ \frac{hq_*}{6\varepsilon}\Big\}.$$

Proof. Let $0 \leq \alpha < h^2$. We get $A_{1,1} = -\frac{6\alpha\varepsilon}{h} + 2q_1 + \frac{6\varepsilon}{h^2} = 2q_1 + 6\varepsilon\Big[\frac{1}{h^2} - \frac{\alpha}{h}\Big] > 2q_1 + 6\varepsilon[\frac{1}{h^2} - h] > 0$ when $\alpha < h^2$, $A_{1,2} = \frac{3\alpha\varepsilon}{2h} + \frac{q_2}{2} - \frac{3\varepsilon}{h^2} < 0$ if $\alpha < \frac{6\varepsilon - h^2 q^*}{3h\varepsilon}$, $A_{1,N-2} = \frac{3\alpha\varepsilon}{2h} \geq 0$

and $A_{1,N-1} = -\frac{6\alpha\varepsilon}{h} \le 0$. We obtain

$$|A_{1,1}| - \sum_{j=2}^{N-1} |A_{1,j}| = -\frac{6\alpha\varepsilon}{h} + 2q_1 + \frac{6\varepsilon}{h^2} + \frac{3\alpha\varepsilon}{2h} + \frac{q_2}{2} - \frac{3\varepsilon}{h^2} - \frac{3\alpha\varepsilon}{2h} - \frac{6\alpha\varepsilon}{h}$$
$$= \left[q_1 - \frac{6\alpha\varepsilon}{h}\right] + \left[q_1 - \frac{6\alpha\varepsilon}{h}\right] + \left[\frac{q_2}{2} + \frac{3\varepsilon}{h^2}\right] > 0 \text{ if } \alpha < \frac{hq_*}{6\varepsilon}.$$

Thus, when $h^2q^* < 6\varepsilon$ and $\alpha < \min\left\{h^2, \frac{6\varepsilon - h^2q^*}{3h\varepsilon}, \frac{hq_*}{6\varepsilon}\right\}$, we get $|A_{1,1}| - \sum_{j=2}^{N-1} |A_{1,j}| > 0$.

We have $A_{2,1} = -\frac{6\alpha\varepsilon}{h} + \frac{q_1}{2} - \frac{3\varepsilon}{h^2} < 0$ if $h^2q^* < 6\varepsilon$, $A_{2,2} = \frac{3\alpha\varepsilon}{2h} + 2q_2 + \frac{6\varepsilon}{h^2} > 0$, $A_{2,3} = \frac{q_3}{2} - \frac{3\varepsilon}{h^2} < 0$ if $h^2q^* < 6\varepsilon$, $A_{2,N-2} = \frac{3\alpha\varepsilon}{2h} \ge 0$ and $A_{2,N-1} = -\frac{6\alpha\varepsilon}{h} \le 0$. We have

$$|A_{2,2}| - \sum_{j=1,\, j\neq 2}^{N-1} |A_{2,j}| = \frac{3\alpha\varepsilon}{2h} + 2q_2 + \frac{6\varepsilon}{h^2} - \frac{6\alpha\varepsilon}{h} + \frac{q_1}{2} - \frac{3\varepsilon}{h^2}$$
$$+ \frac{q_3}{2} - \frac{3\varepsilon}{h^2} - \frac{3\alpha\varepsilon}{2h} - \frac{6\alpha\varepsilon}{h}$$
$$= \left[q_2 - \frac{6\alpha\varepsilon}{h}\right] + \left[q_2 - \frac{6\alpha\varepsilon}{h}\right] + \frac{q_1}{2} + \frac{q_3}{2} > 0$$

if $\alpha < \frac{hq_*}{6\varepsilon}$. Hence, when $h^2q^* < 6\varepsilon$ and $\alpha < \frac{hq_*}{6\varepsilon}$, we obtain $|A_{2,2}| - \sum_{j=1,\, j\neq 2}^{N-1} |A_{2,j}| > 0$.

We obtain $A_{3,1} = -\frac{6\alpha\varepsilon}{h} \le 0$, $A_{3,2} = \frac{3\alpha\varepsilon}{2h} + \frac{q_2}{2} - \frac{3\varepsilon}{h^2} < 0$ if $\alpha < \frac{6\varepsilon - h^2q^*}{3h\varepsilon}$, $A_{3,3} = 2q_3 + \frac{6\varepsilon}{h^2} > 0$, $A_{3,4} = \frac{q_4}{2} - \frac{3\varepsilon}{h^2} < 0$ if $h^2q^* < 6\varepsilon$, $A_{3,N-2} = \frac{3\alpha\varepsilon}{2h} \ge 0$ and $A_{3,N-1} = -\frac{6\alpha\varepsilon}{h} \le 0$. We obtain

$$|A_{3,3}| - \sum_{j=1,\, j\neq 3}^{N-1} |A_{3,j}| = 2q_3 + \frac{6\varepsilon}{h^2} - \frac{6\alpha\varepsilon}{h} + \frac{3\alpha\varepsilon}{2h} + \frac{q_2}{2} - \frac{3\varepsilon}{h^2}$$
$$+ \frac{q_4}{2} - \frac{3\varepsilon}{h^2} - \frac{3\alpha\varepsilon}{2h} - \frac{6\alpha\varepsilon}{h}$$
$$= \left[q_3 - \frac{6\alpha\varepsilon}{h}\right] + \left[q_3 - \frac{6\alpha\varepsilon}{h}\right] + \frac{q_2}{2} + \frac{q_4}{2} > 0$$

if $\alpha < \frac{hq_*}{6\varepsilon}$. Thus, when $h^2q^* < 6\varepsilon$ and $\alpha < \min\left\{\frac{6\varepsilon - h^2q^*}{3h\varepsilon}, \frac{hq_*}{6\varepsilon}\right\}$, we obtain that $|A_{3,3}| - \sum_{j=1,\, j\neq 3}^{N-1} |A_{3,j}| > 0$.

For $i = 4, 5, \ldots, N-4$, we have $A_{i,1} = -\frac{6\alpha\varepsilon}{h} \le 0$, $A_{i,2} = \frac{3\alpha\varepsilon}{2h} \ge 0$, $A_{i,i-1} = \frac{q_{i-1}}{2} - \frac{3\varepsilon}{h^2} < 0$ if $h^2q^* < 6\varepsilon$, $A_{i,i} = 2q_i + \frac{6\varepsilon}{h^2} > 0$, $A_{i,i+1} = \frac{q_{i+1}}{2} - \frac{3\varepsilon}{h^2} < 0$ if $h^2q^* < 6\varepsilon$, $A_{i,N-2} = \frac{3\alpha\varepsilon}{2h} \ge 0$ and $A_{i,N-1} = -\frac{6\alpha\varepsilon}{h} \le 0$. We have

$$|A_{i,i}| - \sum_{j=1,\, j\neq i}^{N-1} |A_{i,j}| = 2q_i + \frac{6\varepsilon}{h^2} - \frac{6\alpha\varepsilon}{h} - \frac{3\alpha\varepsilon}{2h} + \frac{q_{i-1}}{2} - \frac{3\varepsilon}{h^2}$$

$$
\begin{aligned}
&+\frac{q_{i+1}}{2}-\frac{3\varepsilon}{h^2}-\frac{3\alpha\varepsilon}{2h}-\frac{6\alpha\varepsilon}{h}\\
=&\left[q_i-\frac{6\alpha\varepsilon}{h}\right]+\left[q_i-\frac{6\alpha\varepsilon}{h}\right]+\left[\frac{q_{i-1}}{2}-\frac{3\alpha\varepsilon}{h}\right]+\frac{q_{i+1}}{2}\\
>&\,0 \text{ if } \alpha<\frac{hq_*}{6\varepsilon}.
\end{aligned}
$$

Hence, we have $|A_{i,i}| - \sum_{j=1,\, j\neq i}^{N-1} |A_{i,j}| > 0$ if $h^2\ q^* < 6\varepsilon$ and $\alpha < \frac{hq_*}{6\varepsilon}$.

We get $A_{N-3,1} = -\frac{6\alpha\varepsilon}{h} \leq 0$, $A_{N-3,2} = \frac{3\alpha\varepsilon}{2h} \geq 0$, $A_{N-3,N-4} = \frac{q_{N-4}}{2} - \frac{3\varepsilon}{h^2} < 0$ if $h^2q^* < 6\varepsilon$, $A_{N-3,N-3} = 2q_{N-3} + \frac{6\varepsilon}{h^2} > 0$, $A_{N-3,N-2} = \frac{q_{N-2}}{2} - \frac{3\varepsilon}{h^2} + \frac{3\alpha\varepsilon}{2h} < 0$ if $\alpha < \frac{6\varepsilon - h^2q^*}{3h\varepsilon}$, $A_{N-3,N-1} = -\frac{6\alpha\varepsilon}{h} \leq 0$. We get

$$
\begin{aligned}
&|A_{N-3,N-3}| - \sum_{j=1,\, j\neq N-3}^{N-1} |A_{N-3,j}|\\
&= 2q_{N-3}+\frac{6\varepsilon}{h^2}-\frac{6\alpha\varepsilon}{h}-\frac{3\alpha\varepsilon}{2h}+\frac{q_{N-4}}{2}-\frac{3\varepsilon}{h^2}+\frac{q_{N-2}}{2}-\frac{3\varepsilon}{h^2}+\frac{3\alpha\varepsilon}{2h}-\frac{6\alpha\varepsilon}{h}\\
&= \left[q_{N-3}-\frac{6\alpha\varepsilon}{h}\right]+\left[q_{N-3}-\frac{6\alpha\varepsilon}{h}\right]+\frac{q_{N-4}}{2}+\frac{q_{N-2}}{2} > 0 \text{ if } \alpha<\frac{hq_*}{6\varepsilon}.
\end{aligned}
$$

Thus, we obtain $|A_{N-3,N-3}| - \sum_{j=1,\, j\neq N-3}^{N-1} |A_{N-3,j}| > 0$ if $\alpha < \min\left\{\frac{hq_*}{6\varepsilon}, \frac{6\varepsilon-h^2q^*}{3h\varepsilon}\right\}$ and $h^2q^* < 6\varepsilon$.

We obtain $A_{N-2,1} = -\frac{6\alpha\varepsilon}{h} \leq 0$, $A_{N-2,2} = \frac{3\alpha\varepsilon}{2h} \geq 0$, $A_{N-2,N-3} = \frac{q_{N-3}}{2} - \frac{3\varepsilon}{h^2} < 0$ if $h^2q^* < 6\varepsilon$, $A_{N-2,N-2} = 2q_{N-2} + \frac{6\varepsilon}{h^2} + \frac{3\alpha\varepsilon}{2h} > 0$, $A_{N-2,N-1} = \frac{q_{N-1}}{2} - \frac{3\varepsilon}{h^2} - \frac{6\alpha\varepsilon}{h} < 0$ if $h^2q^* < 6\varepsilon$. We have

$$
\begin{aligned}
&|A_{N-2,N-2}| - \sum_{j=1,\, j\neq N-2}^{N-1} |A_{N-2,j}|\\
&= 2q_{N-2}+\frac{6\varepsilon}{h^2}+\frac{3\alpha\varepsilon}{2h}-\frac{6\alpha\varepsilon}{h}-\frac{3\alpha\varepsilon}{2h}+\frac{q_{N-3}}{2}-\frac{3\varepsilon}{h^2}+\frac{q_{N-1}}{2}-\frac{3\varepsilon}{h^2}-\frac{6\alpha\varepsilon}{h}\\
&= \left[q_{N-2}-\frac{6\alpha\varepsilon}{h}\right]+\left[q_{N-2}-\frac{6\alpha\varepsilon}{h}\right]+\frac{q_{N-3}}{2}+\frac{q_{N-1}}{2} > 0 \text{ if } \alpha<\frac{hq_*}{6\varepsilon}.
\end{aligned}
$$

Hence, we get $|A_{N-2,N-2}| - \sum_{j=1,\, j\neq N-2}^{N-1} |A_{N-2,j}| > 0$ if $\alpha < \frac{hq_*}{6\varepsilon}$ and $h^2q^* < 6\varepsilon$.

We get $A_{N-1,1} = -\frac{6\alpha\varepsilon}{h} \leq 0$, $A_{N-1,2} = \frac{3\alpha\varepsilon}{2h} \geq 0$, $A_{N-1,N-2} = \frac{q_{N-2}}{2} - \frac{3\varepsilon}{h^2} + \frac{3\alpha\varepsilon}{2h} < 0$ if $\alpha < \frac{6\varepsilon-h^2q^*}{3h\varepsilon}$, $A_{N-1,N-1} = 2q_{N-1} + \frac{6\varepsilon}{h^2} - \frac{6\alpha\varepsilon}{h} > 2q_{N-1} + 6\varepsilon\left[\frac{1}{h^2} - h\right] > 0$ when $\alpha < h^2$. We obtain

$$
\begin{aligned}
&|A_{N-1,N-1}| - \sum_{j=1}^{N-2} |A_{N-1,j}| \\
&= 2q_{N-1} + \frac{6\varepsilon}{h^2} - \frac{6\alpha\varepsilon}{h} - \frac{6\alpha\varepsilon}{h} - \frac{3\alpha\varepsilon}{2h} + \frac{q_{N-2}}{2} - \frac{3\varepsilon}{h^2} + \frac{3\alpha\varepsilon}{2h} \\
&= \left[q_{N-1} - \frac{6\alpha\varepsilon}{h}\right] + \left[q_{N-1} - \frac{6\alpha\varepsilon}{h}\right] + \frac{q_{N-2}}{2} + \frac{3\varepsilon}{h^2} > 0 \text{ if } \alpha < \frac{hq_*}{6\varepsilon}.
\end{aligned}
$$

Hence, we get $|A_{N-1,N-1}| - \sum_{j=1}^{N-2} |A_{N-1,j}| > 0$ if $h^2 q^* < 6\varepsilon$ and $\alpha < \min\left\{h^2, \frac{6\varepsilon - h^2 q^*}{3h\varepsilon}, \frac{hq_*}{6\varepsilon}\right\}$.

Thus the conditions prescribed in the statement of the Proposition 4.1.1 ensure the strictly diagonal dominance of the matrix A. □

The following theorem provides the error estimation of the proposed numerical method.

Theorem 4.1.1. *Let $u(x)$ be the exact solution of the singularly perturbed differential equation given in (4.1). Let U be the approximate solution obtained by the system given in (4.6). If we choose h and α according to Proposition 4.1.1, then $\|E\|_\infty = \mathcal{O}(h^2)$, where $E = (E_1, E_2, \ldots, E_{N-1})^T$, $E_i = u(x_i) - U_i$.*

Proof. The system (4.6) can be written as

$$AU = d$$

and the system (4.6) with exact solution can be written as

$$A\bar{u} = d + T(h),$$

where $\bar{u} = (u(x_1), u(x_2), \ldots, u(x_{N-1}))^T$, $T(h) = (T_1(h), T_2(h), \ldots, T_{N-1}(h))^T$. Hence, we get

$$A(\bar{u} - U) = T(h), \text{ i.e, } AE = T(h).$$

By taking h and α according to the Proposition 4.1.1, we have

$$E = A^{-1}T(h)$$

and hence we get

$$E_j = \sum_{i=1}^{N-1} A^{-1}_{j,i} T_i(h),\ j = 1, 2, \ldots, N-1.$$

Using the theory of matrices, we have

$$\sum_{i=1}^{N-1} A^{-1}_{k,i} \mathcal{S}_i = 1,\ k = 1, 2, \ldots, N-1,$$

where $\mathcal{S}_i$ is the i-th row sum of the matrix A. The row sums

$$\mathcal{S}_1 = \left[q_1 - \frac{6\alpha\varepsilon}{h}\right] + \left[q_1 - \frac{6\alpha\varepsilon}{h}\right] + \frac{q_2}{2} + \frac{3\alpha\varepsilon}{h} + \frac{3\varepsilon}{h^2} > 0 \text{ if } \alpha < \frac{hq_*}{6\varepsilon},$$
$$\mathcal{S}_i = \left[q_i - \frac{6\alpha\varepsilon}{h}\right] + \left[q_i - \frac{6\alpha\varepsilon}{h}\right] + \frac{q_{i-1}}{2} + \frac{q_{i+1}}{2} + \frac{3\alpha\varepsilon}{h} > 0 \text{ if } \alpha < \frac{hq_*}{6\varepsilon},$$
$$i = 2, 3, \ldots, N-2,$$
$$\mathcal{S}_{N-1} = \left[q_{N-1} - \frac{6\alpha\varepsilon}{h}\right] + \left[q_{N-1} - \frac{6\alpha\varepsilon}{h}\right] + \frac{q_{N-2}}{2} + \frac{3\alpha\varepsilon}{h} + \frac{3\varepsilon}{h^2} > 0 \text{ if } \alpha < \frac{hq_*}{6\varepsilon}.$$

It can be observed that, the conditions given in statement of Proposition 4.1.1 ensures the row sums of the matrix A are positive. We have

$$\sum_{i=1}^{N-1} A^{-1}_{k,i} \le \frac{1}{\min_{1\le i\le N-1} \mathcal{S}_i} = \frac{1}{C_{i_0}},$$

where C_{i_0} is a constant. We obtain

$$\left|E_j\right| \le \frac{Kh^2}{C_{i_0}},$$

where K is a constant, and hence we get $\|E\|_\infty = \mathcal{O}(h^2)$ as $h \to 0$. □

Remark 4.1.2. *Let us consider the restriction on α given in Proposition 4.1.1,*

$$0 \le \alpha < \min\left\{h^2, \frac{6\varepsilon - h^2 q^*}{3h\varepsilon}, \frac{hq_*}{6\varepsilon}\right\}.$$

When h is sufficiently small, h^2 will be the minimum in the above range, i.e., $0 \le \alpha < h^2$. For fixed h, by varying α, we can compute the numerical solutions of the BVP.

4.1.2 Numerical examples

In this section, to demonstrate the applicability of the developed method computationally, we consider the two numerical examples whose exact solutions are known. For any value of N and ε, the maximum point-wise error E^N will be calculated by

$$E^N = \max_i |u(x_i) - U_i|$$

and the order of convergence will be calculated by

$$p^N = \log_2\left(\frac{E^N}{E^{2N}}\right).$$

Example 4.1.1. *Consider the following BVP [8]:*

$$\begin{cases} -\varepsilon u''(x) + u(x) = 1 + 2\sqrt{\varepsilon}[\exp((x-1)/\sqrt{\varepsilon}) + \exp(-x/\sqrt{\varepsilon})], \; x \in (0,1), \\ u(0) = 0, \quad u(1) = 0. \end{cases}$$

The analytical solution is $u(x) = 1 - (1-x)\exp(-x/\sqrt{\varepsilon}) - x\exp((x-1)/\sqrt{\varepsilon})$. For each fixed $\varepsilon = 2^{1-i}$, $i = 1, 2, \ldots, 9$, we take $N = 16, 32, 64, 128, 256, 512, 1024$ and we get the α restriction in the range $0 \le \alpha < h^2$. Hence for each fixed ε, we have taken as $\alpha = 0.9999h^2$. The maximum point-wise errors and the order of convergence are calculated and that are given in Table 4.1. Figure 4.1(a) represents the numerical solutions of the Example 4.1.1. Figure 4.1(b) represents Log-log plots of the maximum errors for the Example 4.1.1. The maximum point-wise errors and the order of convergence corresponding to the cubic spline method are given in Table 4.2.

Example 4.1.2. *Consider the following BVP [134]:*

$$\begin{cases} -\varepsilon u''(x) + u(x) = -\cos^2(\pi x) - 2\varepsilon\pi^2\cos(2\pi x), \; x \in (0,1), \\ u(0) = 0, \quad u(1) = 0. \end{cases}$$

The analytical solution is $u(x) = \dfrac{\exp(-(1-x)/\sqrt{\varepsilon}) + \exp(-x/\sqrt{\varepsilon})}{1 + \exp(-1/\sqrt{\varepsilon})} - \cos^2(\pi x)$. For each fixed $\varepsilon = 2^{1-i}$, $i = 1, 2, \ldots, 9$, we take $N = 16, 32, 64, 128, 256, 512, 1024$ and we obtain the α restriction in the range $0 \le \alpha < h^2$. Thus, for each fixed ε, we have taken $\alpha = 0.9995h^2$. Table 4.3 represents the maximum point-wise errors and the

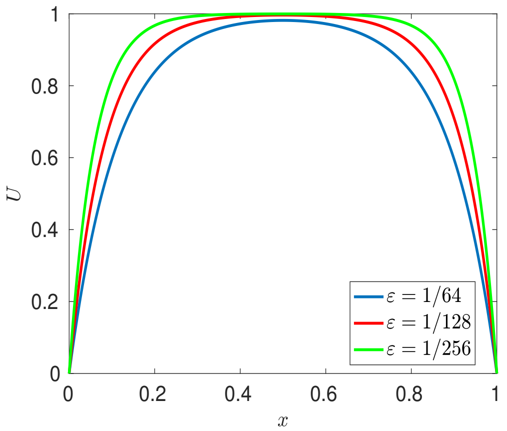
Table 4.1: Maximum point-wise error and order of convergence corresponding to Example 4.1.1.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256	512	1024
1	4.2114e-06	5.1410e-07	9.2670e-08	2.0855e-08	5.0670e-09	1.2575e-09	3.1374e-10
	3.0342	2.4719	2.1517	2.0412	2.0105	2.0030	
1/2	1.2600e-05	1.9419e-06	4.0454e-07	9.5897e-08	2.3641e-08	5.8894e-09	1.4709e-09
	2.6979	2.2631	2.0767	2.0202	2.0051	2.0014	
1/4	3.8924e-05	7.3312e-06	1.6699e-06	4.0685e-07	1.0103e-07	2.5216e-08	6.3009e-09
	2.4086	2.1343	2.0372	2.0096	2.0025	2.0007	
1/8	1.2045e-04	2.6072e-05	6.2366e-06	1.5618e-06	3.9102e-07	9.7796e-08	2.4451e-08
	2.2078	2.0637	1.9976	1.9979	1.9994	1.9999	
1/16	3.5394e-04	9.0815e-05	2.3136e-05	5.7936e-06	1.4490e-06	3.6230e-07	9.0579e-08
	1.9625	1.9728	1.9976	1.9994	1.9998	1.9999	
1/32	1.1473e-03	2.9413e-04	7.3535e-05	1.8450e-05	4.6127e-06	1.1532e-06	2.8829e-07
	1.9637	2.0000	1.9948	2.0000	2.0000	2.0000	
1/64	3.0690e-03	7.9945e-04	1.9890e-04	4.9688e-05	1.2430e-05	3.1074e-06	7.7684e-07
	1.9407	2.0069	2.0011	1.9990	2.0001	2.0000	
1/128	7.7843e-03	1.8606e-03	4.6487e-04	1.1613e-04	2.9033e-05	7.2577e-06	1.8145e-06
	2.0648	2.0009	2.0011	1.9999	2.0001	2.0000	
1/256	1.7587e-02	4.0431e-03	9.9093e-04	2.4802e-04	6.1929e-05	1.5482e-05	3.8703e-06
	2.1209	2.0286	1.9983	2.0018	2.0000	2.0001	

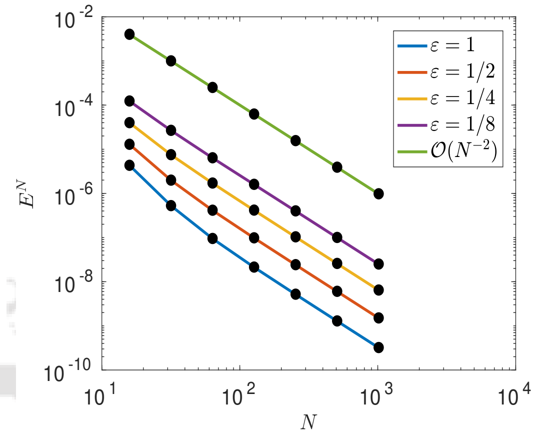
order of convergence corresponding to the Example 4.1.2. Figure 4.2(a) represents the numerical solutions of the Example 4.1.2. Figure 4.2(b) represents Log-log plots of the maximum errors for the Example 4.1.2. The maximum point-wise errors and the order of convergence corresponding to the cubic spline method are given in Table 4.4.

Table 4.2: Maximum point-wise error and order of convergence corresponding to Example 4.1.1 with cubic spline.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256	512	1024
1	2.0728e-04	5.1797e-05	1.2948e-05	3.2369e-06	8.0922e-07	2.0230e-07	5.0576e-08
	2.0006	2.0002	2.0000	2.0000	2.0000	2.0000	
1/2	4.6839e-04	1.1700e-04	2.9245e-05	7.3108e-06	1.8277e-06	4.5692e-07	1.1423e-07
	2.0012	2.0003	2.0001	2.0000	2.0000	2.0000	
1/4	9.4245e-04	2.3526e-04	5.8793e-05	1.4697e-05	3.6742e-06	9.1853e-07	2.2963e-07
	2.0022	2.0005	2.0001	2.0000	2.0000	2.0000	
1/8	1.6114e-03	4.0181e-04	1.0039e-04	2.5093e-05	6.2729e-06	1.5682e-06	3.9205e-07
	2.0037	2.0009	2.0002	2.0001	2.0000	2.0000	
1/16	2.2402e-03	5.5761e-04	1.3925e-04	3.4804e-05	8.7004e-06	2.1751e-06	5.4377e-07
	2.0063	2.0016	2.0004	2.0001	2.0000	2.0000	
1/32	3.2076e-03	7.9459e-04	1.9820e-04	4.9521e-05	1.2380e-05	3.0949e-06	7.7372e-07
	2.0132	2.0033	2.0008	2.0000	2.0001	2.0000	
1/64	5.4621e-03	1.3399e-03	3.3339e-04	8.3250e-05	2.0817e-05	5.2038e-06	1.3009e-06
	2.0274	2.0068	2.0017	1.9997	2.0001	2.0000	
1/128	9.8006e-03	2.4459e-03	6.0625e-04	1.5179e-04	3.7926e-05	9.4800e-06	2.3699e-06
	2.0025	2.0124	1.9978	2.0008	2.0002	2.0001	
1/256	2.0154e-02	4.6639e-03	1.1446e-03	2.8485e-04	7.1131e-05	1.7783e-05	4.4455e-06
	2.1114	2.0267	2.0066	2.0016	1.9999	2.0001	

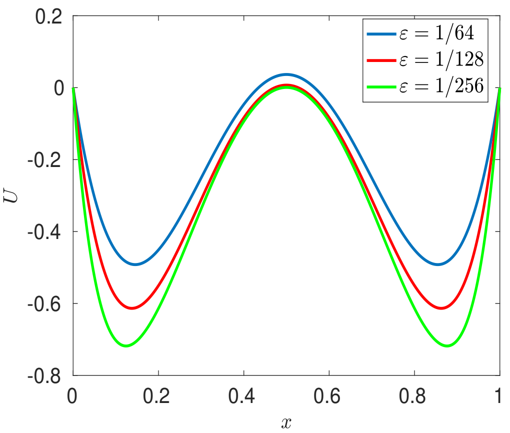


(a) Numerical solutions for Example 4.1.1 with $N = 512$.

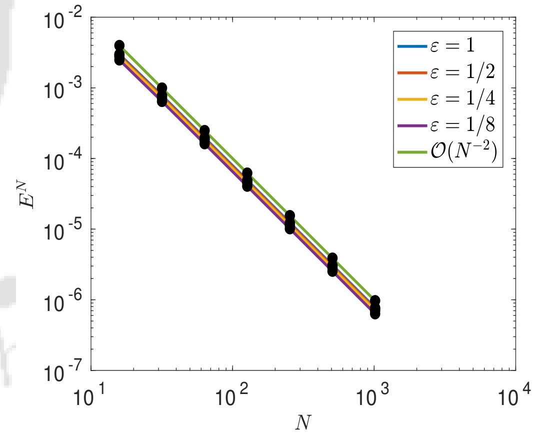


(b) Log-log plot for Example 4.1.1.

Figure 4.1: Numerical results of Example 4.1.1.



(a) Numerical solutions for Example 4.1.2 with $N = 512$.



(b) Log-log plot for Example 4.1.2.

Figure 4.2: Numerical results of Example 4.1.2.

Table 4.3: Maximum point-wise error and order of convergence corresponding to Example 4.1.2.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256	512	1024
1	2.9355e-03	7.6628e-04	1.9332e-04	4.8430e-05	1.2113e-05	3.0287e-06	7.5721e-07
	1.9377	1.9869	1.9970	1.9993	1.9998	2.0000	
1/2	2.8473e-03	7.4168e-04	1.8704e-04	4.6852e-05	1.1718e-05	2.9300e-06	7.3251e-07
	1.9407	1.9875	1.9971	1.9993	1.9998	2.0000	
1/4	2.6944e-03	6.9966e-04	1.7633e-04	4.4164e-05	1.1046e-05	2.7618e-06	6.9047e-07
	1.9452	1.9884	1.9973	1.9994	1.9998	2.0000	
1/8	2.4411e-03	6.3155e-04	1.5906e-04	3.9834e-05	9.9627e-06	2.4909e-06	6.2274e-07
	1.9506	1.9893	1.9975	1.9994	1.9999	2.0000	
1/16	2.0383e-03	5.2037e-04	1.3102e-04	3.2811e-05	8.2062e-06	2.0518e-06	5.1296e-07
	1.9698	1.9897	1.9976	1.9994	1.9999	2.0000	
1/32	1.3014e-03	3.2864e-04	8.2905e-05	2.0773e-05	5.1962e-06	1.2992e-06	3.2482e-07
	1.9855	1.9870	1.9967	1.9992	1.9998	1.9999	
1/64	8.1959e-04	1.9765e-04	4.9015e-05	1.2231e-05	3.0563e-06	7.6411e-07	1.9104e-07
	2.0519	2.0117	2.0027	2.0007	1.9999	1.9999	
1/128	4.9785e-03	1.1875e-03	2.9359e-04	7.3590e-05	1.8385e-05	4.5960e-06	1.1490e-06
	2.0678	2.0161	1.9962	2.0010	2.0001	2.0000	
1/256	1.3927e-02	3.2067e-03	7.8628e-04	1.9688e-04	4.9161e-05	1.2290e-05	3.0723e-06
	2.1187	2.0280	1.9977	2.0017	2.0000	2.0001	

Table 4.4: Maximum point-wise error and order of convergence corresponding to Example 4.1.2 with cubic spline.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256	512	1024
1	1.1769e-02	2.9589e-03	7.4076e-04	1.8525e-04	4.6318e-05	1.1580e-05	2.8949e-06
	1.9919	1.9980	1.9995	1.9999	2.0000	2.0000	
1/2	1.0999e-02	2.7647e-03	6.9209e-04	1.7308e-04	4.3274e-05	1.0819e-05	2.7047e-06
	1.9922	1.9981	1.9995	1.9999	2.0000	2.0000	
1/4	9.8720e-03	2.4802e-03	6.2079e-04	1.5525e-04	3.8814e-05	9.7037e-06	2.4259e-06
	1.9929	1.9982	1.9996	1.9999	2.0000	2.0000	
1/8	8.5011e-03	2.1339e-03	5.3400e-04	1.3353e-04	3.3385e-05	8.3464e-06	2.0866e-06
	1.9942	1.9986	1.9996	1.9999	2.0000	2.0000	
1/16	7.0988e-03	1.7791e-03	4.4506e-04	1.1128e-04	2.7821e-05	6.9554e-06	1.7389e-06
	1.9964	1.9991	1.9998	1.9999	2.0000	2.0000	
1/32	5.6878e-03	1.4224e-03	3.5563e-04	8.8909e-05	2.2227e-05	5.5569e-06	1.3892e-06
	1.9995	1.9999	2.0000	2.0000	2.0000	2.0000	
1/64	4.0736e-03	1.0170e-03	2.5415e-04	6.3533e-05	1.5883e-05	3.9707e-06	9.9267e-07
	2.0020	2.0005	2.0001	2.0000	2.0000	2.0000	
1/128	6.9756e-03	1.7515e-03	4.3394e-04	1.0853e-04	2.7117e-05	6.7785e-06	1.6946e-06
	1.9937	2.0131	1.9994	2.0008	2.0002	2.0000	
1/256	1.6386e-02	3.8017e-03	9.3366e-04	2.3239e-04	5.8034e-05	1.4504e-05	3.6261e-06
	2.1077	2.0257	2.0064	2.0016	2.0004	2.0000	

4.2 Neumann Boundary-Value Problem

Consider the Neumann BVP

$$\begin{cases} -\varepsilon u''(x) + q(x)u(x) = f(x), \ x \in (0,1), \\ -u'(0) = \eta_0, \quad u'(1) = \eta_1, \end{cases} \tag{4.11}$$

where $0 < \varepsilon \leq 1$, q and f are sufficiently smooth functions in $[0,1]$ and $q(x) > 0$ in $[0,1]$.

Let us consider a uniform mesh on the interval $I = [0,1]$ such that $0 = x_0 < x_1 < \cdots < x_N = 1$, where $x_i = ih$, $i = 0,1,\ldots,N$ and $h = 1/N$. Let U_i be the approximation of $u(x_i)$, U_i' be the approximation of $u'(x_i)$ and M_i be the approximation of $u''(x_i)$ for $i = 0,1,\ldots,N$. Let $q_i = q(x_i)$, $f_i = f(x_i)$, $i = 0,1,\ldots,N$.

From Section 4.1, it can be seen that, at the left end point $x = x_0$, $\Phi^{(1)}(x)$ should satisfy

$$6(1-h\alpha)\Phi^{(1)}(x_0) + 2(1-\alpha)hM_0 + hM_1 - \alpha hM_N = \frac{6}{h}[U_1 - U_0 - \alpha h^2(U_N - U_0)]. \tag{4.12}$$

At the right end point $x = x_N$, $\Phi^{(1)}(x)$ should satisfy

$$\begin{aligned} &-\alpha hM_0 + hM_{N-1} + 2(1-\alpha)hM_N - 6(1-h\alpha)\Phi^{(1)}(x_N) \\ &= -\frac{6}{h}[U_N - U_{N-1} - \alpha h^2(U_N - U_0)]. \end{aligned} \tag{4.13}$$

The differential equation given in (4.11) is discretized at x_i as $-\varepsilon M_i + q_iU_i = f_i$, $i = 0,1,\ldots,N$. The boundary conditions are discretized as $U_0' = -\eta_0$ and $U_N' = \eta_1$. We substitute

$$M_i = \frac{q_iU_i - f_i}{\varepsilon}, \ \Phi^{(1)}(x_0) = -\eta_0, \ \Phi^{(1)}(x_N) = \eta_1,$$

in (4.12), (4.5) and (4.13), and hence we get the following system:

$$\begin{aligned} &\Big[2(1-\alpha)h^2q_0 + 6\varepsilon - 6\varepsilon\alpha h^2\Big]U_0 + \Big[h^2q_1 - 6\varepsilon\Big]U_1 + \Big[-\alpha h^2q_N + 6\varepsilon\alpha h^2\Big]U_N \\ &= 6h\varepsilon(1-\alpha h)\eta_0 + 2(1-\alpha)h^2f_0 + h^2f_1 - \alpha h^2f_N, \end{aligned} \tag{4.14}$$

$$-\frac{3\alpha h^2q_0}{2}U_0 + \Big[\frac{h^2q_{i-1}}{2} - 3\varepsilon\Big]U_{i-1} + \Big[2h^2q_i + 6\varepsilon\Big]U_i + \Big[\frac{h^2q_{i+1}}{2} - 3\varepsilon\Big]U_{i+1}$$

$$- \frac{3\alpha h^2 q_N}{2} U_N = -3\alpha\varepsilon h^2 \eta_0 - 3\alpha\varepsilon h^2 \eta_1 - \frac{3\alpha h^2 f_0}{2} + \frac{h^2 f_{i-1}}{2} + 2h^2 f_i$$
$$+ \frac{h^2 f_{i+1}}{2} - \frac{3\alpha h^2 f_N}{2}, \ i = 1, 2, \ldots, N-1, \tag{4.15}$$

$$\Big[- \alpha h^2 q_0 + 6\varepsilon\alpha h^2 \Big] U_0 + \Big[h^2 q_{N-1} - 6\varepsilon \Big] U_{N-1} + \Big[2(1-\alpha) h^2 q_N + 6\varepsilon$$
$$- 6\varepsilon\alpha h^2 \Big] U_N = -\alpha h^2 f_0 + h^2 f_{N-1} + 2(1-\alpha) h^2 f_N + 6\varepsilon h (1 - \alpha h) \eta_1. \tag{4.16}$$

The system given in (4.14), (4.15) and (4.16) provides the approximations U_0, U_1, $\ldots$, U_N for the solution $u(x)$ at $x_0, x_1, \ldots, x_N$.

Remark 4.2.1. *If $\alpha = 0$, then the system given in (4.14), (4.15) and (4.16) will reduce to the cubic spline system as*

$$\Big[2h^2 q_0 + 6\varepsilon \Big] U_0 + \Big[h^2 q_1 - 6\varepsilon \Big] U_1 = 6h\varepsilon\eta_0 + 2h^2 f_0 + h^2 f_1,$$
$$\Big[\frac{h^2 q_{i-1}}{2} - 3\varepsilon \Big] U_{i-1} + \Big[2h^2 q_i + 6\varepsilon \Big] U_i + \Big[\frac{h^2 q_{i+1}}{2} - 3\varepsilon \Big] U_{i+1}$$
$$= \frac{h^2 f_{i-1}}{2} + 2h^2 f_i + \frac{h^2 f_{i+1}}{2}, \ i = 1, 2, \ldots, N-1,$$
$$\Big[h^2 q_{N-1} - 6\varepsilon \Big] U_{N-1} + \Big[2h^2 q_N + 6\varepsilon \Big] U_N = h^2 f_{N-1} + 2h^2 f_N + 6\varepsilon h \eta_1.$$

4.2.1 Convergence analysis

In this section, convergence analysis of the proposed scheme is carried out. Let $|\alpha| < h^2$. At the beginning, truncation error is derived. The truncation error $T_0(h)$ associated with the difference equation given in (4.14) is defined as

$$T_0(h) = \Big[2(1-\alpha) h^2 q_0 + 6\varepsilon - 6\varepsilon\alpha h^2 \Big] u(x_0) + \Big[h^2 q_1 - 6\varepsilon \Big] u(x_1)$$
$$+ \Big[- \alpha h^2 q_N + 6\varepsilon\alpha h^2 \Big] u(x_N) - \Big[- 6h\varepsilon(1 - \alpha h) u'(x_0)$$
$$+ 2(1-\alpha) h^2 f_0 + h^2 f_1 - \alpha h^2 f_N \Big]. \tag{4.17}$$

Substitute $f_i = -\varepsilon u''(x_i) + q_i u(x_i)$ in (4.17), we get

$$T_0(h) = - 6\varepsilon\alpha h^2 u(x_0) + 6\varepsilon\alpha h^2 u(x_N) - 6\varepsilon\alpha h^2 u'(x_0) - 2\varepsilon\alpha h^2 u''(x_0) - \varepsilon\alpha h^2 u''(x_N)$$
$$+ 6\varepsilon u(x_0) - 6\varepsilon u(x_1) + 6h\varepsilon u'(x_0) + 2h^2 \varepsilon u''(x_0) + h^2 \varepsilon u''(x_1). \tag{4.18}$$

Now, using the Taylor expansions of $u(x_1)$ and $u''(x_1)$ about the point $x = x_0$ in (4.18) and after simplifying we get

$$\begin{aligned} T_0(h) = &-6\varepsilon\alpha h^2 u(x_0) + 6\varepsilon\alpha h^2 u(x_N) - 6\varepsilon\alpha h^2 u'(x_0) \\ &- 2\varepsilon\alpha h^2 u''(x_0) - \varepsilon\alpha h^2 u''(x_N) + \frac{\varepsilon h^4}{4} u^{(4)}(x_0) + \mathcal{O}(h^5). \end{aligned}$$

When $h \to 0$, we get $|T_0(h)| \leq K_0 h^4$, where K_0 is a constant. Hence we get $T_0(h) = \mathcal{O}(h^4)$ as $h \to 0$. The truncation error $T_i(h)$, $i = 1, 2, \ldots, N-1$ associated with the difference equations given in (4.15) are defined as

$$\begin{aligned} T_i(h) = &-\frac{3\alpha h^2 q_0}{2} u(x_0) + \Big[\frac{h^2 q_{i-1}}{2} - 3\varepsilon\Big] u(x_{i-1}) + \Big[2h^2 q_i + 6\varepsilon\Big] u(x_i) \\ &+ \Big[\frac{h^2 q_{i+1}}{2} - 3\varepsilon\Big] u(x_{i+1}) - \frac{3\alpha h^2 q_N}{2} u(x_N) - \Big[3\alpha\varepsilon h^2 u'(x_0) - 3\alpha\varepsilon h^2 u'(x_N) \\ &- \frac{3\alpha h^2 f_0}{2} + \frac{h^2 f_{i-1}}{2} + 2h^2 f_i + \frac{h^2 f_{i+1}}{2} - \frac{3\alpha h^2 f_N}{2}\Big]. \end{aligned} \tag{4.19}$$

Now substitute $f_i = -\varepsilon u''(x_i) + q_i u(x_i)$ in (4.19), we get

$$\begin{aligned} T_i(h) = &-\frac{3\alpha\varepsilon h^2}{2} u''(x_0) - \frac{3\alpha\varepsilon h^2}{2} u''(x_N) - 3\alpha\varepsilon h^2 u'(x_0) + 3\alpha\varepsilon h^2 u'(x_N) - 3\varepsilon u(x_{i-1}) \\ &+ 6\varepsilon u(x_i) - 3\varepsilon u(x_{i+1}) + \frac{h^2\varepsilon}{2} u''(x_{i-1}) + 2h^2\varepsilon u''(x_i) + \frac{h^2\varepsilon}{2} u''(x_{i+1}). \end{aligned} \tag{4.20}$$

Using the Taylor expansions of $u(x_{i-1})$, $u(x_{i+1})$, $u''(x_{i-1})$ and $u''(x_{i+1})$ about the point $x = x_i$ in (4.20) and after simplifying we get

$$\begin{aligned} T_i(h) = &-\frac{3\alpha\varepsilon h^2}{2} u''(x_0) - \frac{3\alpha\varepsilon h^2}{2} u''(x_N) - 3\alpha\varepsilon h^2 u'(x_0) \\ &+ 3\alpha\varepsilon h^2 u'(x_N) + \frac{\varepsilon h^4}{4} u^{(4)}(x_0) + \mathcal{O}(h^5). \end{aligned}$$

It can be seen that, when $h \to 0$, we get $|T_i(h)| \leq K_i h^4$, where K_i is a constant. Thus $T_i(h) = \mathcal{O}(h^4)$ as $h \to 0$ for $i = 1, 2, \ldots, N-1$. Using similar procedure followed to compute $T_0(h)$, we get the truncation error associated with the difference equation given in (4.16) as

$$\begin{aligned} T_N(h) = &\,6\varepsilon\alpha h^2 u(x_0) - 6\varepsilon\alpha h^2 u(x_N) + 6\varepsilon\alpha h^2 u'(x_N) \\ &- 2\varepsilon\alpha h^2 u''(x_N) - \varepsilon\alpha h^2 u''(x_0) + \frac{\varepsilon h^4}{4} u^{(4)}(x_N) + \mathcal{O}(h^5). \end{aligned}$$

It can be observed that, when $h \to 0$, we get $|T_N(h)| \leq K_N h^4$, where K_N is a constant. Thus $T_N(h) = \mathcal{O}(h^4)$ as $h \to 0$.

The system given in (4.14), (4.15) and (4.16) can be written in the matrix form $AU = d$, where

$$A = \begin{bmatrix} A_{0,0} & A_{0,1} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{0,N} \\ A_{1,0} & A_{1,1} & A_{1,2} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{1,N} \\ A_{2,0} & A_{2,1} & A_{2,2} & A_{2,3} & 0 & \dots & 0 & 0 & 0 & 0 & A_{2,N} \\ A_{3,0} & 0 & A_{3,2} & A_{3,3} & A_{3,4} & \dots & 0 & 0 & 0 & 0 & A_{3,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{N-3,0} & 0 & 0 & 0 & 0 & \dots & A_{N-3,N-4} & A_{N-3,N-3} & A_{N-3,N-2} & 0 & A_{N-3,N} \\ A_{N-2,0} & 0 & 0 & 0 & 0 & \dots & 0 & A_{N-2,N-3} & A_{N-2,N-2} & A_{N-2,N-1} & A_{N-2,N} \\ A_{N-1,0} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & A_{N-1,N-2} & A_{N-1,N-1} & A_{N-1,N} \\ A_{N,0} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & A_{N,N-1} & A_{N,N} \end{bmatrix},$$

$A_{i,j}$ is the coefficient of U_j, $U = (U_0, U_1, \dots, U_N)^T$, $d = (d_0, d_1, \dots, d_N)^T$, d_i, $i = 0, 1, \dots, N$ are the right hand side of the system.

In the following proposition, h and scaling factor α are restricted suitably so that the matrix A would be strictly diagonally dominant. Let $q_* = \min\{q(x) : x \in [0,1]\}$ and $q^* = \max\{q(x) : x \in [0,1]\}$.

Proposition 4.2.1. *Let A be the matrix corresponding to the system given in (4.14), (4.15) and (4.16). Let h satisfying the condition $h^2 q^* < 6\varepsilon$. The matrix A is strictly diagonally dominant if the scaling factor satisfies the following conditions:*

$$0 \leq \alpha < \min\left\{h^2, \frac{q_*}{12\varepsilon}, \frac{q_*}{3q^*}\right\}.$$

Proof. Let $0 \leq \alpha < h^2$. In order to prove $|A_{0,0}| - \sum_{j=1}^{N} |A_{0,j}| > 0$, we consider two cases. Let $q_N < 6\varepsilon$. The entries $A_{0,0} = 2(1-\alpha)h^2 q_0 + 6\varepsilon(1-\alpha h^2) > 0$, $A_{0,1} = h^2 q_1 - 6\varepsilon < 0$ by the assumption on h, $A_{0,N} = \alpha h^2[6\varepsilon - q_N] \geq 0$. Then

$$\begin{aligned} |A_{0,0}| - \sum_{j=1}^{N} |A_{0,j}| &= 2(1-\alpha)h^2 q_0 + 6\varepsilon(1-\alpha h^2) + h^2 q_1 - 6\varepsilon + h^2\alpha[q_N - 6\varepsilon] \\ &= h^2[2(1-\alpha)q_0 + \alpha q_N] + h^2[q_1 - 12\varepsilon\alpha]. \end{aligned}$$

Hence, the expression $|A_{0,0}| - \sum_{j=1}^{N} |A_{0,j}| > 0$ if $\alpha < \frac{q_*}{12\varepsilon}$.

Let $q_N \geq 6\varepsilon$.

The entries $A_{0,0} = 2(1-\alpha)h^2 q_0 + 6\varepsilon(1-\alpha h^2) > 0$, $A_{0,1} = h^2 q_1 - 6\varepsilon < 0$ by the assumption on h, $A_{0,N} = \alpha h^2[6\varepsilon - q_N] \leq 0$. We have

$$\begin{aligned}|A_{0,0}| - \sum_{j=1}^{N} |A_{0,j}| &= 2(1-\alpha)h^2 q_0 + 6\varepsilon(1-\alpha h^2) + h^2 q_1 - 6\varepsilon + h^2\alpha[6\varepsilon - q_N] \\ &= 2(1-\alpha_0)h^2 q_0 + h^2[q_1 - \alpha q_N].\end{aligned}$$

It is easy to see that $|A_{0,0}| - \sum_{j=1}^{N} |A_{0,j}| > 0$ if $\alpha < \frac{q_*}{q^*}$.

Thus for $q(x) > 0$, we obtain $|A_{0,0}| - \sum_{j=1}^{N} |A_{0,j}| > 0$ if $h^2 q^* < 6\varepsilon$ and $\alpha < \min\{\frac{q_*}{12\varepsilon}, \frac{q_*}{q^*}\}$.

The entry $A_{1,0} = -\frac{3\alpha h^2 q_0}{2} + \left[\frac{h^2 q_0}{2} - 3\varepsilon\right] < 0$ by the assumption on h. One can see that $A_{1,1} = 2h^2 q_1 + 6\varepsilon > 0$, $A_{1,2} = \frac{h^2 q_2}{2} - 3\varepsilon < 0$ by assumption on h, $A_{1,N} = -\frac{3\alpha h^2 q_N}{2} \leq 0$. We obtain

$$\begin{aligned}&|A_{1,1}| - \sum_{j=0,\, j\neq 1}^{N} |A_{1,j}| \\ &= 2h^2 q_1 + 6\varepsilon - \frac{3\alpha h^2 q_0}{2} + \frac{h^2 q_0}{2} - 3\varepsilon + \frac{h^2 q_2}{2} - 3\varepsilon - \frac{3\alpha h^2 q_N}{2} \\ &= 2h^2 q_1 + \frac{h^2}{2}\Big[q_0 - 3\alpha q_0\Big] + \frac{h^2}{2}\Big[q_2 - 3\alpha q_N\Big].\end{aligned}$$

Hence the quantity $|A_{1,1}| - \sum_{j=0,\, j\neq 1}^{N} |A_{1,j}| > 0$ if $\alpha < \frac{q_*}{3q^*}$.

For $i = 2, 3, \ldots, N-2$, the entries $A_{i,i} = 2h^2 q_i + 6\varepsilon > 0$, $A_{i,i-1} = \frac{h^2 q_{i-1}}{2} - 3\varepsilon < 0$, $A_{i,i+1} = \frac{h^2 q_{i+1}}{2} - 3\varepsilon < 0$, $A_{i,0} = -\frac{3\alpha h^2 q_0}{2} \leq 0$, $A_{i,N} = -\frac{3\alpha h^2 q_N}{2} \leq 0$.

We can see that

$$\begin{aligned}&|A_{i,i}| - \sum_{j=0,\, j\neq i}^{N} |A_{i,j}| \\ &= 2h^2 q_i + 6\varepsilon - \frac{3\alpha h^2 q_0}{2} + \frac{h^2 q_{i-1}}{2} - 3\varepsilon + \frac{h^2 q_{i+1}}{2} - 3\varepsilon - \frac{3\alpha h^2 q_N}{2} \\ &= 2h^2 q_i + \frac{h^2}{2}\Big[q_{i-1} - 3\alpha q_0\Big] + \frac{h^2}{2}\Big[q_{i+1} - 3\alpha q_N\Big].\end{aligned}$$

Consequently, the quantity $|A_{i,i}| - \sum_{j=0,\, j\neq i}^{N} |A_{i,j}| > 0$ if $\alpha < \frac{q_*}{3q^*}$.

The entries $A_{N-1,0} = -\frac{3\alpha h^2 q_0}{2} \leq 0$, $A_{N-1,N-2} = \frac{h^2 q_{N-2}}{2} - 3\varepsilon < 0$, $A_{N-1,N-1} = 2h^2 q_{N-1} + 6\varepsilon > 0$, $A_{N-1,N} = \frac{h^2 q_N}{2} - 3\varepsilon - \frac{3\alpha h^2 q_N}{2} < 0$ by the assumption on h. Thus, we obtain

$$\begin{aligned}
&|A_{N-1,N-1}| - \sum_{j=0,\, j\neq N-1}^{N} |A_{N-1,j}| \\
&= 2h^2 q_{N-1} + 6\varepsilon - \frac{3\alpha h^2 q_0}{2} + \frac{h^2 q_{N-2}}{2} - 3\varepsilon + \frac{h^2 q_N}{2} - 3\varepsilon - \frac{3\alpha h^2 q_N}{2} \\
&= 2h^2 q_{N-1} + \frac{h^2}{2}\Big[q_{N-2} - 3\alpha q_0\Big] + \frac{h^2}{2}\Big[q_N - 3\alpha q_N\Big] > 0 \quad \text{if } \alpha < \frac{q_*}{3q^*}.
\end{aligned}$$

In order to prove $|A_{N,N}| - \sum_{j=0}^{N-1} |A_{N,j}| > 0$, we consider two cases.

Let $q_0 < 6\varepsilon$. The entries $A_{N,0} = \alpha h^2[6\varepsilon - q_0] \geq 0$, $A_{N,N-1} = h^2 q_{N-1} - 6\varepsilon < 0$, $A_{N,N} = 2h^2(1-\alpha)q_N + 6\varepsilon(1-\alpha h^2) > 0$. We get

$$\begin{aligned}
&|A_{N,N}| - \sum_{j=0}^{N-1} |A_{N,j}| \\
&= 2h^2(1-\alpha)q_N + 6\varepsilon(1-\alpha h^2) + \alpha h^2[q_0 - 6\varepsilon] + h^2 q_{N-1} - 6\varepsilon \\
&= h^2[2(1-\alpha)q_N + \alpha q_0] + h^2[q_{N-1} - 12\varepsilon\alpha] > 0 \text{ if } \alpha < \frac{q_*}{12\varepsilon}.
\end{aligned}$$

Let $q_0 \geq 6\varepsilon$. The entries $A_{N,0} = \alpha h^2[6\varepsilon - q_0] \leq 0$, $A_{N,N-1} = h^2 q_{N-1} - 6\varepsilon < 0$, $A_{N,N} = 2h^2(1-\alpha)q_N + 6\varepsilon(1-\alpha h^2) > 0$. We obtain

$$\begin{aligned}
&|A_{N,N}| - \sum_{j=0}^{N-1} |A_{N,j}| \\
&= 2h^2(1-\alpha)q_N + 6\varepsilon(1-\alpha h^2) - \alpha h^2 q_0 + 6h^2\varepsilon\alpha + h^2 q_{N-1} - 6\varepsilon \\
&= 2h^2(1-\alpha)q_N + h^2[q_{N-1} - \alpha q_0] > 0 \text{ if } \alpha < \frac{q_*}{q^*}.
\end{aligned}$$

Thus for $q(x) > 0$, we get $|A_{N,N}| - \sum_{j=0}^{N-1} |A_{N,j}| > 0$ if $h^2 q^* < 6\varepsilon$ and $\alpha < \min\left\{\frac{q_*}{12\varepsilon}, \frac{q_*}{q^*}\right\}$.
Thus, the matrix A is strictly diagonally dominant when h and α satisfy the conditions given in statement of the Proposition 4.2.1. □

Theorem 4.2.1. *Let $u(x)$ be the exact solution of the differential equation given in (4.11). Let U be the approximate solution obtained by the system given in (4.14), (4.15)* *and (4.16). If we choose h and α according to Proposition 4.2.1, then $\|E\|_\infty = \mathcal{O}(h^2)$, where $E = (E_0, E_1, \ldots, E_N)^T$, $E_i = u(x_i) - U_i$.*

Proof. The system (4.14), (4.15) and (4.16) with exact solution can be written as

$$A\bar{u} = d + T(h),$$

where $T(h) = (T_0(h), T_1(h), \ldots, T_N(h))^T$, $\bar{u} = (u(x_0), u(x_1), \ldots, u(x_N))^T$. Since $AU = d$, we get

$$AE = T(h)$$

By the assumption on h and α given in Proposition 4.2.1, we have

$$E = A^{-1}T(h).$$

We get

$$E_j = \sum_{i=0}^{N} A^{-1}_{j,i} T_i(h), \; j = 0, 1, \ldots, N.$$

Using the theory of matrices, we have

$$\sum_{i=0}^{N} A^{-1}_{k,i} \mathcal{S}_i = 1, \; k = 0, 1, \ldots, N,$$

where $\mathcal{S}_i = \sum_{j=0}^{N} A_{i,j}$, $i = 0, 1, \ldots, N$. The row sums

$$\mathcal{S}_0 = 2h^2 q_0(1-\alpha) + h^2[q_1 - \alpha q_N] > 0 \text{ if } \alpha < \frac{q_*}{q^*},$$

$$\mathcal{S}_i = 2h^2 q_i + \frac{h^2}{2}\Big[q_{i-1} - 3\alpha q_0\Big] + \frac{h^2}{2}\Big[q_{i+1} - 3\alpha q_N\Big] > 0 \text{ if } \alpha < \frac{q_*}{3q^*},$$

$i = 1, 2, \ldots, N-1,$

$$\mathcal{S}_N = 2h^2 q_N(1-\alpha) + h^2[q_{N-1} - \alpha q_0] > 0 \text{ if } \alpha < \frac{q_*}{q^*}.$$

It can be observed that, the conditions given in statement of Proposition 4.2.1 ensures the row sums of the matrix A are positive. We have

$$\sum_{i=0}^{N} A^{-1}_{k,i} \leq \frac{1}{\min_{0\leq i\leq N} \mathcal{S}_i} = \frac{1}{C_{i_0} h^2},$$

where C_{i_0} is a constant. We obtain

$$\Big|E_j\Big| \leq \frac{Kh^4}{C_{i_0} h^2},$$

where K is a constant. Hence we get $\|E\|_\infty = \mathcal{O}(h^2)$ as $h \to 0$. □

Remark 4.2.2. *Let us consider the restriction of* α *given in Proposition 4.2.1,*

$$0 \leq \alpha < \min\left\{h^2,\ \frac{q_*}{12\varepsilon},\ \frac{q_*}{3q^*}\right\}.$$

When h *is sufficiently small,* h^2 *will be the minimum in the above range. For fixed* h*, by varying* α*, we can compute the numerical solutions.*

4.2.2 Numerical examples

In this section, the presented numerical method is tested on the following singularly perturbed two-point BVP. If the exact solution $u(x)$ is unknown, then the following double mesh method is used to compute the maximum point-wise errors:

For fixed ε, let U^N and U^{2N} be the numerical solutions of N and $2N$ numbers of mesh intervals respectively. Define the maximum point-wise error by

$$E^N = \max_{0\leq i\leq N} |U_i^N - U_{2i}^{2N}|.$$

Example 4.2.1. *Consider the following BVP [98]:*

$$\begin{cases} -\varepsilon u''(x) + [1 + x^2 + \cos x]u(x) = x^{4.5} + \sin x,\ x \in (0, 1), \\ u'(0) = 0, \quad u'(1) = 0. \end{cases}$$

The exact solution is unknown and hence we have used the double mesh method to compute the numerical solutions. For each fixed $\varepsilon = \frac{1}{2^{i-1}}$, $i = 1, 2, \ldots, 9$, we take $N = 16, 32, 64, 128, 256, 512, 1024$ and we obtain the α restriction in the range $0 \leq \alpha < h^2$. For each fixed ε, we taken α as $\alpha = 0.9999h^2$. The maximum point-wise errors and the order of convergence are tabulated in Table 4.5. Figure 4.3 represents Log-log plots of the maximum errors for the Example 4.2.1. Table 4.6 represents the maximum point-wise errors and the order of convergence corresponding to the cubic spline method.

Table 4.5: Maximum point-wise error and order of convergence corresponding to Example 4.2.1.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256	512	1024
1	2.3236e-04	5.8471e-05	1.4657e-05	3.6685e-06	9.1763e-07	2.2947e-07	5.7375e-08
	1.9905	1.9962	1.9983	1.9992	1.9996	1.9998	
1/2	3.7014e-04	9.3286e-05	2.3397e-05	5.8575e-06	1.4653e-06	3.6644e-07	9.1624e-08
	1.9883	1.9953	1.9980	1.9991	1.9996	1.9998	
1/4	5.9504e-04	1.5021e-04	3.7693e-05	9.4382e-06	2.3613e-06	5.9052e-07	1.4765e-07
	1.9860	1.9946	1.9977	1.9990	1.9995	1.9998	
1/8	9.1526e-04	2.3160e-04	5.8157e-05	1.4565e-05	3.6441e-06	9.1137e-07	2.2788e-07
	1.9825	1.9936	1.9974	1.9989	1.9995	1.9997	
1/16	1.3289e-03	3.3785e-04	8.4934e-05	2.1277e-05	5.3238e-06	1.3315e-06	3.3292e-07
	1.9758	1.9920	1.9970	1.9988	1.9995	1.9997	
1/32	1.8560e-03	4.7622e-04	1.1999e-04	3.0075e-05	7.5260e-06	1.8822e-06	4.7064e-07
	1.9625	1.9887	1.9963	1.9986	1.9994	1.9997	
1/64	2.5368e-03	6.6280e-04	1.6776e-04	4.2093e-05	1.0536e-05	2.6351e-06	6.5889e-07
	1.9364	1.9822	1.9947	1.9983	1.9994	1.9997	
1/128	3.3888e-03	9.1671e-04	2.3414e-04	5.8880e-05	1.4745e-05	3.6882e-06	9.2222e-07
	1.8862	1.9691	1.9915	1.9976	1.9992	1.9997	
1/256	4.3492e-03	1.2548e-03	3.2629e-04	8.2426e-05	2.0664e-05	5.1701e-06	1.2928e-06
	1.7934	1.9432	1.9850	1.9960	1.9989	1.9997	

Table 4.6: Maximum point-wise error and order of convergence corresponding to Example 4.2.1 with cubic spline.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256	512	1024
1	5.4465e-04	1.3644e-04	3.4127e-05	8.5330e-06	2.1333e-06	5.3333e-07	1.3333e-07
	1.9971	1.9993	1.9998	2.0000	2.0000	2.0000	
1/2	6.6380e-04	1.6643e-04	4.1636e-05	1.0411e-05	2.6028e-06	6.5072e-07	1.6268e-07
	1.9959	1.9990	1.9997	1.9999	2.0000	2.0000	
1/4	8.6228e-04	2.1647e-04	5.4174e-05	1.3547e-05	3.3870e-06	8.4677e-07	2.1169e-07
	1.9940	1.9985	1.9996	1.9999	2.0000	2.0000	
1/8	1.1497e-03	2.8932e-04	7.2449e-05	1.8120e-05	4.5304e-06	1.1326e-06	2.8316e-07
	1.9905	1.9976	1.9994	1.9999	2.0000	2.0000	
1/16	1.5272e-03	3.8618e-04	9.6821e-05	2.4223e-05	6.0567e-06	1.5143e-06	3.7857e-07
	1.9836	1.9959	1.9990	1.9997	1.9999	2.0000	
1/32	2.0184e-03	5.1533e-04	1.2952e-04	3.2423e-05	8.1085e-06	2.0273e-06	5.0683e-07
	1.9697	1.9923	1.9981	1.9995	1.9999	2.0000	
1/64	2.6667e-03	6.9369e-04	1.7519e-04	4.3910e-05	1.0985e-05	2.7466e-06	6.8668e-07
	1.9427	1.9853	1.9963	1.9991	1.9998	1.9999	
1/128	3.4903e-03	9.4074e-04	2.3984e-04	6.0259e-05	1.5084e-05	3.7721e-06	9.4309e-07
	1.8915	1.9717	1.9928	1.9982	1.9996	1.9999	
1/256	4.4262e-03	1.2733e-03	3.3063e-04	8.3461e-05	2.0916e-05	5.2322e-06	1.3083e-06
	1.7975	1.9453	1.9860	1.9965	1.9991	1.9998	

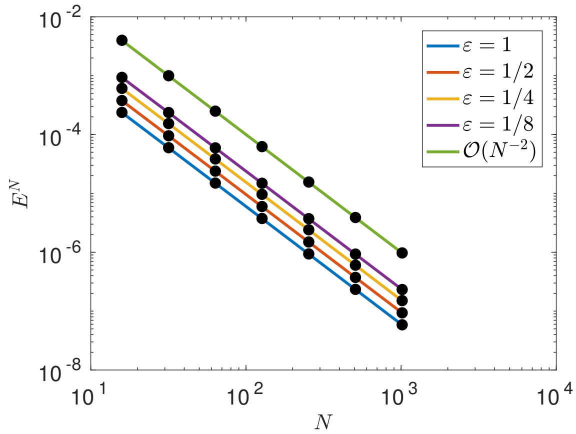


Figure 4.3: Log-log plot of Example 4.2.1.

4.3 Conclusion

We have described and demonstrated the applicability of the fractal cubic spline method for solving self-adjoint singularly perturbed BVPs. Proposed method is accurate, and easy to implement on computer. To implement the developed method, we considered model problems and have tabulated the numerical results. Also, numerical results are compared with the cubic spline method results. Proposed method has second-order convergent. The important feature is that the developed method has second order convergence for Neumann BVPs. It will be great interest if we extend the proposed methods for the functions f which are discontinuous.

Chapter 5

Fractal Cubic Spline Solutions for Nonself-Adjoint Boundary-Value Problems

In this chapter, we consider the nonself-adjoint BVPs and numerical solutions of these BVPs are obtained. At beginning, we consider the singularly perturbed nonself-adjoint BVPs. These problems arising in transport phenomena in chemistry and biology [22,23, 87]. Aziz and Khan [9] used spline in compression to solve the nonself-adjoint singularly perturbed BVPs and this method has quadratic convergence. Kadalbajoo and Bawa [74] proposed second-order convergent method using cubic spline with variable mesh to solve the nonself-adjoint singularly perturbed BVPs. Lin et al. [87] used B-spline method to get the numerical solutions of singularly perturbed BVPs. In this chapter, we have used fractal cubic spline method to solve the nonself-adjoint singularly perturbed BVPs and the proposed method has second-order convergence.

Next, we consider the nonself-adjoint singular BVPs. Many problems in applied mathematics leads to the singular BVPs. These BVPs arise in the study of generalized axially symmetric potentials after separation of variables has been employed. These problems also occur very frequently in the study of electrohydrodynamics and the theory of thermal explosions [117]. To solve the singular BVPs, various authors proposed different methods. For example, Chawla and Katti [40] used three-point finite difference method to solve the singular BVPs. To solve singular BVPs, Ravi Kanth and Reddy [117] developed a numerical method using cubic spline. Abukhaled et al. [2] used cubic B-splines and Rashidinia et al. [110] used spline in compression to solve the singular BVPs. Also, one can see that in References [41, 65, 73, 77, 97, 112, 113, 115, 116], various numerical methods have been used to solve the singular BVPs. Many researchers attempted to solve the singular BVPs by considering the series expansion procedures in the neighborhood of the singularity and then solve the regular BVP in rest of the interval using any numerical method. In this chapter, we discuss a direct method based on fractal cubic splines for a class of singular two-point BVPs.

In Sections 5.1, 5.2 and 5.3, numerical methods have been developed for solving singularly perturbed BVPs, linear singular BVPs and non-linear singular BVPs respectively. Continuity conditions of the fractal cubic spline are taken as the discretization equations to solve these BVPs. Convergence analysis of the proposed methods were developed. In order to test the proposed method, numerical results are tabulated and compared with the numerical solutions corresponding to the cubic spline method.

5.1 Singularly Perturbed Nonself-Adjoint Boundary-Value Problems

Let us consider the singularly perturbed two-point BVPs of the form

$$\begin{cases} \varepsilon u''(x) = p(x)u'(x) + q(x)u(x) + f(x), \; x \in (0,1), \\ u(0) = \eta_0, \quad u(1) = \eta_1, \end{cases} \tag{5.1}$$

where ε is the perturbation parameter such that $0 < \varepsilon \leq 1$, functions p, q, f are sufficiently smooth in $[0,1]$, $p(x)$ is either strictly positive ($p(x) > 0$ for all $x \in [0,1]$) or strictly negative ($p(x) < 0$ for all $x \in [0,1]$) and $q(x) > 0$ for all $x \in [0,1]$. Solutions of the BVPs exhibit boundary layer at $x = 0$ or $x = 1$ if $p(x)$ is strictly negative or $p(x)$ is strictly positive respectively [7].

Let us consider a uniform mesh on the interval $I = [0,1]$ such that $0 = x_0 < x_1 <$

$\cdots < x_N = 1$, where $x_i = ih$, $i = 0, 1, \ldots, N$ and $h = 1/N$. Let U_i be the approximation of $u(x_i)$, U_i' be the approximation of $u'(x_i)$ and M_i be the approximation of $u''(x_i)$.

The differential equation given in (5.1) is discretized at $x = x_i$ as $\varepsilon M_i = p_i U_i' + q_i U_i + f_i$, where $p_i = p(x_i)$, $q_i = q(x_i)$, $f_i = f(x_i)$. We consider the following approximations:

$$\Phi'(x_0) = \frac{-U_2 + 4U_1 - 3U_0}{2h},\ \Phi'(x_N) = \frac{3U_N - 4U_{N-1} + U_{N-2}}{2h},\ U_i' = \frac{U_{i+1} - U_{i-1}}{2h},$$
$$U_{i+1}' = \frac{3U_{i+1} - 4U_i + U_{i-1}}{2h},\ U_{i-1}' = \frac{-U_{i+1} + 4U_i - 3U_{i-1}}{2h}.$$

Substituting M_i and the above approximations in (4.5), we obtain the following system of equations:

$$\begin{aligned}
&2h\Big[-\frac{3\alpha p_0}{2} - 3\alpha\varepsilon\Big]U_1 - \frac{h}{2}\Big[-\frac{3\alpha p_0}{2} - 3\alpha\varepsilon\Big]U_2 + \Big[-3\varepsilon \\
&+ \frac{h^2 q_{i-1}}{2} - \frac{3hp_{i-1}}{4} - hp_i + \frac{hp_{i+1}}{4}\Big]U_{i-1} + \Big[6\varepsilon + 2h^2 q_i \\
&+ hp_{i-1} - hp_{i+1}\Big]U_i + \Big[-3\varepsilon + \frac{h^2 q_{i+1}}{2} - \frac{hp_{i-1}}{4} + hp_i \\
&+ \frac{3hp_{i+1}}{4}\Big]U_{i+1} + \frac{h}{2}\Big[-\frac{3\alpha p_N}{2} + 3\alpha\varepsilon\Big]U_{N-2} - 2h\Big[-\frac{3\alpha p_N}{2} \\
&+ 3\alpha\varepsilon\Big]U_{N-1} = \frac{3h}{2}\Big[-\frac{3\alpha p_0}{2} - 3\alpha\varepsilon\Big]\eta_0 + \frac{3h^2\alpha q_0\eta_0}{2} + \frac{3h^2\alpha f_0}{2} \\
&- \frac{h^2 f_{i-1}}{2} - 2h^2 f_i - \frac{h^2 f_{i+1}}{2} - \frac{3h}{2}\Big[-\frac{3\alpha p_N}{2} + 3\alpha\varepsilon\Big]\eta_1 \\
&+ \frac{3h^2\alpha q_N\eta_1}{2} + \frac{3h^2\alpha f_N}{2}, \quad i = 1, 2, \ldots, N-1,
\end{aligned} \tag{5.2}$$

where $U_0 = \eta_0$ and $U_N = \eta_1$ are the discretizations of the Dirichlet boundary conditions. The system (5.2) gives the approximations $U_1, U_2, \ldots, U_{N-1}$ of the solution $u(x)$ at $x_1, x_2, \ldots, x_{N-1}$.

Remark 5.1.1. *If $\alpha = 0$, then the system (5.2) will reduce to the system corresponding to the cubic spline:*

$$\begin{aligned}
&\Big[-3\varepsilon + \frac{h^2 q_{i-1}}{2} - \frac{3hp_{i-1}}{4} - hp_i + \frac{hp_{i+1}}{4}\Big]U_{i-1} + \Big[6\varepsilon + 2h^2 q_i \\
&+ hp_{i-1} - hp_{i+1}\Big]U_i + \Big[-3\varepsilon + \frac{h^2 q_{i+1}}{2} - \frac{hp_{i-1}}{4} + hp_i \\
&+ \frac{3hp_{i+1}}{4}\Big]U_{i+1} = -\frac{h^2 f_{i-1}}{2} - 2h^2 f_i - \frac{h^2 f_{i+1}}{2},\ i = 1, 2, \ldots, N-1.
\end{aligned}$$

5.1.1 Error analysis

In this section, convergence analysis of the proposed numerical scheme is discussed. Let $|\alpha|< h^2$. Assume that the exact solution $u(x)$ of (5.1) is sufficiently differentiable. Substitute the original function value $u(x_i)$ instead of the approximate value U_i in (5.2) and hence we get the truncation error $T_i(h)$, $i = 1, 2, \dots, N-1$ of the equations given in (5.2) as

$$\begin{aligned}T_i(h) = {} & [r_i^0 u(x_0) + r_i^1 u(x_1) + r_i^2 u(x_2) + r_i^- u(x_{i-1}) + r_i^c u(x_i) + r_i^+ u(x_{i+1}) \\ & + r_i^{N-2} u(x_{N-2}) + r_i^{N-1} u(x_{N-1}) + r_i^N u(x_N)] - [-q_i^0 f_0 - q_i^- f_{i-1} \\ & - q_i^c f_i - q_i^+ f_{i+1} - q_i^N f_N],\end{aligned} \tag{5.3}$$

where

$$\begin{aligned}& r_i^0 = \frac{3h}{2}\Big[\frac{3\alpha p_0}{2} + 3\alpha\varepsilon\Big] - \frac{3h^2\alpha q_0}{2},\ r_i^1 = 2h\Big[-\frac{3\alpha p_0}{2} - 3\alpha\varepsilon\Big], \\ & r_i^2 = -\frac{h}{2}\Big[-\frac{3\alpha p_0}{2} - 3\alpha\varepsilon\Big],\ r_i^- = -3\varepsilon + \frac{h^2 q_{i-1}}{2} - \frac{3hp_{i-1}}{4} - hp_i \\ & + \frac{hp_{i+1}}{4},\ r_i^c = 6\varepsilon + 2h^2 q_i + hp_{i-1} - hp_{i+1},\ r_i^+ = -3\varepsilon + \frac{h^2 q_{i+1}}{2} \\ & - \frac{hp_{i-1}}{4} + hp_i + \frac{3hp_{i+1}}{4},\ r_i^{N-2} = \frac{h}{2}\Big[-\frac{3\alpha p_N}{2} + 3\alpha\varepsilon\Big], \\ & r_i^{N-1} = -2h\Big[-\frac{3\alpha p_N}{2} + 3\alpha\varepsilon\Big],\ r_i^N = \frac{3h}{2}\Big[-\frac{3\alpha p_N}{2} + 3\alpha\varepsilon\Big] - \frac{3h^2\alpha q_N}{2}, \\ & q_i^0 = -\frac{3\alpha h^2}{2},\ q_i^- = \frac{h^2}{2},\ q_i^c = 2h^2,\ q_i^+ = \frac{h^2}{2},\ q_i^N = -\frac{3\alpha h^2}{2}.\end{aligned}$$

Using the differential equation (5.1) in (5.3), we get

$$\begin{aligned}T_i(h) = {} & \Big[r_i^0 u(x_0) + r_i^1 u(x_1) + r_i^2 u(x_2) + r_i^- u(x_{i-1}) + r_i^c u(x_i) + r_i^+ u(x_{i+1}) \\ & + r_i^{N-2} u(x_{N-2}) + r_i^{N-1} u(x_{N-1}) + r_i^N u(x_N)\Big] - \Big[q_i^0(-\varepsilon u''(x_0) \\ & + p_0 u'(x_0) + q_0 u(x_0)) + q_i^-(-\varepsilon u''(x_{i-1}) + p_{i-1} u'(x_{i-1}) \\ & + q_{i-1} u(x_{i-1})) + q_i^c(-\varepsilon u''(x_i) + p_i u'(x_i) + q_i u(x_i)) \\ & + q_i^+(-\varepsilon u''(x_{i+1}) + p_{i+1} u'(x_{i+1}) + q_{i+1} u(x_{i+1})) \\ & + q_i^N(-\varepsilon u''(x_N) + p_N u'(x_N) + q_N u(x_N))\Big].\end{aligned} \tag{5.4}$$

After simplifying (5.4), we obtain

$$\begin{aligned}
T_i(h) =& -3\alpha h^2\varepsilon\Big[\frac{-3u(x_0)+4u(x_1)-u(x_2)}{2h}\Big] - \frac{3\alpha h^2\varepsilon}{2}u''(x_0)\\
&+3\alpha\varepsilon h^2\Big[\frac{3u(x_N)-4u(x_{N-1})+u(x_{N-2})}{2h}\Big] - \frac{3\alpha h^2\varepsilon}{2}u''(x_N)\\
&-\frac{3\alpha h^2 p_0}{2}\Big[\frac{-3u(x_0)+4u(x_1)-u(x_2)}{2h} - u'(x_0)\Big]\\
&-\frac{3\alpha h^2 p_N}{2}\Big[\frac{3u(x_N)-4u(x_{N-1})+u(x_{N-2})}{2h} - u'(x_N)\Big]\\
&+\Big[-3\varepsilon - \frac{3hp_{i-1}}{4} - hp_i + \frac{hp_{i+1}}{4}\Big]u(x_{i-1}) + \Big[6\varepsilon\\
&+hp_{i-1} - hp_{i+1}\Big]u(x_i) + \Big[-3\varepsilon - \frac{hp_{i-1}}{4} + hp_i\\
&+\frac{3hp_{i+1}}{4}\Big]u(x_{i+1}) - \frac{h^2}{2}\Big[-\varepsilon u''(x_{i-1}) + p_{i-1}u'(x_{i-1})\Big]\\
&-2h^2\Big[-\varepsilon u''(x_i) + p_i u'(x_i)\Big] - \frac{h^2}{2}\Big[-\varepsilon u''(x_{i+1}) + p_{i+1}u'(x_{i+1})\Big]. \qquad (5.5)
\end{aligned}$$

Substituting the Taylor expansions for $u(x_{i-1})$, $u(x_{i+1})$, $u'(x_{i-1})$, $u'(x_{i+1})$, $u''(x_{i-1})$, $u''(x_{i+1})$ about the point x_i,

$$\begin{aligned}
\frac{-3u(x_0)+4u(x_1)-u(x_2)}{2h} &= u'(x_0) - \frac{h^2}{3}u'''(x_0) - \dots,\\
\frac{3u(x_N)-4u(x_{N-1})+u(x_{N-2})}{2h} &= u'(x_N) - \frac{h^2}{3}u'''(x_N) + \dots
\end{aligned}$$

in (5.5), we get

$$\begin{aligned}
T_i(h) =& -3\alpha\varepsilon h^2\Big[u'(x_0) - \frac{h^2}{3}u'''(x_0) + \mathcal{O}(h^3)\Big] + 3\alpha\varepsilon h^2\Big[u'(x_N) - \frac{h^2}{3}u'''(x_N) + \mathcal{O}(h^3)\Big]\\
&-\frac{3\alpha h^2 p_0}{2}\Big[-\frac{h^2}{3}u'''(x_0) + \mathcal{O}(h^3)\Big] - \frac{3\alpha h^2 p_N}{2}\Big[-\frac{h^2}{3}u'''(x_N) + \mathcal{O}(h^3)\Big]\\
&-\frac{3\alpha\varepsilon h^2}{2}\Big[u''(x_0) + u''(x_N)\Big] + \frac{h^4}{6}\Big[-p_{i-1} + 2p_i - p_{i+1}\Big]u'''(x_i)\\
&+h^4\Big[\frac{hp_{i-1}}{24} - \frac{hp_{i+1}}{24} + \frac{\varepsilon}{4}\Big]u^{(iv)}(x_i) + \mathcal{O}(h^6).
\end{aligned}$$

Since $|\alpha| < h^2$, when $h \to 0$, it can be seen that $|T_i(h)| \leq K_i h^4$ where K_i is a constant. Hence $T_i(h) = \mathcal{O}(h^4)$ as $h \to 0$ for $i = 1, 2, \dots, N-1$.

The system given in (5.2) can be written as $AU = d$, where

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{1,N-2} & A_{1,N-1} \\ A_{2,1} & A_{2,2} & A_{2,3} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{2,N-2} & A_{2,N-1} \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{3,N-2} & A_{3,N-1} \\ A_{4,1} & A_{4,2} & A_{4,3} & A_{4,4} & A_{4,5} & 0 & \dots & 0 & 0 & 0 & 0 & A_{4,N-2} & A_{4,N-1} \\ A_{5,1} & A_{5,2} & 0 & A_{5,4} & A_{5,5} & A_{5,6} & \dots & 0 & 0 & 0 & 0 & A_{5,N-2} & A_{5,N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{N-5,1} & A_{N-5,2} & 0 & 0 & 0 & 0 & \dots & A_{N-5,N-6} & A_{N-5,N-5} & A_{N-5,N-4} & 0 & A_{N-5,N-2} & A_{N-5,N-1} \\ A_{N-4,1} & A_{N-4,2} & 0 & 0 & 0 & 0 & \dots & 0 & A_{N-4,N-5} & A_{N-4,N-4} & A_{N-4,N-3} & A_{N-4,N-2} & A_{N-4,N-1} \\ A_{N-3,1} & A_{N-3,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & A_{N-3,N-4} & A_{N-3,N-3} & A_{N-3,N-2} & A_{N-3,N-1} \\ A_{N-2,1} & A_{N-2,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & A_{N-2,N-3} & A_{N-2,N-2} & A_{N-2,N-1} \\ A_{N-1,1} & A_{N-1,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{N-1,N-2} & A_{N-1,N-1} \end{bmatrix}$$

$A_{i,j}$ is the coefficient of the U_j of the i-th equation, $U = (U_1, U_2, \dots, U_{N-1})^T$, $d = (d_1, d_2, \dots, d_{N-1})^T$, d_i, $i = 1, 2, \dots, N-1$ are the right hand side of the system (5.2). Let B be the matrix corresponding to $\alpha = 0$.

By taking the matrices A, B, $n = N-1$, $\|.\|_\infty$ norm in the Theorem 1.7.4, we obtain

$$\max_i |\lambda_i' - \mu_i'| \leq 2^{\frac{2N-3}{N-1}} (N-1)^{\frac{1}{N-1}} (2Q)^{\frac{N-2}{N-1}} \|A - B\|_\infty^{\frac{1}{N-1}}, \tag{5.6}$$

where $Q = \max\{\|A\|_\infty,\ \|B\|_\infty\}$, λ_i' and μ_i', $i = 1, 2, \dots, N-1$ are the eigenvalues of A and B respectively. Let $\mathcal{R}_i = \sum_{j=1}^{N-1} B_{i,j}$ be the row sum of B. Note that, when h is sufficiently small, the matrix B is an irreducible matrix with

- $B_{i,i} = 6\varepsilon + 2h^2 q_i + hp_{i-1} - hp_{i+1} > 0$, $i = 1, 2, \dots, N-1$,
- $B_{i,i-1} = -3\varepsilon + \frac{h^2 q_{i-1}}{2} - \frac{3hp_{i-1}}{4} - hp_i + \frac{hp_{i+1}}{4} < 0$, $i = 2, 3, \dots, N-1$,
- $B_{i,i+1} = -3\varepsilon + \frac{h^2 q_{i+1}}{2} - \frac{hp_{i-1}}{4} + hp_i + \frac{3hp_{i+1}}{4} < 0$, $i = 1, 2, \dots, N-2$,
- $\mathcal{R}_1 = 3\varepsilon + 2h^2 q_1 + \frac{h^2 q_2}{2} + \frac{3hp_0}{4} + hp_1 - \frac{hp_2}{4} > 0$,
- $\mathcal{R}_i = \frac{h^2 q_{i-1}}{2} + 2h^2 q_i + \frac{h^2 q_{i+1}}{2} > 0$, $i = 2, 3, \dots, N-2$,
- $\mathcal{R}_{N-1} = 3\varepsilon + \frac{h^2 q_{N-2}}{2} + 2h^2 q_{N-1} + \frac{hp_{N-2}}{4} - hp_{N-1} - \frac{3hp_N}{4} > 0$.

Hence, B is a monotone matrix [67]. Thus B^{-1} exists and eigenvalues μ_i', $i = 1, 2, \dots, N-1$ of B are non-zero. Thus, when h is sufficiently small, B is invertible and its row sums are positive. Once we fix h (which gives B is monotone), then the eigenvalues of

B are non-zero. Now the α can vary in the region $(-h^2, h^2)$. We can choose sufficiently small α in the region $(-h^2, h^2)$ to satisfy the following conditions:

- A is invertible because $\|A - B\|_\infty = \Big| -h\alpha\Big| \Big(\Big|3p_0 + 6\varepsilon\Big| + \Big| -3p_N + 6\varepsilon\Big| \Big) + \Big|\frac{h\alpha}{4}\Big| \Big(\Big|3p_0 + 6\varepsilon\Big| + \Big| -3p_N + 6\varepsilon\Big| \Big)$ and from (5.6), it can been seen that, when α is sufficiently small, eigenvalues of A are non-zero.

- The row sum $\mathcal{S}_i$ of the matrix A
$$\mathcal{S}_i = \mathcal{R}_i - \frac{9h\alpha p_0}{4} + \frac{9h\alpha p_N}{4} - 9h\alpha\varepsilon > 0, \ i = 1, 2, \ldots, N-1,$$
when α is sufficiently small.

Hence by choosing sufficiently small h (which gives B is monotone), choosing α in $(-h^2, h^2)$ (which gives A is invertible and row sums of A are positive), we get the following:

The system given in (5.2) with exact solutions can be written as

$$A\bar{u} = d + T(h),$$

where $\bar{u} = (u(x_1), u(x_2), \ldots, u(x_{N-1}))^T$, $T(h) = (T_1(h), T_2(h), \ldots, T_{N-1}(h))^T$. Since $AU = d$, we get $A(\bar{u} - U) = T(h)$, i.e, $AE = T(h)$ where $E = (E_1, E_2, \ldots, E_{N-1})^T$, $E_i = u(x_i) - U_i$. Consequently, we have

$$E = A^{-1}T(h).$$

From the theory of matrices, we have

$$\sum_{i=1}^{N-1} A^{-1}_{k,i}\mathcal{S}_i = 1, \ k = 1, 2, \ldots, N-1,$$

where $A^{-1}_{k,i}$ is the (k, i)-th element of the matrix A^{-1}. We have

$$\mathcal{S}_1 = 3\varepsilon + 2h^2 q_1 + \frac{h^2 q_2}{2} + \frac{3hp_0}{4} + hp_1 - \frac{hp_2}{4} - \frac{9h\alpha p_0}{4} + \frac{9h\alpha p_N}{4} - 9h\alpha\varepsilon,$$
$$\mathcal{S}_i = h^2\Big[\frac{q_{i-1}}{2} + 2q_i + \frac{q_{i+1}}{2} - \frac{9\alpha p_0}{4h} + \frac{9\alpha p_N}{4h} - \frac{9\alpha\varepsilon}{h}\Big],$$
$$\mathcal{S}_{N-1} = 3\varepsilon + \frac{h^2 q_{N-2}}{2} + 2h^2 q_{N-1} + \frac{hp_{N-2}}{4} - hp_{N-1} - \frac{3hp_N}{4}$$

$$-\frac{9h\alpha p_0}{4}+\frac{9h\alpha p_N}{4}-9h\alpha\varepsilon.$$

Hence, we have

$$\sum_{i=1}^{N-1} A_{k,i}^{-1} \leq \frac{1}{\min_{1\leq i\leq N-1}\mathcal{S}_i} = \frac{1}{h^2 C_{i_0}},$$

where C_{i_0} is a constant. We have

$$E_j = \sum_{i=1}^{N-1} A_{j,i}^{-1} T_i(h),\ j=1,2,\ldots,N-1,$$

and hence $|E_j|\leq \frac{Kh^4}{C_{i_0}h^2}$, where K is a constant. Therefore, it follows that as $h\to 0$

$$\|E\|_\infty = \mathcal{O}(h^2).$$

Thus, we have the second-order convergence method to get the numerical solutions of the BVP given in (5.1).

5.1.2 Numerical examples

In order to see the effectiveness of our method, we compute the numerical solutions of few singularly perturbed differential equations.

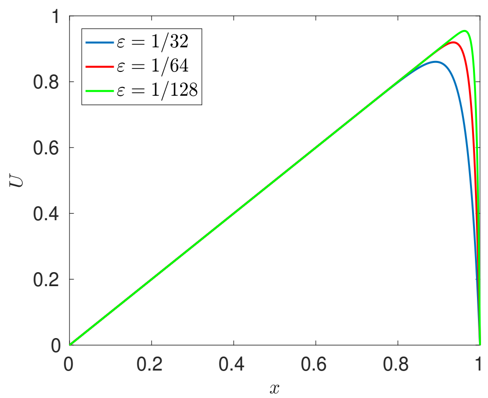
Example 5.1.1. *Consider the singularly perturbed BVP*

$$\begin{cases} \varepsilon u''(x) = u'(x) + u(x) + f(x),\ x\in(0,1), \\ u(0)=0, \quad u(1)=0. \end{cases}$$

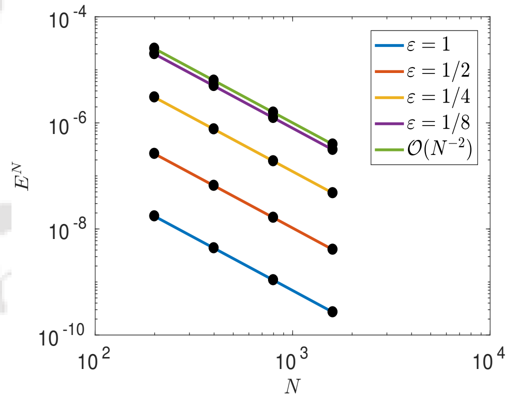
The exact solution is $u(x) = x - \frac{e^{-(1-x)/\varepsilon} - e^{-1/\varepsilon}}{1-e^{-1/\varepsilon}}$. The scaling factors are taken as $\alpha = 0.9999h^2$ for $\varepsilon = 2^{-(i-1)}$, $i=1,2,\ldots,8$. The maximum point-wise error and the order of convergence are given in Table 5.1. Table 5.2 represents the maximum point-wise error and the order of convergence of the cubic spline method. Figure 5.1(a) represents the numerical solutions corresponding to Example 5.1.1. Figure 5.1(b) represents the Log-log plots of the maximum point-wise errors corresponding to Example 5.1.1.

Table 5.1: Maximum point-wise error and order of convergence corresponding to Example 5.1.1.

$\downarrow \varepsilon / N \rightarrow$	200	400	800	1600
1	1.7289e-08	4.3228e-09	1.0802e-09	2.6838e-10
	1.9998	2.0007	2.0090	
1/2	2.6005e-07	6.5022e-08	1.6257e-08	4.0578e-09
	1.9998	1.9999	2.0023	
1/4	3.0208e-06	7.5533e-07	1.8885e-07	4.7207e-08
	1.9998	1.9999	2.0001	
1/8	1.9783e-05	4.9478e-06	1.2371e-06	3.0928e-07
	1.9994	1.9999	1.9999	
1/16	1.1922e-04	2.9776e-05	7.4434e-06	1.8607e-06
	2.0014	2.0001	2.0001	
1/32	6.1082e-04	1.5287e-04	3.8184e-05	9.5439e-06
	1.9985	2.0012	2.0003	
1/64	2.8078e-03	6.9459e-04	1.7318e-04	4.3274e-05
	2.0152	2.0039	2.0007	
1/128	1.1815e-02	2.9861e-03	7.3947e-04	1.8443e-04
	1.9843	2.0137	2.0034	



(a) Numerical solutions with $N = 401$.



(b) Log-log plot.

Figure 5.1: Numerical results of Example 5.1.1.

Table 5.2: Maximum point-wise error and order of convergence corresponding to Example 5.1.1 with $\alpha = 0$.

$\downarrow \varepsilon/N \rightarrow$	200	400	800	1600
1	2.2877e-07	5.7192e-08	1.4299e-08	3.5772e-09
	2.0000	1.9999	1.9990	
1/2	1.5552e-06	3.8880e-07	9.7198e-08	2.4309e-08
	2.0000	2.0000	1.9994	
1/4	8.8620e-06	2.2155e-06	5.5386e-07	1.3847e-07
	2.0000	2.0000	1.9999	
1/8	4.1721e-05	1.0430e-05	2.6074e-06	6.5183e-07
	2.0001	2.0000	2.0000	
1/16	1.8010e-04	4.5000e-05	1.1249e-05	2.8122e-06
	2.0008	2.0001	2.0000	
1/32	7.5223e-04	1.8766e-04	4.6900e-05	1.1725e-05
	2.0031	2.0005	2.0000	
1/64	3.1015e-03	7.6877e-04	1.9179e-04	4.7946e-05
	2.0124	2.0031	2.0000	
1/128	1.2497e-02	3.1365e-03	7.7744e-04	1.9395e-04
	1.9943	2.0123	2.0031	

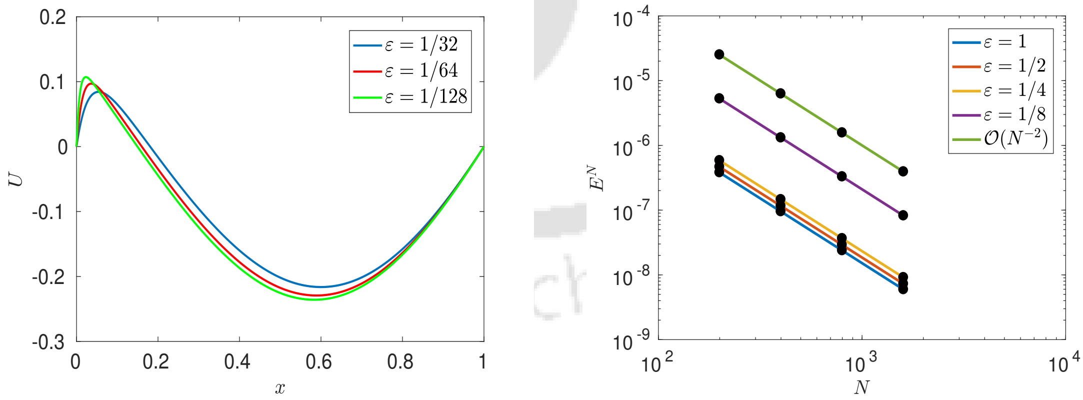
Example 5.1.2. *Consider the singularly perturbed BVP [75]*

$$\begin{cases} \varepsilon u''(x) = -u'(x) + u(x) - \cos(\pi x), \; x \in (0,1), \\ u(0) = 0, \quad u(1) = 0. \end{cases}$$

The solution of the BVP is $u(x) = a\cos(\pi x) + b\sin(\pi x) + Ae^{m_1 x} + Be^{-m_2(1-x)}$ where $a = \frac{\varepsilon\pi^2+1}{\pi^2+(\varepsilon\pi^2+1)^2}$, $b = \frac{-\pi}{\pi^2+(\varepsilon\pi^2+1)^2}$, $m_1 = \frac{-1-\sqrt{1+4\varepsilon}}{2\varepsilon}$, $m_2 = \frac{-1+\sqrt{1+4\varepsilon}}{2\varepsilon}$, $A = -a\frac{1+e^{-m_2}}{1-e^{m_1-m_2}}$, $B = a\frac{1+e^{m_1}}{1-e^{m_1-m_2}}$. The scaling factors are taken as $\alpha = 0.9999h^2$ for $\varepsilon = 2^{-(i-1)}$, $i = 1, 2, \ldots, 8$. The maximum point-wise error and the order of convergence are given in Table 5.3. The maximum point-wise error and the order of convergence corresponding to the cubic spline are given in Table 5.4. The numerical solutions corresponding to Example 5.1.2 are given in Figure 5.2(a). Figure 5.2(b) represents the Log-log plots of the maximum point-wise errors.

Table 5.3: Maximum point-wise error and order of convergence corresponding to Example 5.1.2.

$\downarrow \varepsilon/N \rightarrow$	200	400	800	1600
1	3.7952e-07	9.4892e-08	2.3723e-08	5.9306e-09
	1.9998	2.0000	2.0000	
1/2	4.6300e-07	1.1577e-07	2.8942e-08	7.2346e-09
	1.9998	2.0000	2.0002	
1/4	5.8122e-07	1.4534e-07	3.6337e-08	9.0855e-09
	1.9996	1.9999	1.9998	
1/8	5.2439e-06	1.3102e-06	3.2750e-07	8.1872e-08
	2.0009	2.0002	2.0000	
1/16	2.5093e-05	6.2741e-06	1.5680e-06	3.9197e-07
	1.9998	2.0005	2.0001	
1/32	1.0315e-04	2.5779e-05	6.4390e-06	1.6096e-06
	2.0005	2.0013	2.0001	
1/64	4.1337e-04	1.0224e-04	2.5492e-05	6.3718e-06
	2.0154	2.0039	2.0003	
1/128	1.6115e-03	4.0770e-04	1.0095e-04	2.5178e-05
	1.9828	2.0138	2.0035	



(a) Numerical solutions with $N = 401$.

(b) Log-log plot.

Figure 5.2: Numerical results of Example 5.1.2.

Table 5.4: Maximum point-wise error and order of convergence corresponding to Example 5.1.2 with $\alpha = 0$.

$\downarrow \varepsilon / N \rightarrow$	200	400	800	1600
1	8.0294e-07	2.0074e-07	5.0185e-08	1.2546e-08
	2.0000	2.0000	2.0000	
1/2	2.0076e-06	5.0191e-07	1.2548e-07	3.1368e-08
	2.0000	2.0000	2.0000	
1/4	4.7328e-06	1.1832e-06	2.9580e-07	7.3949e-08
	2.0000	2.0000	2.0000	
1/8	1.2717e-05	3.1788e-06	7.9468e-07	1.9867e-07
	2.0002	2.0001	2.0000	
1/16	3.8196e-05	9.5438e-06	2.3856e-06	5.9638e-07
	2.0008	2.0002	2.0000	
1/32	1.2681e-04	3.1632e-05	7.9038e-06	1.9758e-06
	2.0032	2.0008	2.0001	
1/64	4.5651e-04	1.1314e-04	2.8224e-05	7.0537e-06
	2.0125	2.0031	2.0005	
1/128	1.7043e-03	4.2817e-04	1.0612e-04	2.6474e-05
	1.9929	2.0124	2.0031	

5.2 Linear Singular Boundary-Value Problems

We consider the singular BVPs of the form

$$\begin{cases} u''(x) + \dfrac{k}{x}\, u'(x) - q(x)u(x) = f(x), \;\; x \in (0,1), \\ u'(0) = 0, \quad u(1) = \eta_1, \end{cases} \tag{5.7}$$

where $k = 1, 2$, functions q and f are sufficiently smooth in $[0,1]$, $q(x) > 0$ in $[0,1]$.

Let us consider a uniform mesh on the interval $I = [0,1]$ such that $0 = x_0 < x_1 < \cdots < x_N = 1$, where $x_i = ih$, $i = 0, 1, \ldots, N$ and $h = 1/N$. Let U_i be the approximation of $u(x_i)$, U_i' be the approximation of $u'(x_i)$ and M_i be the approximation of $u''(x_i)$ for $i = 0, 1, \ldots, N$. Let $p(x) = \frac{k}{x}$, $p_i = p(x_i)$, $i = 1, 2, \ldots, N$, $q_i = q(x_i)$, $f_i = f(x_i)$, $i = 0, 1, \ldots, N$.

Using L'Hospital's rule, the differential equation given in (5.7) at $x = x_0$ can be written as $u''(x_0)(k+1) - q_0 u(x_0) = f_0$. Now, at $x = x_0$, the differential equation given in (5.7) is discretized as $M_0(k+1) - q_0 U_0 = f_0$ and at $x = x_i$, $i = 1, 2, \ldots, N$, the differential equation given in (5.7) is discretized as $M_i + p_i U_i' - q_i U_i = f_i$, $i = 1, 2, \ldots, N$. The boundary conditions are discretized as $U_0' = 0$ and $U_N = \eta_1$. We consider the following approximations:

$$U_i' = \frac{U_{i+1} - U_{i-1}}{2h},\ U_{i+1}' = \frac{3U_{i+1} - 4U_i + U_{i-1}}{2h},\ U_{i-1}' = \frac{-U_{i+1} + 4U_i - 3U_{i-1}}{2h}.$$

Substitute the following

$$\begin{aligned}
&\Phi'(x_0) = 0 \text{ (because } U_0' = 0), M_0 = (f_0 + q_0 U_0)/(k+1),\\
&M_1 = f_1 - p_1 U_1' + q_1 U_1 \text{ with } U_1' = (U_2 - U_0)/2h,\\
&M_N = f_N - p_N U_N' + q_N U_N \text{ with } U_N' = (3U_N - 4U_{N-1} + U_{N-2})/2h,
\end{aligned}$$

in (4.12), we get the relation as

$$\begin{aligned}
&\Big[\frac{2(1-\alpha)h^2 q_0}{k+1} + \frac{hp_1}{2} + 6 - 6\alpha h^2\Big] U_0 + \Big[h^2 q_1 - 6\Big] U_1\\
&- \frac{hp_1}{2} U_2 + \frac{\alpha h p_N}{2} U_{N-2} - 2\alpha h p_N U_{N-1} = -\frac{2(1-\alpha)h^2 f_0}{k+1}\\
&- h^2 f_1 + \alpha h^2 f_N - \frac{3\alpha h p_N \eta_1}{2} + \alpha h^2 q_N \eta_1 - 6\alpha h^2 \eta_1,
\end{aligned} \tag{5.8}$$

where $U_N = \eta_1$ is the discretization for the boundary condition.

Put $i = 1$ in (4.5), we get

$$\begin{aligned}
&- 3\alpha\Phi'(x_0) + \Big[-\frac{3\alpha}{2} + \frac{1}{2}\Big] M_0 + 2M_1 + \frac{M_2}{2} - \frac{3\alpha}{2} M_N + 3\alpha\Phi'(x_N)\\
&= \frac{3}{h^2}[U_2 - 2U_1 + U_0].
\end{aligned} \tag{5.9}$$

Substituting

$$\begin{aligned}
&\Phi'(x_0) = 0,\ \Phi'(x_N) = U_N' = (3U_N - 4U_{N-1} + U_{N-2})/2h,\\
&M_0 = (f_0 + q_0 U_0)/(k+1),\ M_i = f_i - p_i U_i' + q_i U_i,\ i = 1, 2, N,\\
&U_1' = (U_2 - U_0)/2h,\ U_2' = (3U_2 - 4U_1 + U_0)/2h,
\end{aligned}$$

in (5.9), we obtain the relation as

$$\Big[-\frac{3\alpha h^2 q_0}{2(k+1)}+\frac{h^2 q_0}{2(k+1)}+hp_1-\frac{hp_2}{4}-3\Big]U_0+\Big[2h^2q_1+hp_2+6\Big]U_1$$
$$+\Big[-hp_1-\frac{3hp_2}{4}+\frac{h^2q_2}{2}-3\Big]U_2+\frac{h}{2}\Big[\frac{3\alpha p_N}{2}+3\alpha\Big]U_{N-2}$$
$$-2h\Big[\frac{3\alpha p_N}{2}+3\alpha\Big]U_{N-1}=\frac{3\alpha h^2 f_0}{2(k+1)}-\frac{h^2 f_0}{2(k+1)}-2h^2f_1-\frac{h^2f_2}{2}$$
$$+\frac{3h^2\alpha f_N}{2}-\frac{3h\eta_1}{2}\Big[\frac{3\alpha p_N}{2}+3\alpha\Big]+\frac{3h^2\alpha\eta_1 q_N}{2}. \tag{5.10}$$

Substitute the following

$$\Phi'(x_0)=0, \Phi'(x_N)=U'_N=(3U_N-4U_{N-1}+U_{N-2})/2h, U'_i=(U_{i+1}-U_{i-1})/2h,$$
$$M_0=(f_0+q_0U_0)/(k+1), M_j=f_j-p_jU'_j+q_jU_j, j=i-1,i,i+1,N,$$
$$U'_{i+1}=(3U_{i+1}-4U_i+U_{i-1})/2h,\ U'_{i-1}=(-U_{i+1}+4U_i-3U_{i-1})/2h,$$

in (4.5), we get the relation as

$$-\frac{3h^2\alpha q_0}{2(k+1)}U_0+\Big[-3+\frac{h^2q_{i-1}}{2}+\frac{3hp_{i-1}}{4}+hp_i-\frac{hp_{i+1}}{4}\Big]U_{i-1}+\Big[6$$
$$+2h^2q_i-hp_{i-1}+hp_{i+1}\Big]U_i+\Big[-3+\frac{h^2q_{i+1}}{2}+\frac{hp_{i-1}}{4}-hp_i$$
$$-\frac{3hp_{i+1}}{4}\Big]U_{i+1}+\frac{h}{2}\Big[\frac{3\alpha p_N}{2}+3\alpha\Big]U_{N-2}-2h\Big[\frac{3\alpha p_N}{2}+3\alpha\Big]U_{N-1}=$$
$$\frac{3\alpha h^2f_0}{2(k+1)}-\frac{h^2f_{i-1}}{2}-2h^2f_i-\frac{h^2f_{i+1}}{2}+\frac{3h^2\alpha f_N}{2}-\frac{3h\eta_1}{2}\Big[\frac{3\alpha p_N}{2}$$
$$+3\alpha\Big]+\frac{3h^2\alpha\eta_1q_N}{2},\quad i=2,3,\ldots,N-1. \tag{5.11}$$

The system given in (5.8), (5.10) and (5.11), gives the approximations $U_0, U_1, \ldots, U_{N-1}$ of the solution $u(x)$ at x_0, x_1, $\ldots$, x_{N-1}.

If $\alpha=0$, then the system given in (5.8), (5.10) and (5.11) reduces into the system corresponding to the cubic spline.

5.2.1 Convergence analysis

In this section, convergence analysis of the proposed numerical scheme is investigated. Let $|\alpha|< h^2$. We get the truncation error $T_0(h)$ associated with the equation given in

(5.8) as

$$T_0(h) = \Big[\frac{2(1-\alpha)h^2q_0}{k+1} + \frac{hp_1}{2} + 6 - 6\alpha h^2\Big]u(x_0) + \Big[h^2q_1 - 6\Big]u(x_1)$$
$$- \frac{hp_1}{2}u(x_2) + \frac{\alpha hp_N}{2}u(x_{N-2}) - 2\alpha hp_N u(x_{N-1}) - \Big[-\frac{2(1-\alpha)h^2f_0}{k+1}$$
$$- h^2f_1 + \alpha h^2 f_N - \frac{3\alpha hp_N u(x_N)}{2} + \alpha h^2 q_N u(x_N) - 6\alpha h^2 u(x_N)\Big]. \tag{5.12}$$

Substituting $f_0 = u''(x_0)(k+1) - q_0u(x_0)$, $f_i = u''(x_i) + p_iu'(x_i) - q_iu(x_i)$, $i = 1, N$ in (5.12) and after simplifying we get

$$T_0(h) = -6\alpha h^2\Big[u(x_0) - u(x_N)\Big] + \alpha h^2 p_N\Big[\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h} - u'(x_N)\Big]$$
$$- 2\alpha h^2u''(x_0) - \alpha h^2u''(x_N) + \Big[\frac{hp_1}{2} + 6\Big]u(x_0) - 6u(x_1) - \frac{hp_1}{2}u(x_2)$$
$$+ 2h^2u''(x_0) + h^2\Big[u''(x_1) + p_1u'(x_1)\Big]. \tag{5.13}$$

Using the Taylor series expansion for $u(x_1)$, $u(x_2)$, $u'(x_1)$, $u''(x_1)$ about the point x_0 and

$$\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h} = u'(x_N) - \frac{h^2}{3}u'''(x_N) + \dots$$

in (5.13), we get

$$T_0(h) = -6\alpha h^2\Big[u(x_0) - u(x_N)\Big] + \alpha h^2 p_N\Big[-\frac{h^2}{3}u'''(x_N) + \dots\Big]$$
$$- 2\alpha h^2u''(x_0) - \alpha h^2u''(x_N) - \frac{h^4p_1}{6}u'''(x_0) + \Big[\frac{h^4}{4} - \frac{h^5p_1}{6}\Big]u^{(4)}(x_0) + \dots.$$

When $h \to 0$, we get $|T_0(h)| \leq K_0h^4$, where K_0 is a constant. Thus, $T_0(h) = \mathcal{O}(h^4)$ as $h \to 0$. We get the truncation error $T_1(h)$ associated with the equation given in (5.10) as

$$T_1(h) = \Big[-\frac{3\alpha h^2q_0}{2(k+1)} + \frac{h^2q_0}{2(k+1)} + hp_1 - \frac{hp_2}{4} - 3\Big]u(x_0) + \Big[2h^2q_1$$
$$+ hp_2 + 6\Big]u(x_1) + \Big[-hp_1 - \frac{3hp_2}{4} + \frac{h^2q_2}{2} - 3\Big]u(x_2)$$
$$+ \frac{h}{2}\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big]u(x_{N-2}) - 2h\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big]u(x_{N-1})$$
$$- \Big(\frac{3\alpha h^2 f_0}{2(k+1)} - \frac{h^2f_0}{2(k+1)} - 2h^2f_1 - \frac{h^2f_2}{2} + \frac{3h^2\alpha f_N}{2}$$
$$- \frac{3h}{2}\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big]u(x_N) + \frac{3h^2\alpha q_N}{2}u(x_N)\Big). \tag{5.14}$$

Substituting $f_0 = u''(x_0)(k+1) - q_0 u(x_0)$, $f_i = u''(x_i) + p_i u'(x_i) - q_i u(x_i)$, $i = 1, 2, N$ in (5.14) and after simplifying, we get

$$
\begin{aligned}
T_1(h) = {}& \frac{3h^2\alpha p_N}{2}\Big[\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h} - u'(x_N)\Big] \\
&+ 3h^2\alpha\Big[\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h}\Big] - \frac{3h^2\alpha}{2}\Big[u''(x_0) + u''(x_N)\Big] \\
&+ \Big[hp_1 - \frac{hp_2}{4} - 3\Big]u(x_0) + \Big[hp_2 + 6\Big]u(x_1) + \Big[-hp_1 \\
&- \frac{3hp_2}{4} - 3\Big]u(x_2) + \frac{h^2}{2}u''(x_0) + 2h^2\Big[u''(x_1) + p_1 u'(x_1)\Big] \\
&+ \frac{h^2}{2}\Big[u''(x_2) + p_2 u'(x_2)\Big].
\end{aligned} \tag{5.15}
$$

Substituting the Taylor expansion for $u(x_0)$, $u(x_2)$, $u'(x_2)$, $u''(x_0)$ and $u''(x_2)$ about the point x_1 and

$$
\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h} = u'(x_N) - \frac{h^2}{3}u'''(x_N) + \ldots
$$

in (5.15), we get

$$
\begin{aligned}
T_1(h) = {}& \frac{3h^2\alpha p_N}{2}\Big[-\frac{h^2}{3}u'''(x_N) + \ldots\Big] + 3h^2\alpha\Big[u'(x_N) - \frac{h^2}{3}u'''(x_N) + \ldots\Big] \\
&- \frac{3h^2\alpha}{2}\Big[u''(x_0) + u''(x_N)\Big] + \Big[-\frac{h^4 p_1}{3} + \frac{h^4 p_2}{6}\Big]u'''(x_1) \\
&+ \Big[\frac{h^4}{4} + \frac{h^5 p_2}{24}\Big]u^{(4)}(x_1) + \ldots.
\end{aligned}
$$

When $h \to 0$, it can be observed that $|T_1(h)| \leq K_1 h^4$, where K_1 is a constant. Therefore, $T_1(h) = \mathcal{O}(h^4)$ as $h \to 0$. Using the procedure followed to compute $T_1(h)$, we get the truncation error $T_i(h)$, $i = 2, 3, \ldots, N-1$ associated with the equations given in (5.11) as

$$
\begin{aligned}
T_i(h) = {}& \frac{3h^2\alpha p_N}{2}\Big[-\frac{h^2}{3}u'''(x_N) + \ldots\Big] + 3h^2\alpha\Big[u'(x_N) - \frac{h^2}{3}u'''(x_N) + \ldots\Big] \\
&- \frac{3h^2\alpha}{2}\Big[u''(x_0) + u''(x_N)\Big] + \Big[\frac{h^4 p_{i-1}}{6} - \frac{h^4 p_i}{3} + \frac{h^4 p_{i+1}}{6}\Big]u'''(x_i) \\
&+ \Big[\frac{h^4}{4} - \frac{h^5 p_{i-1}}{24} + \frac{h^5 p_{i+1}}{24}\Big]u^{(4)}(x_i) + \ldots, \quad i = 2, 3, \ldots, N-1.
\end{aligned}
$$

When $h \to 0$, it can be seen that $|T_i(h)| \leq K_i h^4$, where K_i is a constant. Thus, $T_i(h) = \mathcal{O}(h^4)$, $i = 2, 3, \ldots, N-1$ as $h \to 0$.

The system (5.8), (5.10) and (5.11) can be written as $AU = d$, where

$$A = \begin{bmatrix} A_{0,0} & A_{0,1} & A_{0,2} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{0,N-2} & A_{0,N-1} \\ A_{1,0} & A_{1,1} & A_{1,2} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{1,N-2} & A_{1,N-1} \\ A_{2,0} & A_{2,1} & A_{2,2} & A_{2,3} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{2,N-2} & A_{2,N-1} \\ A_{3,0} & 0 & A_{3,2} & A_{3,3} & A_{3,4} & 0 & \dots & 0 & 0 & 0 & 0 & A_{3,N-2} & A_{3,N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{N-5,0} & 0 & 0 & 0 & 0 & 0 & \dots & A_{N-5,N-6} & A_{N-5,N-5} & A_{N-5,N-4} & 0 & A_{N-5,N-2} & A_{N-5,N-1} \\ A_{N-4,0} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & A_{N-4,N-5} & A_{N-4,N-4} & A_{N-4,N-3} & A_{N-4,N-2} & A_{N-4,N-1} \\ A_{N-3,0} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & A_{N-3,N-4} & A_{N-3,N-3} & A_{N-3,N-2} & A_{N-3,N-1} \\ A_{N-2,0} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & A_{N-2,N-3} & A_{N-2,N-2} & A_{N-2,N-1} \\ A_{N-1,0} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{N-1,N-2} & A_{N-1,N-1} \end{bmatrix},$$

$A_{i,j}$ is the coefficient of U_j, $U = (U_0, U_1, \dots, U_{N-1})^T$, $d = (d_0, d_1, \dots, d_{N-1})^T$, d_i is the right hand side of the equations given in (5.8), (5.10) and (5.11). Let B be the matrix corresponding to $\alpha = 0$.

We obtain $\|A - B\|_\infty = \max\{C_1, C_2\}$,

$$C_1 = \Big| -\frac{2\alpha h^2 q_0}{k+1} - 6\alpha h^2 \Big| + \Big|\frac{\alpha h p_N}{2}\Big| + \Big| -2\alpha h p_N \Big|,$$
$$C_2 = \Big| -\frac{3\alpha h^2 q_0}{2(k+1)}\Big| + \Big|\frac{h}{2}\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big]\Big| + \Big| -2h\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big]\Big|.$$

By taking the matrices A, B, $n = N$, $\|.\|_\infty$ norm in the Theorem 1.7.4, we get

$$\max_i |\lambda_i' - \mu_i'| \leq 2^{\frac{2N-1}{N}} (N)^{\frac{1}{N}} (2Q)^{\frac{N-1}{N}} \|A - B\|_\infty^{\frac{1}{N}} \tag{5.16}$$

where $Q = \max\{\|A\|_\infty, \|B\|_\infty\}$, λ_i' and μ_i', $i = 0, 1, \dots, N-1$ are the eigenvalues of A and B respectively.

It can be observed that, if h is sufficiently small, we get $B_{0,0} = \frac{2h^2 q_0}{k+1} + \frac{hp_1}{2} + 6 > 0$, $B_{0,1} = h^2 q_1 - 6 < 0$, $B_{0,2} = -\frac{hp_1}{2} < 0$, $B_{1,0} = -3 + \frac{h^2 q_0}{2(k+1)} + hp_1 - \frac{hp_2}{4} = -3 + \frac{h^2 q_0}{2(k+1)} + k - \frac{hp_2}{4} < 0$, $B_{1,1} = 2h^2 q_1 + hp_2 + 6 > 0$, $B_{1,2} = -hp_1 - \frac{3hp_2}{4} + \frac{h^2 q_2}{2} - 3 < 0$, for $i = 2, 3, \dots, N-1$, we have $B_{i,i} = 6 + 2h^2 q_i - hp_{i-1} + hp_{i+1} = 6 + 2h^2 q_i - \frac{k}{i-1} + hp_{i+1} > 0$, $B_{i,i-1} = -3 + \frac{h^2 q_{i-1}}{2} + \frac{3k}{4(i-1)} + \frac{k}{i} - \frac{hp_{i+1}}{4} < 0$, $B_{i,i+1} = -3 + \frac{h^2 q_{i+1}}{2} + \frac{hp_{i-1}}{4} - hp_i - \frac{3hp_{i+1}}{4} = -3 + \frac{h^2 q_{i+1}}{2} + \frac{k}{4(i-1)} - hp_i - \frac{3hp_{i+1}}{4} < 0$, the row sums $\mathcal{R}_i$, $i = 0, 1, \dots, N-1$ of the matrix B, i.e.,

$$\mathcal{R}_0 = \frac{2h^2 q_0}{k+1} + h^2 q_1 > 0,$$

$$\mathcal{R}_1 = \frac{h^2 q_0}{2(k+1)} + 2h^2 q_1 + \frac{h^2 q_2}{2} > 0,$$
$$\mathcal{R}_i = \frac{h^2 q_{i-1}}{2} + 2h^2 q_i + \frac{h^2 q_{i+1}}{2} > 0,\ i = 2, 3, \ldots, N-2,$$
$$\mathcal{R}_{N-1} = 3 + \frac{h^2 q_{N-2}}{2} + 2h^2 q_{N-1} - \frac{hp_{N-2}}{4} + hp_{N-1} + \frac{3hp_N}{4} > 0.$$

Thus, when h is sufficiently small, the matrix B is irreducible, $B_{i,i} > 0$, $B_{i,j} \leq 0$, $i \neq j$; $i, j = 0, 1, \ldots, N-1$, row sums $\mathcal{R}_i > 0$, $i = 0, 1, \ldots, N-1$. Hence, B is a monotone matrix when h is sufficiently small [67]. Consequently B^{-1} exists and the eigenvalues μ'_i, $i = 0, 1, \ldots, N-1$ of B are non-zero when h is sufficiently small. Once we fix the sufficiently small h (which gives B is monotone), then α can vary in the region $(-h^2, h^2)$. Now we can choose α sufficiently small in $(-h^2, h^2)$ to satisfy the following conditions:

- A is invertible because $\|A - B\|_\infty = \max\{C_1, C_2\}$ and from (5.16), it can been seen that, the eigenvalues of A are non-zero when α is sufficiently small.

- The row sums of A

  $$\mathcal{S}_0 = \mathcal{R}_0 - \frac{2\alpha h^2 q_0}{k+1} - 6\alpha h^2 + \frac{\alpha h p_N}{2} - 2\alpha h p_N > 0,$$
  $$\mathcal{S}_i = \mathcal{R}_i - \frac{3\alpha h^2 q_0}{2(k+1)} + \frac{h}{2}\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big] - 2h\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big] > 0,$$
  $$i = 1, 2, \ldots, N-1,$$

  when α is sufficiently small.

Hence, when h is sufficiently small (which gives B is monotone) and sufficiently small $\alpha \in (-h^2, h^2)$ (which gives A is invertible and the row sums of A are positive), we get the following:

It can be seen that $AE = T(h)$ where $T(h) = (T_0(h), T_1(h), \ldots, T_{N-1}(h))^T$, $E = (E_0, E_1, \ldots, E_{N-1})^T$, $E_i = u(x_i) - U_i$. Consequently, we obtain

$$E = A^{-1} T(h).$$

By the definition of multiplication of a matrix by its inverse, we have

$$\sum_{i=0}^{N-1} A^{-1}_{j,i} \mathcal{S}_i = 1,\ j = 0, 1, \ldots, N-1,$$

where $A^{-1}_{j,i}$ is the (j,i)-th element of the matrix A^{-1}.

$$\begin{aligned}
\mathcal{S}_0 &= h^2\Big[\frac{2q_0}{k+1} + q_1 - \frac{2\alpha q_0}{k+1} - 6\alpha + \frac{\alpha p_N}{2h} - \frac{2\alpha p_N}{h}\Big],\\
\mathcal{S}_1 &= h^2\Big[\frac{q_0}{2(k+1)} + 2q_1 + \frac{q_2}{2} - \frac{3\alpha q_0}{2(k+1)} + \frac{1}{2h}\Big(\frac{3\alpha p_N}{2} + 3\alpha\Big) - \frac{2}{h}\Big(\frac{3\alpha p_N}{2} + 3\alpha\Big)\Big],\\
\mathcal{S}_i &= h^2\Big[\frac{q_{i-1}}{2} + 2q_i + \frac{q_{i+1}}{2} - \frac{3\alpha q_0}{2(k+1)} + \frac{1}{2h}\Big(\frac{3\alpha p_N}{2} + 3\alpha\Big) - \frac{2}{h}\Big(\frac{3\alpha p_N}{2} + 3\alpha\Big)\Big],\\
\mathcal{S}_{N-1} &= 3 + \frac{h^2 q_{N-2}}{2} + 2h^2 q_{N-1} - \frac{hp_{N-2}}{4} + hp_{N-1} + \frac{3hp_N}{4}\\
&\quad - \frac{3\alpha h^2 q_0}{2(k+1)} + \frac{h}{2}\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big] - 2h\Big[\frac{3\alpha p_N}{2} + 3\alpha\Big].
\end{aligned}$$

It can be observed that $\min\limits_{0\le i\le N-1} \mathcal{S}_i = h^2 C_{i_0}$, where C_{i_0} is a constant. Hence, we have

$$\sum_{i=0}^{N-1} A^{-1}_{j,i} \le 1/\min_{0\le i\le N-1} \mathcal{S}_i = 1/h^2 C_{i_0}.$$

From

$$E_j = \sum_{i=0}^{N-1} A^{-1}_{j,i} T_i(h),\ j = 0, 1, \ldots, N-1,$$

it follows that $|E_j| \le \frac{Kh^4}{C_{i_0} h^2}$, where K is a constant. Therefore, as $h \to 0$,

$$\|E\|_\infty = \mathcal{O}(h^2).$$

5.2.2 Numerical examples

In this section, to implement the proposed method, few singular BVPs are considered.

Example 5.2.1. *Consider the singular BVPs*

$$\begin{cases} u''(x) + \dfrac{k}{x} u'(x) - u(x) = f(x),\ x \in (0,1),\\ u'(0) = 0, \quad u(1) = e, \end{cases}$$

where $k = 1, 2$.

The exact solutions of the BVPs are $u(x) = \exp(x^2)$. The value $\alpha = 0.9999h^2$ is used to calculate the numerical solutions. The maximum point-wise error and the order of convergence are given in Table 5.5. Table 5.6 represents the maximum point-wise error and the order of convergence of the cubic spline method.

Table 5.5: Maximum point-wise error and order of convergence corresponding to Example 5.2.1.

N	150	300	600	1200	2400
$k = 1,\ E^N$	5.7873e-06	1.4464e-06	3.6164e-07	9.0399e-08	2.2363e-08
$k = 1,\ p^N$	2.0005	1.9998	2.0001	2.0152	
$k = 2,\ E^N$	6.1198e-06	1.5304e-06	3.8263e-07	9.5665e-08	2.4009e-08
$k = 2,\ p^N$	1.9996	1.9999	1.9999	1.9944	

Table 5.6: Maximum point-wise error and order of convergence corresponding to Example 5.2.1 with $\alpha = 0$.

N	150	300	600	1200	2400
$k = 1,\ E^N$	2.8955e-05	7.2361e-06	1.8088e-06	4.5219e-07	1.1329e-07
$k = 1,\ p^N$	2.0005	2.0002	2.0000	1.9969	
$k = 2,\ E^N$	2.4097e-05	6.0215e-06	1.5052e-06	3.7629e-07	9.4250e-08
$k = 2,\ p^N$	2.0006	2.0002	2.0000	1.9973	

Example 5.2.2. *Consider the singular BVP [117]*

$$\begin{cases} u''(x) + \dfrac{2}{x}u'(x) - 4u(x) = -2, \ x \in (0,1), \\ u'(0) = 0, \quad u(1) = 5.5. \end{cases}$$

The exact solution is $u(x) = 0.5 + \dfrac{5\sinh(2x)}{x\sinh(2)}$. The $\alpha = 0.999h^2$ is used to calculate the numerical approximations. The maximum point-wise error and order of convergence are given in Table 5.7. The maximum point-wise error and order of convergence corresponding to the cubic spline method are given in Table 5.8.

Table 5.7: Maximum point-wise error and order of convergence corresponding to Example 5.2.2.

N	150	300	600	1200	2400
E^N	5.7271e-07	1.4324e-07	3.5820e-08	8.8861e-09	2.2698e-09
p^N	1.9994	1.9995	2.0112	1.9690	

Table 5.8: Maximum point-wise error and order of convergence corresponding to Example 5.2.2 with $\alpha = 0$.

N	150	300	600	1200	2400
E^N	5.2695e-06	1.3158e-06	3.2883e-07	8.2066e-08	2.0602e-08
p^N	2.0018	2.0005	2.0025	1.9940	

5.3 Non-Linear Singular Boundary-Value Problems

Consider the non-linear singular BVP of the form

$$\begin{cases} u_{xx}(x) + \dfrac{k}{x}\, u_x(x) = F(x, u(x)),\ x \in (0,1), \\ u_x(0) = 0, \quad u(1) = \eta_1, \end{cases} \tag{5.17}$$

where $k = 1, 2$. We assume that for $(x, u(x)) \in D = \{0 \le x \le 1,\ -\infty < u(x) < \infty\}$, the functions F and $\partial F/\partial u$ are continuous, $\partial F/\partial u \ge 0$ on D and $\partial F/\partial u > 0$ on $D^\circ = \{0 < x < 1,\ -\infty < u(x) < \infty\}$. Here the notation u_{xx} is used for u'' mainly for the sake of convenience in the future.

Let us consider a uniform mesh on the interval $I = [0, 1]$ such that $0 = x_0 < x_1 < \cdots < x_N = 1$, where $x_i = ih$, $i = 0, 1, \ldots, N$ and $h = 1/N$.

We use the quasi-linearization technique [17, 113, 115] to reduce the non-linear problem into sequence of linear problems. We choose the reasonable initial approximation for the function $u(x)$ in $F(x, u(x))$, call it $u^{(0)}(x)$ and expand $F(x, u(x))$ around the function $u^{(0)}(x)$ to obtain

$$F(x, u^{(1)}(x)) = F(x, u^{(0)}(x)) + (u^{(1)}(x) - u^{(0)}(x))\Big(\frac{\partial F}{\partial u}\Big)_{(x, u^{(0)}(x))} + \ldots,$$

or in general, we can write for $r = 0, 1, \ldots$ (r is iteration index)

$$F(x, u^{(r+1)}(x)) = F(x, u^{(r)}(x)) + (u^{(r+1)}(x) - u^{(r)}(x))\Big(\frac{\partial F}{\partial u}\Big)_{(x, u^{(r)}(x))} + \ldots.$$

From (5.17), we have

$$\begin{cases} u_{xx}^{(r+1)}(x) + \dfrac{k}{x}\, u_x^{(r+1)}(x) = F(x, u^{(r+1)}(x)),\ x \in (0,1), \\ u_x^{(r+1)}(0) = 0, \quad u^{(r+1)}(1) = \eta_1. \end{cases} \tag{5.18}$$

After substituting

$$F(x, u^{(r+1)}(x)) = F(x, u^{(r)}(x)) + (u^{(r+1)}(x) - u^{(r)}(x))\Big(\frac{\partial F}{\partial u}\Big)_{(x,u^{(r)}(x))}$$

in (5.18), we get

$$\begin{cases} u_{xx}^{(r+1)}(x) + p^{(r)}(x)u_x^{(r+1)}(x) - q^{(r)}(x)u^{(r+1)}(x) = f^{(r)}(x), \quad x \in (0,1), \\ u_x^{(r+1)}(0) = 0, \quad u^{(r+1)}(1) = \eta_1, \end{cases} \tag{5.19}$$

where the functions $p^{(r)}(x) = \dfrac{k}{x}$, $q^{(r)}(x) = \Big(\dfrac{\partial F}{\partial u}\Big)_{(x,u^{(r)}(x))}$ and $f^{(r)}(x) = F(x, u^{(r)}(x)) - u^{(r)}(x)\Big(\dfrac{\partial F}{\partial u}\Big)_{(x,u^{(r)}(x))}$. Thus, we have reduced non-linear problem into sequence of linear problems. Now our aim is to solve these linear problems into numerically.

Let $U_i^{(r)}$ be the approximation of $u^{(r)}(x_i)$, $U_i^{'(r)}$ be the approximation of $u_x^{(r)}(x_i)$ and $M_i^{(r)}$ be the approximation of $u_{xx}^{(r)}(x_i)$. Using L'Hospital's rule, the differential equation given in (5.19) at $x = x_0$ can be written as $u_{xx}^{(r+1)}(x_0)(k+1) - q^{(r)}(x_0)u^{(r+1)}(x_0) = f^{(r)}(x_0)$. Now, at $x = x_0$, the differential equation given in (5.19) is discretized as $M_0^{(r+1)}(k+1) - q_0^{(r)}U_0^{(r+1)} = f_0^{(r)}$ and at $x = x_i$, $i = 1, 2, \ldots, N$, the differential equation given in (5.19) is discretized as $M_i^{(r+1)} + p_i^{(r)}U_i^{'(r+1)} - q_i^{(r)}U_i^{(r)} = f_i^{(r)}$, $i = 1, 2, \ldots, N$, where $p_i^{(r)} = \dfrac{k}{x_i}$, $q_i^{(r)} = \Big(\dfrac{\partial F}{\partial u}\Big)_{(x_i,U_i^{(r)})}$ and $f_i^{(r)} = F(x_i, U_i^{(r)}) - U_i^{(r)}\Big(\dfrac{\partial F}{\partial u}\Big)_{(x_i,U_i^{(r)})}$. The boundary conditions are discretized as $U_0^{'(r+1)} = 0$ and $U_N^{(r+1)} = \eta_1$.

Using the procedure followed in Section 5.2, we get the following system:

$$\begin{aligned} &\Big[\frac{2(1-\alpha)h^2q_0^{(r)}}{k+1} + \frac{hp_1^{(r)}}{2} + 6 - 6\alpha h^2\Big]U_0^{(r+1)} + \Big[h^2q_1^{(r)} - 6\Big]U_1^{(r+1)} \\ &- \frac{hp_1^{(r)}}{2}U_2^{(r+1)} + \frac{\alpha hp_N^{(r)}}{2}U_{N-2}^{(r+1)} - 2\alpha hp_N^{(r)}U_{N-1}^{(r+1)} = -\frac{2(1-\alpha)h^2f_0^{(r)}}{k+1} \\ &- h^2f_1^{(r)} + \alpha h^2f_N^{(r)} - \frac{3\alpha hp_N^{(r)}\eta_1}{2} + \alpha h^2q_N^{(r)}\eta_1 - 6\alpha h^2\eta_1, \end{aligned} \tag{5.20}$$

$$\Big[-\frac{3\alpha h^2 q_0^{(r)}}{2(k+1)}+\frac{h^2 q_0^{(r)}}{2(k+1)}+hp_1^{(r)}-\frac{hp_2^{(r)}}{4}-3\Big]U_0^{(r+1)}+\Big[2h^2q_1^{(r)}+hp_2^{(r)}$$
$$+6\Big]U_1^{(r+1)}+\Big[-hp_1^{(r)}-\frac{3hp_2^{(r)}}{4}+\frac{h^2q_2^{(r)}}{2}-3\Big]U_2^{(r+1)}+\frac{h}{2}\Big[\frac{3\alpha p_N^{(r)}}{2}+3\alpha\Big]U_{N-2}^{(r+1)}$$
$$-2h\Big[\frac{3\alpha p_N^{(r)}}{2}+3\alpha\Big]U_{N-1}^{(r+1)}=\frac{3\alpha h^2 f_0^{(r)}}{2(k+1)}-\frac{h^2 f_0^{(r)}}{2(k+1)}-2h^2f_1^{(r)}-\frac{h^2f_2^{(r)}}{2}$$
$$+\frac{3h^2\alpha f_N^{(r)}}{2}-\frac{3h\eta_1}{2}\Big[\frac{3\alpha p_N^{(r)}}{2}+3\alpha\Big]+\frac{3h^2\alpha\eta_1 q_N^{(r)}}{2}, \tag{5.21}$$

and

$$-\frac{3h^2\alpha q_0^{(r)}}{2(k+1)}U_0^{(r+1)}+\Big[-3+\frac{h^2q_{i-1}^{(r)}}{2}+\frac{3hp_{i-1}^{(r)}}{4}+hp_i^{(r)}-\frac{hp_{i+1}^{(r)}}{4}\Big]U_{i-1}^{(r+1)}$$
$$+\Big[6+2h^2q_i^{(r)}-hp_{i-1}^{(r)}+hp_{i+1}^{(r)}\Big]U_i^{(r+1)}+\Big[-3+\frac{h^2q_{i+1}^{(r)}}{2}+\frac{hp_{i-1}^{(r)}}{4}$$
$$-hp_i^{(r)}-\frac{3hp_{i+1}^{(r)}}{4}\Big]U_{i+1}^{(r+1)}+\frac{h}{2}\Big[\frac{3\alpha p_N^{(r)}}{2}+3\alpha\Big]U_{N-2}^{(r+1)}-2h\Big[\frac{3\alpha p_N^{(r)}}{2}+3\alpha\Big]U_{N-1}^{(r+1)}=$$
$$\frac{3\alpha h^2 f_0^{(r)}}{2(k+1)}-\frac{h^2f_{i-1}^{(r)}}{2}-2h^2f_i^{(r)}-\frac{h^2f_{i+1}^{(r)}}{2}+\frac{3h^2\alpha f_N^{(r)}}{2}-\frac{3h\eta_1}{2}\Big[\frac{3\alpha p_N^{(r)}}{2}$$
$$+3\alpha\Big]+\frac{3h^2\alpha\eta_1 q_N^{(r)}}{2},\quad i=2,3,\ldots,N-1. \tag{5.22}$$

The system (5.20), (5.21) and (5.22) gives the approximations $U_0^{(r+1)}, U_1^{(r+1)}, \ldots, U_{N-1}^{(r+1)}$ of $u^{(r+1)}(x)$ at x_i, $i=0,1,\ldots,N-1$.

5.3.1 Error analysis

Let $|\alpha|< h^2$. The truncation error $T_i^{(r)}(h)$, $i=0,1,\ldots,N-1$ associated with the equations (5.20), (5.21) and (5.22) are obtained as

$$T_0^{(r)}(h)=-6\alpha h^2\Big[u^{(r+1)}(x_0)-u^{(r+1)}(x_N)\Big]+\alpha h^2p_N^{(r)}\Big[-\frac{h^2}{3}u_{xxx}^{(r+1)}(x_N)+\ldots\Big]$$
$$-2\alpha h^2u_{xx}^{(r+1)}(x_0)-\alpha h^2u_{xx}^{(r+1)}(x_N)-\frac{h^4p_1^{(r)}}{6}u_{xxx}^{(r+1)}(x_0)$$
$$+\Big[\frac{h^4}{4}-\frac{h^5p_1^{(r)}}{6}\Big]u_{xxxx}^{(r+1)}(x_0)+\ldots,$$

$$T_1^{(r)}(h)=\frac{3h^2\alpha p_N^{(r)}}{2}\Big[-\frac{h^2}{3}u_{xxx}^{(r+1)}(x_N)+\ldots\Big]+3h^2\alpha\Big[u_x^{(r+1)}(x_N)$$
$$-\frac{h^2}{3}u_{xxx}^{(r+1)}(x_N)+\ldots\Big]-\frac{3h^2\alpha}{2}\Big[u_{xx}^{(r+1)}(x_0)+u_{xx}^{(r+1)}(x_N)\Big]$$

$$+\Big[-\frac{h^4p_1^{(r)}}{3}+\frac{h^4p_2^{(r)}}{6}\Big]u_{xxx}^{(r+1)}(x_1)+\Big[\frac{h^4}{4}+\frac{h^5p_2^{(r)}}{24}\Big]u_{xxxx}^{(r+1)}(x_1)+\dots,$$

and

$$\begin{aligned}
T_i^{(r)}(h) = {} & \frac{3h^2\alpha p_N^{(r)}}{2}\Big[-\frac{h^2}{3}u_{xxx}^{(r+1)}(x_N)+\dots\Big]+3h^2\alpha\Big[u_x^{(r+1)}(x_N) \\
& -\frac{h^2}{3}u_{xxx}^{(r+1)}(x_N)+\dots\Big]-\frac{3h^2\alpha}{2}\Big[u_{xx}^{(r+1)}(x_0)+u_{xx}^{(r+1)}(x_N)\Big] \\
& +\Big[\frac{h^4p_{i-1}^{(r)}}{6}-\frac{h^4p_i^{(r)}}{3}+\frac{h^4p_{i+1}^{(r)}}{6}\Big]u_{xxx}^{(r+1)}(x_i) \\
& +\Big[\frac{h^4}{4}-\frac{h^5p_{i-1}^{(r)}}{24}+\frac{h^5p_{i+1}^{(r)}}{24}\Big]u_{xxxx}^{(r+1)}(x_i)+\dots, \\
& i=2,3,\dots,N-1.
\end{aligned}$$

Since $|\alpha|< h^2$, when $h\to 0$, we have $|T_i^{(r)}(h)|\le K_i^{(r)}h^4$, where $K_i^{(r)}$ is a constant. Hence $T_i^{(r)}(h)=\mathcal{O}(h^4)$ as $h\to 0$ for $i=0,1,\dots,N-1$.

The system (5.20), (5.21) and (5.22) can be written as $A^{(r)}U^{(r+1)}=d^{(r)}$, where

$$A^{(r)}=\begin{bmatrix}
A_{0,0}^{(r)} & A_{0,1}^{(r)} & A_{0,2}^{(r)} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{0,N-2}^{(r)} & A_{0,N-1}^{(r)} \\
A_{1,0}^{(r)} & A_{1,1}^{(r)} & A_{1,2}^{(r)} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{1,N-2}^{(r)} & A_{1,N-1}^{(r)} \\
A_{2,0}^{(r)} & A_{2,1}^{(r)} & A_{2,2}^{(r)} & A_{2,3}^{(r)} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{2,N-2}^{(r)} & A_{2,N-1}^{(r)} \\
A_{3,0}^{(r)} & 0 & A_{3,2}^{(r)} & A_{3,3}^{(r)} & A_{3,4}^{(r)} & 0 & \dots & 0 & 0 & 0 & 0 & A_{3,N-2}^{(r)} & A_{3,N-1}^{(r)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
A_{N-5,0}^{(r)} & 0 & 0 & 0 & 0 & 0 & \dots & A_{N-5,N-6}^{(r)} & A_{N-5,N-5}^{(r)} & A_{N-5,N-4}^{(r)} & 0 & A_{N-5,N-2}^{(r)} & A_{N-5,N-1}^{(r)} \\
A_{N-4,0}^{(r)} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & A_{N-4,N-5}^{(r)} & A_{N-4,N-4}^{(r)} & A_{N-4,N-3}^{(r)} & A_{N-4,N-2}^{(r)} & A_{N-4,N-1}^{(r)} \\
A_{N-3,0}^{(r)} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & A_{N-3,N-4}^{(r)} & A_{N-3,N-3}^{(r)} & A_{N-3,N-2}^{(r)} & A_{N-3,N-1}^{(r)} \\
A_{N-2,0}^{(r)} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & A_{N-2,N-3}^{(r)} & A_{N-2,N-2}^{(r)} & A_{N-2,N-1}^{(r)} \\
A_{N-1,0}^{(r)} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{N-1,N-2}^{(r)} & A_{N-1,N-1}^{(r)}
\end{bmatrix},$$

$A_{i,j}^{(r)}$ is the coefficient of $U_j^{(r+1)}$ in the equations given in (5.20), (5.21) and (5.22), $U^{(r+1)}=(U_0^{(r+1)},\ U_1^{(r+1)},\ \dots,\ U_{N-1}^{(r+1)})^T$, $d^{(r)}=(d_0^{(r)},\ d_1^{(r)},\ \dots,\ d_{N-1}^{(r)})^T$, $d_i^{(r)}$ is the right hand side of the system given in (5.20), (5.21) and (5.22). Let $B^{(r)}$ be the matrix corresponding to $\alpha=0$.

Using the procedure followed in Section 5.2, it can be seen that, when h is sufficiently small, we get the matrix $B^{(r)}$ is monotone. We can choose α in $(-h^2,h^2)$ so that the matrix $A^{(r)}$ is invertible and the row sums $\mathcal{S}_i^{(r)}$, $i=0,1,\dots,N-1$ of $A^{(r)}$ are positive.

Hence, by choosing sufficiently small h (which gives $B^{(r)}$ is monotone) and sufficiently small α in $(-h^2,h^2)$ (which gives $A^{(r)}$ is invertible and row sums of $A^{(r)}$ are positive), we get the following:

It can be seen that

$$A^{(r)}E^{(r+1)} = T^{(r)}(h)$$

where $T^{(r)}(h) = (T_0^{(r)}(h), T_1^{(r)}(h), \dots, T_{N-1}^{(r)}(h))^T$, $E^{(r+1)} = (E_0^{(r+1)}, E_1^{(r+1)}, \dots, E_{N-1}^{(r+1)})^T$, $E_i^{(r+1)} = u^{(r+1)}(x_i) - U_i^{(r+1)}$. Consequently, we get

$$E^{(r+1)} = {A^{(r)}}^{-1}T^{(r)}(h).$$

By the definition of multiplication of a matrix by its inverse, we get

$$\sum_{i=0}^{N-1} {A^{(r)}}^{-1}_{j,i}\mathcal{S}_i^{(r)} = 1, \ j = 0, 1, \dots, N-1,$$

where ${A^{(r)}}^{-1}_{j,i}$ is the (j,i)-th element of the matrix ${A^{(r)}}^{-1}$. We have $\min\limits_{0\le i\le N-1} \mathcal{S}_i^{(r)} = C_{i_0}^{(r)}h^2$, where $C_{i_0}^{(r)}$ is a constant. Hence, we have

$$\sum_{i=0}^{N-1} {A^{(r)}}^{-1}_{j,i} \le 1/\min_{0\le i\le N-1} \mathcal{S}_i^{(r)} = 1/C_{i_0}^{(r)}h^2.$$

We have

$$E_j^{(r+1)} = \sum_{i=0}^{N-1} {A^{(r)}}^{-1}_{j,i}T_i^{(r)}(h), \ j = 0, 1, \dots, N-1,$$

and hence $|E_j^{(r+1)}| \le \frac{K^{(r)}h^4}{C_{i_0}^{(r)}h^2}$, where $K^{(r)}$ is a constant. Therefore, as $h \to 0$, we get

$$\|E^{(r+1)}\|_\infty = \mathcal{O}(h^2).$$

5.3.2 Numerical examples

The numerical solutions of the non-linear singular BVPs are computed as follows: For each fixed N, by taking initial approximation $U_0^{(0)}, U_1^{(0)}, \dots, U_N^{(0)}$, we compute the numerical solutions $U^{(r+1)} = (U_0^{(r+1)}, U_1^{(r+1)}, \dots, U_{N-1}^{(r+1)})$, $r = 0, 1, \dots$. We can take sufficient number of iterations so that maximum error between two successive iterations $\max\limits_i |U_i^{(r+1)} - U_i^{(r)}| < \delta$, where δ is a small tolerance value that is prescribed. Once the criterion is satisfied, we consider $U^{(r+1)}$ as the numerical solution U to the non-linear singular BVPs. To compute the numerical solutions of the following non-linear singular BVPs, we take $\delta = 10^{-15}$ and 8 iterations.

Example 5.3.1. *Consider the BVPs*

$$\begin{cases} u''(x) + \dfrac{k}{x}u'(x) = \dfrac{5x^3[5x^5e^{u(x)} - k - 4]}{4 + x^5}, \; x \in (0,1), \\ u'(0) = 0, \quad u(1) = \log(1/5), \end{cases}$$

where $k = 1, 2$.

The exact solutions of the BVPs are $u(x) = \log(1/(4 + x^5))$. The value $\alpha = 0.99h^2$ is used to get the numerical solutions of the BVPs. The maximum point-wise error and the order of convergence are given in Table 5.9. The maximum point-wise error and the order of convergence corresponding to the cubic spline method are tabulated in Table 5.10. The values $U_i^{(0)} = 1$, $i = 0, 1, \ldots, N-1$, $U_N^{(0)} = \log(1/5)$ are taken to compute the numerical solutions.

Table 5.9: Maximum point-wise error and order of convergence corresponding to Example 5.3.1.

N	100	200	400	800
$k=1$, E^N	2.2900e-06	5.7273e-07	1.4318e-07	3.5795e-08
$k=1$, p^N	1.9994	2.0000	2.0000	
$k=2$, E^N	2.6318e-06	6.5796e-07	1.6449e-07	4.1123e-08
$k=2$, p^N	2.0000	2.0000	2.0000	

Table 5.10: Maximum point-wise error and order of convergence corresponding to Example 5.3.1 with $\alpha = 0$.

N	100	200	400	800
$k=1$, E^N	1.3455e-05	3.3625e-06	8.4054e-07	2.1013e-07
$k=1$, p^N	2.0006	2.0002	2.0000	
$k=2$, E^N	8.3361e-06	2.0825e-06	5.2053e-07	1.3013e-07
$k=2$, p^N	2.0011	2.0003	2.0001	

5.4 Conclusion

In this chapter, fractal cubic spline method is used to solve the nonself-adjoint BVPs. Proposed method has second-order convergent. To illustrate our method and to demonstrate the applicability of our presented method computationally, we considered the nonself-adjoint BVPs whose exact solutions are known. The numerical results of the BVPs are obtained and tabulated. From numerical results, it can be observed that the proposed method is accurate than the cubic spline method.

Chapter 6

Fractal Non-Polynomial Cubic Spline Solutions for Singularly Perturbed Boundary-Value Problems

In this chapter, a fractal non-polynomial cubic spline method is developed to get the numerical solutions for both self-adjoint and nonself-adjoint singularly perturbed BVPs. The numerical method developed in Chapter 4 for the self-adjoint singularly perturbed BVPs and the numerical method developed in Chapter 5 for the nonself-adjoint singularly perturbed BVPs are of second-order convergence. In this chapter, we have developed the fourth-order convergence numerical methods for these BVPs.

To get the numerical approximations for the self-adjoint singularly perturbed BVPs, various kinds of non-polynomial splines have been used. For example, Tirmizi et al. [134] used non-polynomial quartic spline method to solve the self-adjoint singularly perturbed BVPs and the developed method has sixth-order convergence. Khan and Khandelwal [82] developed sixth-order convergence numerical method with the help of non-polynomial sextic spline to obtain the numerical solutions for the self-adjoint singularly perturbed BVPs. Also one can see in References [108, 111, 114] different kinds of non-polynomial splines are used. To obtain the numerical solutions for nonself-adjoint singularly perturbed BVPs, Aziz and Khan [9] developed second-order convergence method using spline in compression. By modifying Aziz and Khan method, Bawa [12] developed fourth-order convergence method using spline in compression.

In Section 6.1, we have taken the continuity conditions of the fractal non-polynomial cubic spline as the discretized equations for computing the numerical solutions of the self-adjoint singularly perturbed BVPs. Convergence analysis of the proposed method is established and computational efficiency of the method is verified through numerical examples. In Section 6.2, we have solved the nonself-adjoint singularly perturbed BVPs through the fractal non-polynomial cubic spline method and the fourth-order convergence is achieved by introducing the parameter in the approximations of derivative values. Numerical examples are given to illustrate the efficiency of the proposed method.

6.1 Self-Adjoint Boundary-Value Problems

Let us consider the BVP of the form

$$\begin{cases} -\varepsilon u''(x) + q(x)u(x) = f(x), \; x \in (0,1), \\ u(0) = \eta_0, \quad u(1) = \eta_1, \end{cases} \tag{6.1}$$

where $0 < \varepsilon \leq 1$, functions q and f are sufficiently smooth in $[0,1]$ and $q(x) > 0$ for all $x \in [0,1]$.

Let us consider a uniform mesh on the interval $I = [0,1]$ such that $0 = x_0 < x_1 < \cdots < x_N = 1$, where $x_i = ih$, $i = 0,1,\ldots,N$ and $h = 1/N$. Let $U_0, U_1, \ldots, U_N$ be the approximations of $u(x)$ and $M_0, M_1, \ldots, M_N$ be the approximations of $u''(x)$ at $x_0, x_1, \ldots, x_N$. Let $q_i = q(x_i)$ and $f_i = f(x_i)$, $i = 0,1,\ldots,N$.

Consider the IFS $\left\{I \times \mathbb{R}; w_i(x,y) = (L_i(x), F_i(x,y)) : i = 1,2,\ldots,N\right\}$, where $L_i : I \to [x_{i-1}, x_i]$ defined by $L_i(x) = hx + x_{i-1}$, $x \in I$, $F_i : I \times \mathbb{R} \to \mathbb{R}$ defined by $F_i(x,y) = \alpha y + r_i(x)$, $(x,y) \in I \times \mathbb{R}$ with $r_i(x) = A_i + B_i(x - x_0) + C_i \sin \mu(x - x_0) + D_i \cos \mu(x - x_0)$, $\mu > 0$ and α is the scaling factor such that $|\alpha| < h^2$.

We assume the following conditions on the IFS:

$$F_i(x_0, U_0) = U_{i-1},\ F_i(x_N, U_N) = U_i, \quad i = 1, 2, \ldots, N,$$

$$F_{i,1}(x_N, U_{N,1}) = F_{i+1,1}(x_0, U_{0,1}),\ i = 1, 2, \ldots, N-1,$$

$$F_{i,2}(x_0, M_0) = M_{i-1},\ F_{i,2}(x_N, M_N) = M_i, \quad i = 1, 2, \ldots, N,$$

where $F_{i,k}(x, y) = \frac{\alpha y + r_i^{(k)}(x)}{h^k}$, $U_{0,1} = \frac{r_1^{(1)}(x_0)}{h-\alpha}$, $U_{N,1} = \frac{r_N^{(1)}(x_N)}{h-\alpha}$, $k = 1, 2$.

Notice that, the conditions prescribed on the IFS $\Big\{I \times \mathbb{R}; w_i(x, y) = (L_i(x), F_i(x, y)) : i = 1, 2, \ldots, N\Big\}$ satisfy the conditions for constructing two times continuously differentiable fractal functions (see [11]).

Let $\mathscr{F} = \{\phi \in \mathcal{C}^2(I,\ \mathbb{R}) \mid \phi(x_0) = U_0,\ \phi(x_N) = U_N,\ \phi''(x_0) = M_0,\ \phi''(x_N) = M_N\}$. Then $(\mathscr{F}, \rho)$ is a complete metric space where ρ is the metric induced by the norm $\|\phi\| = \|\phi\|_\infty + \|\phi'\|_\infty + \|\phi''\|_\infty$. Define the Read-Bajraktarević operator on $(\mathscr{F}, \rho)$ as

$$T\phi(L_i(x)) = F_i(x, \phi(x)) = \alpha\phi(x) + r_i(x),\ x \in I,\ i = 1, 2, \ldots, N.$$

The fixed-point of T satisfies the functional equation

$$\Phi(L_i(x)) = F_i(x, \Phi(x)) = \alpha\Phi(x) + r_i(x),\ x \in I,\ i = 1, 2, \ldots, N. \tag{6.2}$$

Since $\Phi \in \mathcal{C}^2(I,\ \mathbb{R})$, it satisfies the following functional equation:

$$\Phi''(L_i(x)) = \frac{\alpha\Phi''(x) + r_i''(x)}{h^2},\ x \in I,\ i = 1, 2, \ldots, N. \tag{6.3}$$

The conditions $F_i(x_0, U_0) = U_{i-1}$, $F_i(x_N, U_N) = U_i$, $F_{i,2}(x_0,\ M_0) = M_{i-1}$, $F_{i,2}(x_N, M_N) = M_i$ can be reformulated as $\Phi(x_{i-1}) = U_{i-1}$, $\Phi(x_i) = U_i$, $\Phi''(x_{i-1}) = M_{i-1}$, $\Phi''(x_i) = M_i$ respectively.

Substituting $x = x_0$ and $x = x_N$ in (6.2), we get

$$A_i + D_i = U_{i-1} - \alpha U_0, \tag{6.4}$$

$$A_i + B_i + C_i \sin\mu + D_i \cos\mu = U_i - \alpha U_N. \tag{6.5}$$

Substituting $x = x_0$ and $x = x_N$ in (6.3), we obtain

$$-\frac{D_i\mu^2}{h^2} = M_{i-1} - \frac{\alpha M_0}{h^2}, \tag{6.6}$$

$$-\frac{C_i\mu^2 \sin\mu}{h^2} - \frac{D_i\mu^2 \cos\mu}{h^2} = M_i - \frac{\alpha M_N}{h^2}. \tag{6.7}$$

Solving the system given in (6.4)-(6.7), we get

$$\begin{aligned}
A_i &= \Big[U_{i-1} - \alpha U_0\Big] + \frac{h^2}{\mu^2}\Big[M_{i-1} - \frac{\alpha M_0}{h^2}\Big],\\
B_i &= \Big[U_i - \alpha U_N\Big] - \Big[U_{i-1} - \alpha U_0\Big] - \frac{h^2}{\mu^2}\Big[M_{i-1} - \frac{\alpha M_0}{h^2}\Big] + \frac{h^2}{\mu^2}\Big[M_i - \frac{\alpha M_N}{h^2}\Big],\\
C_i &= \frac{h^2 \cot\mu}{\mu^2}\Big[M_{i-1} - \frac{\alpha M_0}{h^2}\Big] - \frac{h^2 \csc\mu}{\mu^2}\Big[M_i - \frac{\alpha M_N}{h^2}\Big],\\
D_i &= -\frac{h^2}{\mu^2}\Big[M_{i-1} - \frac{\alpha M_0}{h^2}\Big].
\end{aligned}$$

The FIF Φ' satisfies the following functional equation:

$$\Phi'(L_i(x)) = \frac{\alpha\Phi'(x) + r_i'(x)}{h},\ i = 1, 2, \dots, N.$$

The condition $F_{i,1}(x_N, U_{N,1}) = F_{i+1,1}(x_0, U_{0,1}),\ i = 1, 2, \dots, N-1$, can be reformulated as

$$\Phi'(L_i(x_N)) = \Phi'(L_{i+1}(x_0)),\ i = 1, 2, \dots, N-1,$$

which coincides with the standard continuity condition

$$\Phi'(x_i^-) = \Phi'(x_i^+),\ i = 1, 2, \dots, N-1. \tag{6.8}$$

On $[x_{i-1}, x_i]$, we have

$$\Phi'(L_i(x)) = \frac{\alpha\Phi'(x)}{h} + \frac{B_i}{h} + \frac{C_i\mu}{h}\cos\mu(x - x_0) - \frac{D_i\mu}{h}\sin\mu(x - x_0),$$

and on $[x_i, x_{i+1}]$, we get

$$\Phi'(L_{i+1}(x)) = \frac{\alpha\Phi'(x)}{h} + \frac{B_{i+1}}{h} + \frac{C_{i+1}\mu}{h}\cos\mu(x - x_0) - \frac{D_{i+1}\mu}{h}\sin\mu(x - x_0).$$

Making use of (6.8), for each node x_i, $i = 1, 2, \dots, N-1$, we get

$$\frac{\alpha}{h}\Phi'(x_N) + \frac{B_i}{h} + \frac{C_i\mu}{h}\cos\mu - \frac{D_i\mu}{h}\sin\mu = \frac{\alpha}{h}\Phi'(x_0) + \frac{B_{i+1}}{h} + \frac{C_{i+1}\mu}{h},$$

which is simplified as

$$-\frac{\alpha}{h^2}\Phi'(x_0) - \frac{\alpha}{h^2}(\omega_1 + \omega_2)M_0 + \omega_1 M_{i-1} + 2\omega_2 M_i + \omega_1 M_{i+1}$$

$$-\frac{\alpha}{h^2}(\omega_1+\omega_2)M_N+\frac{\alpha}{h^2}\Phi'(x_N)=\frac{U_{i+1}-2U_i+U_{i-1}}{h^2}, \tag{6.9}$$

where $\omega_1=-\frac{1}{\mu^2}+\frac{1}{\mu\sin\mu}$ and $\omega_2=\frac{1}{\mu^2}-\frac{\cos\mu}{\mu\sin\mu}$.

The differential equation (6.1) is discretized at $x=x_i$ as $-\varepsilon M_i+q_iU_i=f_i$, $i=0,1,\ldots,N$. By taking the continuity equation given in (6.9) and substituting

$$M_i=\frac{q_iU_i-f_i}{\varepsilon},\ \Phi'(x_0)=\frac{-3U_0+4U_1-U_2}{2h},\ \Phi'(x_N)=\frac{3U_N-4U_{N-1}+U_{N-2}}{2h},$$

we get the following system:

$$\begin{aligned}&-\Big[\frac{2\alpha\varepsilon}{h}\Big]U_1+\Big[\frac{\alpha\varepsilon}{2h}\Big]U_2+\Big[\omega_1h^2q_{i-1}-\varepsilon\Big]U_{i-1}+\Big[2\omega_2h^2q_i+2\varepsilon\Big]U_i\\&+\Big[\omega_1h^2q_{i+1}-\varepsilon\Big]U_{i+1}+\Big[\frac{\alpha\varepsilon}{2h}\Big]U_{N-2}-\Big[\frac{2\alpha\varepsilon}{h}\Big]U_{N-1}=-(\omega_1+\omega_2)\alpha f_0\\&+\omega_1h^2f_{i-1}+2\omega_2h^2f_i+\omega_1h^2f_{i+1}-(\omega_1+\omega_2)\alpha f_N\\&+\Big[-\frac{3\varepsilon\alpha}{2h}+q_0\alpha(\omega_1+\omega_2)\Big]\eta_0+\Big[-\frac{3\varepsilon\alpha}{2h}+q_N\alpha(\omega_1+\omega_2)\Big]\eta_1,\end{aligned} \tag{6.10}$$

for $i=1,2,\ldots,N-1$, where $U_0=\eta_0$ and $U_N=\eta_1$ are the discretizations for the Dirichlet boundary conditions. The system (6.10) gives the approximations $U_1,U_2,\ldots,U_{N-1}$ of the solution $u(x)$ at $x_1,x_2,\ldots,x_{N-1}$.

If $\alpha=0$, then the system given in (6.10) becomes the system corresponding to the non-polynomial cubic spline as

$$\begin{aligned}&\Big[\omega_1h^2q_{i-1}-\varepsilon\Big]U_{i-1}+\Big[2\omega_2h^2q_i+2\varepsilon\Big]U_i+\Big[\omega_1h^2q_{i+1}-\varepsilon\Big]U_{i+1}\\&=\omega_1h^2f_{i-1}+2\omega_2h^2f_i+\omega_1h^2f_{i+1},\end{aligned}$$

for $i=1,2,\ldots,N-1$.

6.1.1 Convergence analysis

In this section, convergence analysis of the proposed numerical scheme is discussed. Assume that the unknown function $u(x)$ is sufficiently differentiable. By replacing U_i by $u(x_i)$ in (6.10), the truncation error $T_i(h)$, $i=1,2,\ldots,N-1$ associated with the system given in (6.10) is defined as

$$T_i(h)=[r_i^0u(x_0)+r_i^1u(x_1)+r_i^2u(x_2)+r_i^-u(x_{i-1})+r_i^cu(x_i)+r_i^+u(x_{i+1})$$

$$+ r_i^{N-2}u(x_{N-2}) + r_i^{N-1}u(x_{N-1}) + r_i^N u(x_N)] - [q_i^0 f_0 + q_i^- f_{i-1} + q_i^c f_i + q_i^+ f_{i+1} + q_i^N f_N], \tag{6.11}$$

where

$$r_i^0 = -\Big[-\frac{3\varepsilon\alpha}{2h} + q_0\alpha(\omega_1+\omega_2)\Big],\ r_i^1 = -\frac{2\alpha\varepsilon}{h},\ r_i^2 = \frac{\alpha\varepsilon}{2h},\ r_i^- = \omega_1 h^2 q_{i-1} - \varepsilon,$$
$$r_i^c = 2\omega_2 h^2 q_i + 2\varepsilon,\ r_i^+ = \omega_1 h^2 q_{i+1} - \varepsilon,\ r_i^{N-2} = \frac{\alpha\varepsilon}{2h},\ r_i^{N-1} = -\frac{2\alpha\varepsilon}{h},$$
$$r_i^N = -\Big[-\frac{3\varepsilon\alpha}{2h} + q_N\alpha(\omega_1+\omega_2)\Big],\ q_i^0 = -(\omega_1+\omega_2)\alpha,\ q_i^- = \omega_1 h^2,\ q_i^c = 2\omega_2 h^2,$$
$$q_i^+ = \omega_1 h^2,\ q_i^N = -(\omega_1+\omega_2)\alpha.$$

Substituting the differential equation (6.1) in (6.11), we get

$$\begin{aligned} T_i(h) = {} & [r_i^0 u(x_0) + r_i^1 u(x_1) + r_i^2 u(x_2) + r_i^- u(x_{i-1}) + r_i^c u(x_i) + r_i^+ u(x_{i+1}) \\ & + r_i^{N-2}u(x_{N-2}) + r_i^{N-1}u(x_{N-1}) + r_i^N u(x_N)] - [q_i^0(-\varepsilon u''(x_0) + q_0 u(x_0)) \\ & + q_i^-(-\varepsilon u''(x_{i-1}) + q_{i-1}u(x_{i-1})) + q_i^c(-\varepsilon u''(x_i) + q_i u(x_i)) \\ & + q_i^+(-\varepsilon u''(x_{i+1}) + q_{i+1}u(x_{i+1})) + q_i^N(-\varepsilon u''(x_N) + q_N u(x_N))]. \end{aligned} \tag{6.12}$$

After simplifying (6.12), we get

$$\begin{aligned} T_i(h) = & - \alpha\varepsilon\Big[\frac{-3u(x_0) + 4u(x_1) - u(x_2)}{2h}\Big] + \alpha\varepsilon\Big[\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h}\Big] \\ & - \alpha\varepsilon(\omega_1+\omega_2)u''(x_0) - \alpha\varepsilon(\omega_1+\omega_2)u''(x_N) - \varepsilon u(x_{i-1}) + 2\varepsilon u(x_i) \\ & - \varepsilon u(x_{i+1}) + \varepsilon\omega_1 h^2 u''(x_{i-1}) + 2\varepsilon\omega_2 h^2 u''(x_i) + \varepsilon\omega_1 h^2 u''(x_{i+1}). \end{aligned} \tag{6.13}$$

Substituting the Taylor expansions for $u(x_{i-1})$, $u(x_{i+1})$, $u''(x_{i-1})$, $u''(x_{i+1})$ about the point x_i,

$$\frac{-3u(x_0) + 4u(x_1) - u(x_2)}{2h} = u'(x_0) - \frac{h^2}{3}u'''(x_0) - \dots,$$
$$\frac{3u(x_N) - 4u(x_{N-1}) + u(x_{N-2})}{2h} = u'(x_N) - \frac{h^2}{3}u'''(x_N) + \dots$$

in (6.13), we get

$$\begin{aligned} T_i(h) = & - \alpha\varepsilon\Big[u'(x_0) - \frac{h^2}{3}u'''(x_0) + \mathcal{O}(h^3)\Big] + \alpha\varepsilon\Big[u'(x_N) - \frac{h^2}{3}u'''(x_N) + \mathcal{O}(h^3)\Big] \\ & - \alpha\varepsilon(\omega_1+\omega_2)u''(x_0) - \alpha\varepsilon(\omega_1+\omega_2)u''(x_N) + \Big[-1 + 2(\omega_1+\omega_2)\Big]\varepsilon h^2 u''(x_i) \\ & + \Big(\omega_1 - \frac{1}{12}\Big)\varepsilon h^4 u^{(4)}(x_i) + \Big[\frac{\omega_1}{12} - \frac{1}{360}\Big]\varepsilon h^6 u^{(6)}(x_i) + \mathcal{O}(h^8). \end{aligned}$$

- If $|\alpha|< h^4$, $\omega_1 = \frac{1}{6}$ and $\omega_2 = \frac{1}{3}$, we have $|T_i(h)|\leq K_{1i}h^4$ as $h \to 0$, where K_{1i} is a constant. Hence $T_i(h) = \mathcal{O}(h^4)$ as $h \to 0$ for $i = 1, 2, \ldots, N-1$.

- If $|\alpha|< h^6$, $\omega_1 = \frac{1}{12}$ and $\omega_2 = \frac{5}{12}$, we get $|T_i(h)|\leq K_{2i}h^6$ as $h \to 0$, where K_{2i} is a constant. Thus $T_i(h) = \mathcal{O}(h^6)$ as $h \to 0$ for $i = 1, 2, \ldots, N-1$.

The system (6.10) can be written as $AU = d$, where

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{1,N-2} & A_{1,N-1} \\ A_{2,1} & A_{2,2} & A_{2,3} & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{2,N-2} & A_{2,N-1} \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{3,N-2} & A_{3,N-1} \\ A_{4,1} & A_{4,2} & A_{4,3} & A_{4,4} & A_{4,5} & 0 & \ldots & 0 & 0 & 0 & 0 & A_{4,N-2} & A_{4,N-1} \\ A_{5,1} & A_{5,2} & 0 & A_{5,4} & A_{5,5} & A_{5,6} & \ldots & 0 & 0 & 0 & 0 & A_{5,N-2} & A_{5,N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{N-5,1} & A_{N-5,2} & 0 & 0 & 0 & 0 & \ldots & A_{N-5,N-6} & A_{N-5,N-5} & A_{N-5,N-4} & 0 & A_{N-5,N-2} & A_{N-5,N-1} \\ A_{N-4,1} & A_{N-4,2} & 0 & 0 & 0 & 0 & \ldots & 0 & A_{N-4,N-5} & A_{N-4,N-4} & A_{N-4,N-3} & A_{N-4,N-2} & A_{N-4,N-1} \\ A_{N-3,1} & A_{N-3,2} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & A_{N-3,N-4} & A_{N-3,N-3} & A_{N-3,N-2} & A_{N-3,N-1} \\ A_{N-2,1} & A_{N-2,2} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & A_{N-2,N-3} & A_{N-2,N-2} & A_{N-2,N-1} \\ A_{N-1,1} & A_{N-1,2} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{N-1,N-2} & A_{N-1,N-1} \end{bmatrix}$$

$A_{i,j}$ is the coefficient of U_j, $U = (U_1, U_2, \ldots, U_{N-1})^T$, $d = (d_1, d_2, \ldots, d_{N-1})^T$, d_i, $i = 1, 2, \ldots, N-1$ are the right hand side of the system.

In the following proposition, suitable conditions on h and α are derived to make the matrix A to be strictly diagonally dominant. Let us take the following notations:

$$q_* = \min\{q(x) : x \in [0, 1]\}, \quad q^* = \max\{q(x) : x \in [0, 1]\}.$$

Proposition 6.1.1. *Let A be the matrix corresponding to the system (6.10). Let h satisfy $\omega_1 h^2 q^* < \varepsilon$. Then the following conditions on the scaling factor ensure the matrix A would be strictly diagonally dominant:*

$$0 \leq \alpha < \min\left\{h^4, \frac{\omega_2 h^3 q_*}{2\varepsilon}, \frac{\omega_1 h^3 q_*}{\varepsilon}, \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}\right\}$$

when $\omega_1 = \frac{1}{6}$, $\omega_2 = \frac{1}{3}$,

$$0 \leq \alpha < \min\left\{h^6, \frac{\omega_2 h^3 q_*}{2\varepsilon}, \frac{\omega_1 h^3 q_*}{\varepsilon}, \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}\right\}$$

when $\omega_1 = \frac{1}{12}$, $\omega_2 = \frac{5}{12}$.

Proof. Let $\alpha \geq 0$. We get $A_{1,1} = -\frac{2\alpha\varepsilon}{h} + 2h^2\omega_2 q_1 + 2\varepsilon = 2h^2\omega_2 q_1 + 2\varepsilon\left[1 - \frac{\alpha}{h}\right] > 0$ when $\alpha < h^4$, $A_{1,2} = \omega_1 h^2 q_2 - \varepsilon + \frac{\alpha\varepsilon}{2h} < 0$ if $\alpha < \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}$, $A_{1,N-2} = \frac{\alpha\varepsilon}{2h} \geq 0$ and $A_{1,N-1} = -\frac{2\alpha\varepsilon}{h} \leq 0$. Now,

$$\begin{aligned} &|A_{1,1}| - \sum_{j=2}^{N-1} |A_{1,j}| \\ &= 2\omega_2 h^2 q_1 + 2\varepsilon - \frac{2\alpha\varepsilon}{h} + \omega_1 h^2 q_2 - \varepsilon + \frac{\alpha\varepsilon}{2h} - \frac{\alpha\varepsilon}{2h} - \frac{2\alpha\varepsilon}{h} \\ &= \left[\omega_2 h^2 q_1 - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_1 - \frac{2\alpha\varepsilon}{h}\right] + \left[\varepsilon + \omega_1 h^2 q_2\right] > 0 \text{ if } \alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}. \end{aligned}$$

Therefore if $\alpha < \min\left\{h^4, \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}, \frac{\omega_2 h^3 q_*}{2\varepsilon}\right\}$, then $|A_{1,1}| - \sum_{j=2}^{N-1} |A_{1,j}| > 0$.

Observe that $A_{2,1} = -\frac{2\alpha\varepsilon}{h} + [\omega_1 h^2 q_1 - \varepsilon] < 0$ by the assumption on h, $A_{2,2} = \frac{\alpha\varepsilon}{2h} + 2\omega_2 h^2 q_2 + 2\varepsilon > 0$, $A_{2,3} = \omega_1 h^2 q_3 - \varepsilon < 0$ by the assumption on h, $A_{2,N-2} = \frac{\alpha\varepsilon}{2h} \geq 0$ and $A_{2,N-1} = -\frac{2\alpha\varepsilon}{h} \leq 0$. We obtain

$$\begin{aligned} &|A_{2,2}| - \sum_{j=1,\, j\neq 2}^{N-1} |A_{2,j}| \\ &= \frac{\alpha\varepsilon}{2h} + 2\omega_2 h^2 q_2 + 2\varepsilon - \frac{2\alpha\varepsilon}{h} + \omega_1 h^2 q_1 - \varepsilon + \omega_1 h^2 q_3 - \varepsilon - \frac{\alpha\varepsilon}{2h} - \frac{2\alpha\varepsilon}{h} \\ &= \left[\omega_2 h^2 q_2 - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_2 - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_1 h^2 q_1 + \omega_1 h^2 q_3\right] > 0 \end{aligned}$$

if $\alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}$. Therefore if $\alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}$ then $|A_{2,2}| - \sum_{j=1,\, j\neq 2}^{N-1} |A_{2,j}| > 0$.

We have $A_{3,1} = -\frac{2\alpha\varepsilon}{h} \leq 0$, $A_{3,2} = \frac{\alpha\varepsilon}{2h} + \omega_1 h^2 q_2 - \varepsilon < 0$ if $\alpha < \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}$, $A_{3,3} = 2h^2\omega_2 q_3 + 2\varepsilon > 0$, $A_{3,4} = \omega_1 h^2 q_4 - \varepsilon < 0$ if by the assumption on h, $A_{3,N-2} = \frac{\alpha\varepsilon}{2h} \geq 0$ and $A_{3,N-1} = -\frac{2\alpha\varepsilon}{h} \leq 0$. We get

$$\begin{aligned} &|A_{3,3}| - \sum_{j=1,\, j\neq 3}^{N-1} |A_{3,j}| \\ &= 2h^2\omega_2 q_3 + 2\varepsilon - \frac{2\alpha\varepsilon}{h} + \frac{\alpha\varepsilon}{2h} + \omega_1 h^2 q_2 - \varepsilon + \omega_1 h^2 q_4 - \varepsilon - \frac{\alpha\varepsilon}{2h} - \frac{2\alpha\varepsilon}{h} \\ &= \left[\omega_2 h^2 q_3 - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_3 - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_1 h^2 q_2 + \omega_1 h^2 q_4\right] > 0 \end{aligned}$$

if $\alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}$. Thus, if $\alpha < \min\left\{\frac{\omega_2 h^3 q_*}{2\varepsilon}, \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}\right\}$, then $|A_{3,3}| - \sum_{j=1,\, j\neq 3}^{N-1} |A_{3,j}| > 0$.

For $i = 4, 5, \dots, N-4$, we get $A_{i,1} = -\frac{2\alpha\varepsilon}{h} \leq 0$, $A_{i,2} = \frac{\alpha\varepsilon}{2h} \geq 0$, $A_{i,i-1} = \omega_1 h^2 q_{i-1} - \varepsilon < 0$ by the assumption on h, $A_{i,i} = 2\omega_2 h^2 q_i + 2\varepsilon > 0$, $A_{i,i+1} = \omega_1 h^2 q_{i+1} - \varepsilon < 0$ by the assumption on h, $A_{i,N-2} = \frac{\alpha\varepsilon}{2h} \geq 0$ and $A_{i,N-1} = -\frac{2\alpha\varepsilon}{h} \leq 0$. We have

$$
\begin{aligned}
&|A_{i,i}| - \sum_{j=1,\, j\neq i}^{N-1} |A_{i,j}| \\
&= 2\omega_2 h^2 q_i + 2\varepsilon - \frac{2\alpha\varepsilon}{h} - \frac{\alpha\varepsilon}{2h} + \omega_1 h^2 q_{i-1} - \varepsilon + \omega_1 h^2 q_{i+1} - \varepsilon - \frac{\alpha\varepsilon}{2h} - \frac{2\alpha\varepsilon}{h} \\
&= \left[\omega_2 h^2 q_i - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_i - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_1 h^2 q_{i-1} - \frac{\alpha\varepsilon}{h}\right] + \omega_1 h^2 q_{i+1} > 0
\end{aligned}
$$

if $\alpha < \min\left\{\frac{\omega_2 h^3 q_*}{2\varepsilon}, \frac{\omega_1 h^3 q_*}{\varepsilon}\right\}$. Hence, we obtain the condition $|A_{i,i}| - \sum_{j=1,\, j\neq i}^{N-1} |A_{i,j}| > 0$ if $\alpha < \min\left\{\frac{\omega_2 h^3 q_*}{2\varepsilon}, \frac{\omega_1 h^3 q_*}{\varepsilon}\right\}$.

It is obtained that $A_{N-3,1} = -\frac{2\alpha\varepsilon}{h} \leq 0$, $A_{N-3,2} = \frac{\alpha\varepsilon}{2h} \geq 0$, $A_{N-3,N-4} = \omega_1 h^2 q_{N-4} - \varepsilon < 0$ by the assumption on h, $A_{N-3,N-3} = 2h^2 \omega_2 q_{N-3} + 2\varepsilon > 0$, $A_{N-3,N-2} = \omega_1 h^2 q_{N-4} - \varepsilon + \frac{\alpha\varepsilon}{2h} < 0$ if $\alpha < \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}$, $A_{N-3,N-1} = -\frac{2\alpha\varepsilon}{h} \leq 0$. We get

$$
\begin{aligned}
&|A_{N-3,N-3}| - \sum_{j=1,\, j\neq N-3}^{N-1} |A_{N-3,j}| \\
&= 2h^2 \omega_2 q_{N-3} + 2\varepsilon - \frac{2\alpha\varepsilon}{h} - \frac{\alpha\varepsilon}{2h} + \omega_1 h^2 q_{N-4} - \varepsilon + \omega_1 h^2 q_{N-4} - \varepsilon + \frac{\alpha\varepsilon}{2h} - \frac{2\alpha\varepsilon}{h} \\
&= \left[\omega_2 h^2 q_{N-3} - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_{N-3} - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_1 h^2 q_{N-2} + \omega_1 h^2 q_{N-4}\right] > 0
\end{aligned}
$$

if $\alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}$. Thus $|A_{N-3,N-3}| - \sum_{j=1,\, j\neq N-3}^{N-1} |A_{N-3,j}| > 0$ if $\alpha < \min\left\{\frac{\omega_2 h^3 q_*}{2\varepsilon}, \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}\right\}$.

We have $A_{N-2,1} = -\frac{2\alpha\varepsilon}{h} \leq 0$, $A_{N-2,2} = \frac{\alpha\varepsilon}{2h} \geq 0$, $A_{N-2,N-3} = \omega_1 h^2 q_{N-3} - \varepsilon < 0$ by the assumption on h, $A_{N-2,N-2} = 2\omega_2 h^2 q_{N-2} + 2\varepsilon + \frac{\alpha\varepsilon}{2h} > 0$, $A_{N-2,N-1} = \omega_1 h^2 q_{N-1} - \varepsilon - \frac{2\alpha\varepsilon}{h} < 0$ by the assumption on h. We obtain

$$
\begin{aligned}
&|A_{N-2,N-2}| - \sum_{j=1,\, j\neq N-2}^{N-1} |A_{N-2,j}| \\
&= 2\omega_2 h^2 q_{N-2} + 2\varepsilon + \frac{\alpha\varepsilon}{2h} - \frac{2\alpha\varepsilon}{h} - \frac{\alpha\varepsilon}{2h} + \omega_1 h^2 q_{N-3} - \varepsilon + \omega_1 h^2 q_{N-1} - \varepsilon - \frac{2\alpha\varepsilon}{h} \\
&= \left[\omega_2 h^2 q_{N-2} - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_{N-2} - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_1 h^2 q_{N-3} + \omega_1 h^2 q_{N-1}\right] > 0
\end{aligned}
$$

if $\alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}$. Hence, $|A_{N-2,N-2}| - \sum_{j=1,\, j\neq N-2}^{N-1} |A_{N-2,j}| > 0$ if $\alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}$.

We get $A_{N-1,1} = -\frac{2\alpha\varepsilon}{h} \leq 0$, $A_{N-1,2} = \frac{\alpha\varepsilon}{2h} \geq 0$, $A_{N-1,N-2} = \omega_1 h^2 q_{N-2} - \varepsilon + \frac{\alpha\varepsilon}{2h} < 0$ if $\alpha < \frac{2h(\varepsilon-\omega_1 h^2 q^*)}{\varepsilon}$, $A_{N-1,N-1} = 2\omega_2 h^2 q_{N-1} + 2\varepsilon\Big[1 - \frac{\alpha}{h}\Big] > 0$ when $\alpha < h^4$. We obtain

$$\begin{aligned}
&|A_{N-1,N-1}| - \sum_{j=1}^{N-2} |A_{N-1,j}| \\
&= 2\omega_2 h^2 q_{N-1} + 2\varepsilon - \frac{2\varepsilon\alpha}{h} - \frac{2\alpha\varepsilon}{h} - \frac{\alpha\varepsilon}{2h} + \omega_1 h^2 q_{N-2} - \varepsilon + \frac{\alpha\varepsilon}{2h} \\
&= \Big[\omega_2 h^2 q_{N-1} - \frac{2\alpha\varepsilon}{h}\Big] + \Big[\omega_2 h^2 q_{N-1} - \frac{2\alpha\varepsilon}{h}\Big] + \omega_1 h^2 q_{N-2} + \varepsilon > 0
\end{aligned}$$

if $\alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}$. Hence, $|A_{N-1,N-1}| - \sum_{j=1}^{N-2} |A_{N-1,j}| > 0$ if $\alpha < \min\Big\{h^4, \frac{\omega_2 h^3 q_*}{2\varepsilon}, \frac{2h(\varepsilon-\omega_1 h^2 q^*)}{\varepsilon}\Big\}$.

□

Theorem 6.1.1. *Let $u(x)$ be the solution of the differential equation given in (6.1). Let U be the approximate solution obtained by the system given in (6.10). If h and α satisfy the conditions given in Proposition 6.1.1, then as $h \to 0$*

- $\|E\|_\infty = \mathcal{O}(h^2)$ *if* $\omega_1 = \frac{1}{6}$, $\omega_2 = \frac{1}{3}$,
- $\|E\|_\infty = \mathcal{O}(h^4)$ *if* $\omega_1 = \frac{1}{12}$, $\omega_2 = \frac{5}{12}$,

where $E = (E_1, E_2, \ldots, E_{N-1})^T$, $E_i = u(x_i) - U_i$.

Proof. The system (6.10) can be written as

$$AU = d$$

and the system (6.10) with exact solution can be written as

$$A\bar{u} = d + T(h),$$

where $\bar{u} = (u(x_1), u(x_2), \ldots, u(x_{N-1}))^T$, $T(h) = (T_1(h), T_2(h), \ldots, T_{N-1}(h))^T$. Hence, we get

$$A(\bar{u} - U) = T(h), \text{ i.e, } AE = T(h).$$

By taking h and α according to the Proposition 6.1.1, we have

$$E = A^{-1}T(h)$$

and hence we get

$$E_j = \sum_{i=1}^{N-1} A_{j,i}^{-1} T_i(h), \; j = 1, 2, \ldots, N-1.$$

Using the theory of matrices, we have

$$\sum_{i=1}^{N-1} A_{k,i}^{-1} \mathcal{S}_i = 1, \; k = 1, 2, \ldots, N-1,$$

where $\mathcal{S}_i$ is the i-th row sum of the matrix A. The row sums

$$\begin{aligned}
\mathcal{S}_1 =& \; \varepsilon + \left[\omega_2 h^2 q_1 - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_1 - \frac{2\alpha\varepsilon}{h}\right] + \omega_1 h^2 q_2 + \frac{\alpha\varepsilon}{h} > 0 \text{ if } \alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}, \\
\mathcal{S}_i =& \; \left[\omega_2 h^2 q_i - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_i - \frac{2\alpha\varepsilon}{h}\right] + \omega_1 h^2 q_{i-1} \\
&+ \omega_1 h^2 q_{i+1} + \frac{\alpha\varepsilon}{h} > 0 \text{ if } \alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon}, \\
\mathcal{S}_{N-1} =& \; \varepsilon + \left[\omega_2 h^2 q_{N-1} - \frac{2\alpha\varepsilon}{h}\right] + \left[\omega_2 h^2 q_{N-1} - \frac{2\alpha\varepsilon}{h}\right] \\
&+ \omega_1 h^2 q_{N-2} + \frac{\alpha\varepsilon}{h} > 0 \text{ if } \alpha < \frac{\omega_2 h^3 q_*}{2\varepsilon},
\end{aligned}$$

It can be seen that, the conditions given in statement of Proposition 6.1.1 ensures that the row sums of the matrix A are positive. If $\omega_1 = \frac{1}{6}$, $\omega_2 = \frac{1}{3}$, we have

$$\sum_{i=1}^{N-1} A_{k,i}^{-1} \leq \frac{1}{\min_{1 \leq i \leq N-1} \mathcal{S}_i} = \frac{1}{C_{i_1} h^2},$$

where C_{i_1} is a constant. It follows that

$$\left|E_j\right| \leq \frac{K_1 h^4}{C_{i_1} h^2},$$

where K_1 is a constant. Hence we get $\|E\|_\infty = \mathcal{O}(h^2)$ as $h \to 0$.

If $\omega_1 = \frac{1}{12}$, $\omega_2 = \frac{5}{12}$, we have

$$\sum_{i=1}^{N-1} A_{k,i}^{-1} \leq \frac{1}{\min_{1 \leq i \leq N-1} \mathcal{S}_i} = \frac{1}{C_{i_2} h^2},$$

where C_{i_2} is a constant. It follows that

$$\left|E_j\right| \leq \frac{K_2 h^6}{C_{i_2} h^2},$$

where K_2 is a constant. Thus, we have $\|E\|_\infty = \mathcal{O}(h^4)$ as $h \to 0$. □

Remark 6.1.1. *Consider the restrictions on* α *in the Proposition 6.1.1, i.e.,*

$$0 \le \alpha < \min\left\{h^4,\ \frac{\omega_2 h^3 q_*}{2\varepsilon},\ \frac{\omega_1 h^3 q_*}{\varepsilon},\ \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}\right\}$$

when $\omega_1 = \frac{1}{6}$, $\omega_2 = \frac{1}{3}$,

$$0 \le \alpha < \min\left\{h^6,\ \frac{\omega_2 h^3 q_*}{2\varepsilon},\ \frac{\omega_1 h^3 q_*}{\varepsilon},\ \frac{2h(\varepsilon - \omega_1 h^2 q^*)}{\varepsilon}\right\}$$

when $\omega_1 = \frac{1}{12}$, $\omega_2 = \frac{5}{12}$. *When* $\omega_1 = \frac{1}{6}$, $\omega_2 = \frac{1}{3}$, *if* h *is sufficiently small,* h^4 *will be the minimum in these restrictions. When* $\omega_1 = \frac{1}{12}$, $\omega_2 = \frac{5}{12}$, *if* h *is sufficiently small,* h^6 *will be the minimum in these restrictions. For each fixed* h*, by varying* α*, we can compute the numerical solutions of the BVP.*

6.1.2 Numerical examples

In order to see the computational efficiency of the developed method, two numerical examples are considered and numerical results are tabulated. We consider the case $\omega_1 = \frac{1}{12}$, $\omega_2 = \frac{5}{12}$ to compute the numerical solutions of the following BVPs.

Example 6.1.1. *Consider the following BVP [82]:*

$$\begin{cases} -\varepsilon u''(x) + 4u(x) = 4 + 2\sqrt{\varepsilon}\Big[\exp\left(\frac{-x}{\sqrt{\varepsilon}}\right) + \exp\left(\frac{x-1}{\sqrt{\varepsilon}}\right)\Big] - 3(1-x)\exp\left(\frac{-x}{\sqrt{\varepsilon}}\right) \\ \qquad\qquad -3x\exp\left(\frac{x-1}{\sqrt{\varepsilon}}\right),\ x \in (0,1), \\ u(0) = 0, \quad u(1) = 0. \end{cases}$$

The analytical solution is $u(x) = 1 - (1-x)\exp\left(\frac{-x}{\sqrt{\varepsilon}}\right) - x\exp\left(\frac{x-1}{\sqrt{\varepsilon}}\right)$. For each fixed $\varepsilon = 2^{1-i}$, $i = 1, 2, \ldots, 8$, we take $N = 16, 32, 64, 128, 256$ and we get the α restriction in the range $0 \le \alpha < h^6$. Thus, for each fixed ε, the scaling factor is taken as given in Table 6.1. Table 6.2 represents the maximum point-wise errors and the order of convergence. Table 6.3 represents the maximum point-wise errors and the order of convergence corresponding to non-polynomial cubic spline method. Figure 6.1 represents Log-log plots of the maximum errors for the Example 6.1.1.

Table 6.1: Scaling factor for Example 6.1.1.

ε	1	1/2	1/4	1/8	1/16	1/32	1/64	1/128
α	$0.03h^6$	$0.16h^6$	$0.25h^6$	$0.5h^6$	$0.9999h^6$	$0.9999h^6$	$0.9999h^6$	$0.9999h^6$

Table 6.2: Maximum point-wise error and order of convergence corresponding to Example 6.1.1.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256
1	2.5991e-08	1.6374e-09	1.0253e-10	6.3038e-12	3.8969e-14
	3.9885	3.9973	4.0236	7.3378	
1/2	2.7765e-08	1.6029e-09	9.7933e-11	5.9771e-12	5.6843e-14
	4.1145	4.0327	4.0343	6.7163	
1/4	4.7762e-08	3.2265e-09	2.0572e-10	1.2877e-11	6.1218e-13
	3.8878	3.9713	3.9978	4.3947	
1/8	2.1289e-07	1.3919e-08	8.8045e-10	5.5228e-11	3.5728e-12
	3.9349	3.9827	3.9948	3.9503	
1/16	1.0832e-06	7.0972e-08	4.4884e-09	2.8133e-10	1.7530e-11
	3.9320	3.9830	3.9958	4.0044	
1/32	7.8406e-06	5.0099e-07	3.1539e-08	1.9739e-09	1.2348e-10
	3.9681	3.9896	3.9980	3.9987	
1/64	3.2046e-05	2.0554e-06	1.3021e-07	8.1436e-09	5.0928e-10
	3.9626	3.9805	3.9990	3.9992	
1/128	1.2279e-04	7.7344e-06	4.8522e-07	3.0358e-08	1.9009e-09
	3.9887	3.9946	3.9985	3.9973	

Example 6.1.2. *Consider the following BVP [82]:*

$$\begin{cases} -\varepsilon u''(x) + (1 + x(1-x))u(x) = 1 + x(1-x) + [2\sqrt{\varepsilon} - x(1-x)^2]\exp\left(\frac{-x}{\sqrt{\varepsilon}}\right) \\ \qquad\qquad +[2\sqrt{\varepsilon} - x^2(1-x)]\exp\left(\frac{x-1}{\sqrt{\varepsilon}}\right), \ x \in (0,1), \\ u(0) = 0, \quad u(1) = 0. \end{cases}$$

The analytical solution is $u(x) = 1 - (1-x)\exp\left(\frac{-x}{\sqrt{\varepsilon}}\right) - x\exp\left(\frac{x-1}{\sqrt{\varepsilon}}\right)$. For each fixed

Table 6.3: Maximum point-wise error and order of convergence corresponding to Example 6.1.1 with $\alpha = 0$.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256
1	4.5255e-08	2.8288e-09	1.7679e-10	1.0948e-11	2.6684e-13
	3.9998	4.0001	4.0133	5.3585	
1/2	1.6882e-07	1.0555e-08	6.5973e-10	4.1340e-11	2.9809e-12
	3.9996	3.9998	3.9963	3.7937	
1/4	5.2473e-07	3.2815e-08	2.0512e-09	1.2812e-10	7.6469e-12
	3.9992	3.9998	4.0009	4.0665	
1/8	1.3267e-06	8.3014e-08	5.1899e-09	3.2447e-10	2.0535e-11
	3.9984	3.9996	3.9996	3.9819	
1/16	3.4476e-06	2.1593e-07	1.3503e-08	8.4406e-10	5.2660e-11
	3.9969	3.9992	3.9998	4.0026	
1/32	1.0601e-05	6.6504e-07	4.1609e-08	2.6029e-09	1.6277e-10
	3.9945	3.9985	3.9987	3.9992	
1/64	3.4432e-05	2.2360e-06	1.4053e-07	8.8072e-09	5.5050e-10
	3.9448	3.9920	3.9960	3.9999	
1/128	1.2570e-04	7.9133e-06	4.9634e-07	3.1052e-08	1.9430e-09
	3.9896	3.9949	3.9986	3.9984	

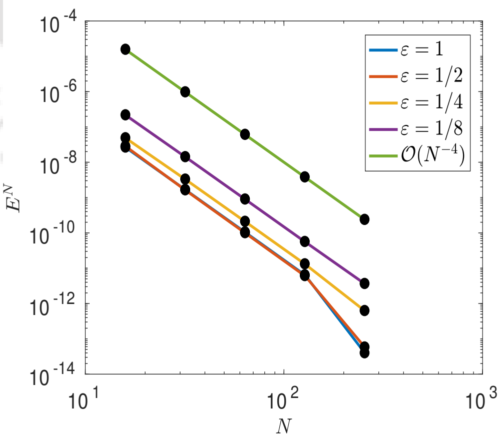


Figure 6.1: Log-log plots of Example 6.1.1.

$\varepsilon = 2^{1-i}$, $i = 1, 2, \ldots, 8$, we take $N = 16, 32, 64, 128, 256$ and we obtain the α restriction in the range $0 \leq \alpha < h^6$. For each fixed ε, the scaling factor is taken as given in Table 6.4. The maximum point-wise errors and the order of convergence are tabulated in Table 6.5. Table 6.6 represents the maximum point-wise errors and the order of convergence corresponding to non-polynomial cubic spline method. Figure 6.2 represents Log-log plots of the maximum errors for the Example 6.1.2.

Table 6.4: Scaling factor for Example 6.1.2.

ε	1	1/2	1/4	1/8	1/16	1/32	1/64	1/128
α	$0.06h^6$	$0.15h^6$	$0.27h^6$	$0.54h^6$	$0.9999h^6$	$0.9999h^6$	$0.9999h^6$	$0.9999h^6$

Table 6.5: Maximum point-wise error and order of convergence corresponding to Example 6.1.2.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256
1	8.5277e-09	5.6511e-10	3.5846e-11	2.2071e-12	2.8200e-14
	3.9156	3.9786	4.0216	6.2903	
1/2	2.1946e-08	1.1885e-09	7.1207e-11	4.3564e-12	3.8414e-14
	4.2067	4.0610	4.0308	6.8254	
1/4	2.8445e-08	2.0975e-09	1.3776e-10	8.6873e-12	4.5525e-13
	3.7614	3.9284	3.9871	4.2542	
1/8	2.5123e-07	1.3186e-08	7.8115e-10	4.8072e-11	3.0113e-12
	4.2519	4.0773	4.0223	3.9967	
1/16	1.4879e-06	9.7308e-08	6.1716e-09	3.8758e-10	2.4198e-11
	3.9346	3.9788	3.9931	4.0015	
1/32	1.4699e-05	9.2969e-07	5.8521e-08	3.6608e-09	2.2900e-10
	3.9828	3.9897	3.9987	3.9987	
1/64	6.4481e-05	4.0751e-06	2.5815e-07	1.6147e-08	1.0093e-09
	3.9840	3.9805	3.9989	3.9998	
1/128	2.4250e-04	1.5730e-05	9.9775e-07	6.2429e-08	3.9058e-09
	3.9464	3.9787	3.9984	3.9985	

Table 6.6: Maximum point-wise error and order of convergence corresponding to Example 6.1.2 with $\alpha = 0$.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256
1	5.7074e-08	3.5676e-09	2.2297e-10	1.3895e-11	7.0982e-13
	3.9998	4.0000	4.0042	4.2909	
1/2	2.4938e-07	1.5591e-08	9.7451e-10	6.0953e-11	4.0483e-12
	3.9996	3.9999	3.9989	3.9123	
1/4	9.5497e-07	5.9721e-08	3.7331e-09	2.3328e-10	1.4368e-11
	3.9991	3.9998	4.0002	4.0212	
1/8	3.0550e-06	1.9116e-07	1.1951e-08	7.4705e-10	4.6880e-11
	3.9983	3.9996	3.9998	3.9942	
1/16	7.9343e-06	4.9702e-07	3.1084e-08	1.9432e-09	1.2131e-10
	3.9967	3.9991	3.9997	4.0017	
1/32	2.2195e-05	1.3932e-06	8.7462e-08	5.4679e-09	3.4183e-10
	3.9937	3.9936	3.9996	3.9996	
1/64	7.3085e-05	4.6055e-06	2.8879e-07	1.8089e-08	1.1307e-09
	3.9881	3.9953	3.9968	3.9998	
1/128	2.4975e-04	1.6298e-05	1.0297e-06	6.4479e-08	4.0339e-09
	3.9377	3.9844	3.9973	3.9986	

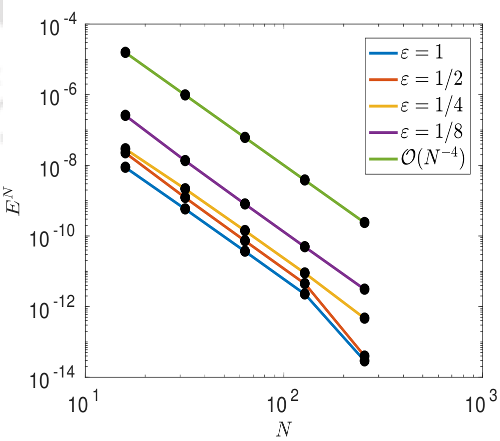


Figure 6.2: Log-log plots of Example 6.1.2.

6.2 Nonself-Adjoint Boundary-Value Problems

Let us consider the singularly perturbed BVPs of the form

$$\begin{cases} \varepsilon u''(x) = p(x)u'(x) + q(x)u(x) + f(x), \ x \in (0,1), \\ u(0) = \eta_0, \quad u(1) = \eta_1, \end{cases} \tag{6.14}$$

where $0 < \varepsilon \leq 1$, functions p, q, f are sufficiently smooth, $p(x) > 0$ or $p(x) < 0$ and $q(x) > 0$ for $x \in [0,1]$.

Let us consider a uniform mesh on the interval $I = [0,1]$ such that $0 = x_0 < x_1 < \cdots < x_N = 1$, where $x_i = ih$, $i = 0, 1, \ldots, N$, $h = 1/N$. Let U_i be the approximation of $u(x_i)$, U_i' be the approximation of $u'(x_i)$ and M_i be the approximation of $u''(x_i)$.

The differential equation given in (6.14) is discretized at $x = x_i$ as $\varepsilon M_i = p_i U_i' + q_i U_i + f_i$, where $p_i = p(x_i)$, $q_i = q_i$, $f_i = f_i$. Substitute $M_i = \frac{p_i U_i' + q_i U_i + f_i}{\varepsilon}$ and the approximations

$$\begin{aligned}
U_{i+1}' &= \frac{3U_{i+1} - 4U_i + U_{i-1}}{2h}, \ U_{i-1}' = \frac{-U_{i+1} + 4U_i - 3U_{i-1}}{2h}, \\
U_i' &= \frac{U_{i+1} - U_{i-1}}{2h} + \frac{[2h^2\sigma q_{i+1} + h\sigma(3p_{i+1} + p_{i-1})]U_{i+1}}{2h} - 2\sigma(p_{i+1} + p_{i-1})U_i \\
&\quad - \frac{[2h^2\sigma q_{i-1} - h\sigma(p_{i+1} + 3p_{i-1})]U_{i-1}}{2h} + h\sigma(f_{i+1} - f_{i-1}), \ \sigma \in \mathbb{R}, \\
\Phi'(x_0) &= \frac{-U_2 + 4U_1 - 3U_0}{2h}, \ \Phi'(x_N) = \frac{3U_N - 4U_{N-1} + U_{N-2}}{2h},
\end{aligned}$$

in (6.9), we get the following system:

$$\begin{aligned}
&-\frac{2}{h}\Big[\alpha\varepsilon + \alpha(\omega_1 + \omega_2)p_0\Big]U_1 + \frac{1}{2h}\Big[\alpha\varepsilon + \alpha(\omega_1 + \omega_2)p_0\Big]U_2 + \Big[-\varepsilon + \omega_1 h^2 q_{i-1} \\
&- \frac{3\omega_1 h p_{i-1}}{2} - \omega_2 h p_i + \frac{\omega_1 h p_{i+1}}{2} - 2h^3\omega_2\sigma p_i q_{i-1} + h^2\omega_2\sigma p_i(p_{i+1} + 3p_{i-1})\Big]U_{i-1} \\
&+ \Big[2\varepsilon + 2\omega_2 h^2 q_i + 2\omega_1 h p_{i-1} - 2\omega_1 h p_{i+1} - 4h^2\omega_2\sigma p_i(p_{i+1} + p_{i-1})\Big]U_i + \Big[-\varepsilon \\
&+ \omega_1 h^2 q_{i+1} - \frac{\omega_1 h p_{i-1}}{2} + \omega_2 h p_i + \frac{3\omega_1 h p_{i+1}}{2} + 2h^3\omega_2\sigma p_i q_{i+1} + h^2\omega_2\sigma p_i(3p_{i+1} \\
&+ p_{i-1})\Big]U_{i+1} + \frac{1}{2h}\Big[\alpha\varepsilon - \alpha(\omega_1 + \omega_2)p_N\Big]U_{N-2} - \frac{2}{h}\Big[\alpha\varepsilon - \alpha(\omega_1 + \omega_2)p_N\Big]U_{N-1} \\
&= -\frac{3}{2h}\Big[\alpha\varepsilon + \alpha(\omega_1 + \omega_2)p_0\Big]\eta_0 + \alpha(\omega_1 + \omega_2)q_0\eta_0 + \alpha(\omega_1 + \omega_2)f_0 - h^2\omega_1 f_{i-1} \\
&- 2h^2\omega_2 f_i - h^2\omega_1 f_{i+1} - 2h^3\omega_2\sigma p_i(f_{i+1} - f_{i-1}) + \alpha(\omega_1 + \omega_2)f_N
\end{aligned}$$

$$-\frac{3}{2h}\Big[-\alpha(\omega_1+\omega_2)p_N+\alpha\varepsilon\Big]\eta_1+\alpha(\omega_1+\omega_2)q_N\eta_1,\ i=1,2,\ldots,N-1, \tag{6.15}$$

where $U_0=\eta_0$ and $U_N=\eta_1$ are the discretizations of the Dirichlet boundary conditions. The system (6.15) gives the approximations $U_1, U_2, \ldots, U_{N-1}$ of the solution $u(x)$ at $x_1, x_2, \ldots, x_{N-1}$.

Remark 6.2.1. *If $\alpha=0$, then the system given in (6.15) becomes to the system that corresponds to the non-polynomial cubic spline as*

$$\Big[-\varepsilon+\omega_1h^2q_{i-1}-\frac{3\omega_1hp_{i-1}}{2}-\omega_2hp_i+\frac{\omega_1hp_{i+1}}{2}-2h^3\omega_2\sigma p_iq_{i-1}+h^2\omega_2\sigma p_i(p_{i+1}$$
$$+3p_{i-1})\Big]U_{i-1}+\Big[2\varepsilon+2\omega_2h^2q_i+2\omega_1hp_{i-1}-2\omega_1hp_{i+1}-4h^2\omega_2\sigma p_i(p_{i+1}+p_{i-1})\Big]U_i$$
$$+\Big[-\varepsilon+\omega_1h^2q_{i+1}-\frac{\omega_1hp_{i-1}}{2}+\omega_2hp_i+\frac{3\omega_1hp_{i+1}}{2}+2h^3\omega_2\sigma p_iq_{i+1}+h^2\omega_2\sigma p_i(3p_{i+1}$$
$$+p_{i-1})\Big]U_{i+1}=-h^2\omega_1f_{i-1}-2h^2\omega_2f_i-h^2\omega_1f_{i+1}-2h^3\omega_2\sigma p_i(f_{i+1}-f_{i-1}),$$
$$i=1,2,\ldots,N-1.$$

6.2.1 Convergence analysis

In this section, convergence analysis of the developed numerical scheme is discussed. Let $|\alpha|< h^6$. Assume that the unknown function $u(x)$ is sufficiently differentiable. Substitute the original function value $u(x_i)$ instead of the approximate value U_i in (6.15) and hence we get the truncation errors $T_i(h)$, $i=1,2,\ldots,N-1$ associated with the equations given in (6.15) as

$$\begin{aligned}T_i(h)=&[r_i^0u(x_0)+r_i^1u(x_1)+r_i^2u(x_2)+r_i^-u(x_{i-1})+r_i^cu(x_i)+r_i^+u(x_{i+1})\\&+r_i^{N-2}u(x_{N-2})+r_i^{N-1}u(x_{N-1})+r_i^Nu(x_N)]-[q_i^0f_0+q_i^-f_{i-1}\\&+q_i^cf_i+q_i^+f_{i+1}+q_i^Nf_N],\end{aligned} \tag{6.16}$$

where

$$r_i^0=\frac{3}{2h}\Big[\alpha\varepsilon+\alpha(\omega_1+\omega_2)p_0\Big]-\alpha(\omega_1+\omega_2)q_0,\ r_i^1=-\frac{2}{h}\Big[\alpha\varepsilon+\alpha(\omega_1+\omega_2)p_0\Big],$$
$$r_i^2=\frac{1}{2h}\Big[\alpha\varepsilon+\alpha(\omega_1+\omega_2)p_0\Big],\ r_i^-=-\varepsilon+\omega_1h^2q_{i-1}-\frac{3\omega_1hp_{i-1}}{2}-\omega_2hp_i+\frac{\omega_1hp_{i+1}}{2}$$
$$-2h^3\omega_2\sigma p_iq_{i-1}+h^2\omega_2\sigma p_i(p_{i+1}+3p_{i-1}),\ r_i^c=2\varepsilon+2\omega_2h^2q_i+2\omega_1hp_{i-1}-2\omega_1hp_{i+1}$$

$$- 4h^2\omega_2\sigma p_i(p_{i+1}+p_{i-1}),\ r_i^{+} = -\varepsilon + \omega_1 h^2 q_{i+1} - \frac{\omega_1 h p_{i-1}}{2} + \omega_2 h p_i + \frac{3\omega_1 h p_{i+1}}{2}$$

$$+ 2h^3\omega_2\sigma p_i q_{i+1} + h^2\omega_2\sigma p_i(3p_{i+1}+p_{i-1}),\ r_i^{N-2} = \frac{1}{2h}\Big[\alpha\varepsilon - \alpha(\omega_1+\omega_2)p_N\Big],$$

$$r_i^{N-1} = -\frac{2}{h}\Big[\alpha\varepsilon - \alpha(\omega_1+\omega_2)p_N\Big],\ r_i^{N} = \frac{3}{2h}\Big[-\alpha(\omega_1+\omega_2)p_N + \alpha\varepsilon\Big] - \alpha(\omega_1+\omega_2)q_N,$$

$$q_i^0 = \alpha(\omega_1+\omega_2),\ q_i^{-} = -h^2\omega_1 + 2h^3\omega_2\sigma p_i,\ q_i^{c} = -2h^2\omega_2,\ q_i^{+} = -h^2\omega_1 - 2h^3\omega_2\sigma p_i,$$

$$q_i^N = \alpha(\omega_1+\omega_2).$$

Substituting the differential equation (6.14) in (6.16), we get

$$\begin{aligned}
T_i(h) =& \Big[r_i^0 u(x_0) + r_i^1 u(x_1) + r_i^2 u(x_2) + r_i^{-} u(x_{i-1}) + r_i^{c} u(x_i) + r_i^{+} u(x_{i+1}) \\
&+ r_i^{N-2} u(x_{N-2}) + r_i^{N-1} u(x_{N-1}) + r_i^{N} u(x_N)\Big] - \Big[q_i^0(\varepsilon u''(x_0) \\
&- p_0 u'(x_0) - q_0 u(x_0)) + q_i^{-}(\varepsilon u''(x_{i-1}) - p_{i-1}u'(x_{i-1}) \\
&- q_{i-1}u(x_{i-1})) + q_i^{c}(\varepsilon u''(x_i) - p_i u'(x_i) - q_i u(x_i)) \\
&+ q_i^{+}(\varepsilon u''(x_{i+1}) - p_{i+1}u'(x_{i+1}) - q_{i+1}u(x_{i+1})) \\
&+ q_i^{N}(\varepsilon u''(x_N) - p_N u'(x_N) - q_N u(x_N))\Big].
\end{aligned} \tag{6.17}$$

After simplifying (6.17), we obtain

$$\begin{aligned}
T_i(h) =& -\alpha\varepsilon\Big[\frac{-3u(x_0)+4u(x_1)-u(x_2)}{2h}\Big] - \alpha(\omega_1+\omega_2)\varepsilon u''(x_0) \\
&+ \alpha\varepsilon\Big[\frac{3u(x_N)-4u(x_{N-1})+u(x_{N-2})}{2h}\Big] - \alpha(\omega_1+\omega_2)\varepsilon u''(x_N) \\
&- \alpha(\omega_1+\omega_2)p_0\Big[\frac{-3u(x_0)+4u(x_1)-u(x_2)}{2h} - u'(x_0)\Big] \\
&- \alpha(\omega_1+\omega_2)p_N\Big[\frac{3u(x_N)-4u(x_{N-1})+u(x_{N-2})}{2h} - u'(x_N)\Big] \\
&+ \Big[-\varepsilon - \frac{3\omega_1 h p_{i-1}}{2} - \omega_2 h p_i + \frac{\omega_1 h p_{i+1}}{2} + h^2\omega_2\sigma p_i(p_{i+1}+3p_{i-1})\Big]u(x_{i-1}) \\
&+ \Big[2\varepsilon + 2\omega_1 h p_{i-1} - 2\omega_1 h p_{i+1} - 4h^2\omega_2\sigma p_i(p_{i+1}+p_{i-1})\Big]u(x_i) + \Big[-\varepsilon \\
&- \frac{\omega_1 h p_{i-1}}{2} + \omega_2 h p_i + \frac{3\omega_1 h p_{i+1}}{2} + h^2\omega_2\sigma p_i(3p_{i+1}+p_{i-1})\Big]u(x_{i+1}) \\
&+ (h^2\omega_1 - 2h^3\omega_2\sigma p_i)\Big[\varepsilon u''(x_{i-1}) - p_{i-1}u'(x_{i-1})\Big] + 2h^2\omega_2\Big[\varepsilon u''(x_i) \\
&- p_i u'(x_i)\Big] + (h^2\omega_1 + 2h^3\omega_2\sigma p_i)\Big[\varepsilon u''(x_{i+1}) - p_{i+1}u'(x_{i+1})\Big].
\end{aligned} \tag{6.18}$$

Using the Taylor expansions for $u(x_{i-1})$, $u(x_{i+1})$, $u'(x_{i-1})$, $u'(x_{i+1})$, $u''(x_{i-1})$, $u''(x_{i+1})$ about the point x_i,

$$\begin{aligned}
\frac{-3u(x_0)+4u(x_1)-u(x_2)}{2h} &= u'(x_0) - \frac{h^2}{3}u'''(x_0) - \dots, \\
\frac{3u(x_N)-4u(x_{N-1})+u(x_{N-2})}{2h} &= u'(x_N) - \frac{h^2}{3}u'''(x_N) + \dots
\end{aligned}$$

in (6.18), we get

$$\begin{aligned}
T_i(h) =& -\alpha\varepsilon\Big[u'(x_0) - \frac{h^2}{3}u'''(x_0) + \mathcal{O}(h^3)\Big] + \alpha\varepsilon\Big[u'(x_N) - \frac{h^2}{3}u'''(x_N) + \mathcal{O}(h^3)\Big] \\
& - \alpha(\omega_1+\omega_2)p_0\Big[-\frac{h^2}{3}u'''(x_0) + \mathcal{O}(h^3)\Big] - \alpha(\omega_1+\omega_2)p_N\Big[-\frac{h^2}{3}u'''(x_N) + \mathcal{O}(h^3)\Big] \\
& - \alpha(\omega_1+\omega_2)\varepsilon\Big[u''(x_0) + u''(x_N)\Big] + \varepsilon h^2\Big[2\omega_1 + 2\omega_2 - 1\Big]u''(x_i) \\
& + \frac{h^4}{3}\Big[(1+12\varepsilon\sigma)\omega_2 p_i - \omega_1(p_{i-1}+p_{i+1}) + 2h\omega_2\sigma(p_ip_{i-1} - p_ip_{i+1})\Big]u'''(x_i) \\
& + h^4\Big[\frac{\varepsilon(12\omega_1-1)}{12} + \frac{h\omega_1 p_{i-1}}{12} - \frac{h\omega_1 p_{i+1}}{12} - \frac{h^2\omega_2\sigma p_i(p_{i-1}+p_{i+1})}{6}\Big]u^{(4)}(x_i) \\
& + \mathcal{O}(h^6).
\end{aligned}$$

Let

$$\begin{aligned}
p_{i-1} &= p_i - hp_i' + \frac{h^2}{2!}p_i'' + \dots, \\
p_{i+1} &= p_i + hp_i' + \frac{h^2}{2!}p_i'' + \dots.
\end{aligned}$$

It follows that

$$\begin{aligned}
T_i(h) =& -\alpha\varepsilon\Big[u'(x_0) - \frac{h^2}{3}u'''(x_0) + \mathcal{O}(h^3)\Big] + \alpha\varepsilon\Big[u'(x_N) - \frac{h^2}{3}u'''(x_N) + \mathcal{O}(h^3)\Big] \\
& - \alpha(\omega_1+\omega_2)p_0\Big[-\frac{h^2}{3}u'''(x_0) + \mathcal{O}(h^3)\Big] - \alpha(\omega_1+\omega_2)p_N\Big[-\frac{h^2}{3}u'''(x_N) \\
& + \mathcal{O}(h^3)\Big] - \alpha(\omega_1+\omega_2)\varepsilon\Big[u''(x_0) + u''(x_N)\Big] + \varepsilon h^2\Big[2\omega_1 + 2\omega_2 - 1\Big]u''(x_i) \\
& + \frac{h^4}{3}\Big[(1+12\varepsilon\sigma)\omega_2 - 2\omega_1\Big]p_i u'''(x_i) + h^4\Big[\frac{\varepsilon(12\omega_1-1)}{12}\Big]u^{(4)}(x_i) + \mathcal{O}(h^6).
\end{aligned}$$

Since $|\alpha| < h^6$, if $\omega_1 = \frac{1}{12}$, $\omega_2 = \frac{5}{12}$ and $\sigma = -\frac{1}{20\varepsilon}$, we get $|T_i(h)| \le K_i h^6$ as $h \to 0$, where K_i is a constant. Thus $T_i(h) = \mathcal{O}(h^6)$ as $h \to 0$ for $i = 1, 2, \dots, N-1$.

The system (6.15) can be written as $AU = d$, where

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{1,N-2} & A_{1,N-1} \\ A_{2,1} & A_{2,2} & A_{2,3} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{2,N-2} & A_{2,N-1} \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{3,N-2} & A_{3,N-1} \\ A_{4,1} & A_{4,2} & A_{4,3} & A_{4,4} & A_{4,5} & 0 & \dots & 0 & 0 & 0 & 0 & A_{4,N-2} & A_{4,N-1} \\ A_{5,1} & A_{5,2} & 0 & A_{5,4} & A_{5,5} & A_{5,6} & \dots & 0 & 0 & 0 & 0 & A_{5,N-2} & A_{5,N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{N-5,1} & A_{N-5,2} & 0 & 0 & 0 & 0 & \dots & A_{N-5,N-6} & A_{N-5,N-5} & A_{N-5,N-4} & 0 & A_{N-5,N-2} & A_{N-5,N-1} \\ A_{N-4,1} & A_{N-4,2} & 0 & 0 & 0 & 0 & \dots & 0 & A_{N-4,N-5} & A_{N-4,N-4} & A_{N-4,N-3} & A_{N-4,N-2} & A_{N-4,N-1} \\ A_{N-3,1} & A_{N-3,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & A_{N-3,N-4} & A_{N-3,N-3} & A_{N-3,N-2} & A_{N-3,N-1} \\ A_{N-2,1} & A_{N-2,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & A_{N-2,N-3} & A_{N-2,N-2} & A_{N-2,N-1} \\ A_{N-1,1} & A_{N-1,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A_{N-1,N-2} & A_{N-1,N-1} \end{bmatrix},$$

$A_{i,j}$ is the coefficient of U_j, $U = (U_1, U_2, \dots, U_{N-1})^T$, $d = (d_1, d_2, \dots, d_{N-1})^T$, d_i, $i = 1, 2, \dots, N-1$ are the right hand side of the system.

Let B be the matrix corresponding to $\alpha = 0$. By taking the matrices A, B, $n = N-1$, $\|.\|_\infty$ norm in the Theorem 1.7.4, we get

$$\max_i |\lambda_i' - \mu_i'| \leq 2^{\frac{2N-3}{N-1}} (N-1)^{\frac{1}{N-1}} (2Q)^{\frac{N-2}{N-1}} \|A - B\|_\infty^{\frac{1}{N-1}} \tag{6.19}$$

where $Q = \max\{\|A\|_\infty, \ \|B\|_\infty\}$, λ_i' and μ_i', $i = 1, 2, \dots, N-1$ are the eigenvalues of A and B respectively.

Note that, when h is sufficiently small, the matrix B is irreducible, $B_{i,i} > 0$, $B_{i,j} \leq 0$, $i \neq j$; $i, j = 1, 2, \dots, N-1$, the row sums of B are such that

$$\begin{aligned} \mathcal{R}_1 &= \varepsilon + 2\omega_2 h^2 q_1 + \omega_1 h^2 q_2 + \frac{3\omega_1 h p_0}{2} + \omega_2 h p_1 - \frac{\omega_1 h p_2}{2} + 2h^3 \omega_2 \sigma p_1 q_2 \\ &\quad - h^2 \omega_2 \sigma p_1 p_2 - 3h^2 \omega_2 \sigma p_1 p_0 > 0, \\ \mathcal{R}_i &= \omega_1 h^2 q_{i-1} + 2\omega_2 h^2 q_i + \omega_1 h^2 q_{i+1} - 2h^3 \omega_2 \sigma p_i q_{i-1} + 2h^3 \omega_2 \sigma p_i q_{i+1} > 0, \\ &\qquad i = 2, 3, \dots, N-2, \\ \mathcal{R}_{N-1} &= \varepsilon + 2\omega_2 h^2 q_{N-1} + \omega_1 h^2 q_{N-2} + \frac{\omega_1 h p_{N-2}}{2} - \omega_2 h p_{N-1} - \frac{3\omega_1 h p_N}{2} \\ &\quad - 2h^3 \omega_2 \sigma p_{N-1} q_{N-2} - 3h^2 \omega_2 \sigma p_{N-1} p_N - h^2 \omega_2 \sigma p_{N-1} p_{N-2} > 0. \end{aligned}$$

Thus, B is a monotone matrix [67] and hence B^{-1} exists and the eigenvalues μ_i', $i = 1, 2, \dots, N-1$ of B are non-zero. Thus, when h is sufficiently small, B is invertible and its row sums are positive. Once we fix h (which gives B is monotone), then the eigenvalues of B are non-zero. Now α can vary in the region $(-h^6, h^6)$. We can choose sufficiently small α in the region $(-h^6, h^6)$ to satisfy the following conditions:

- A is invertible because $\|A-B\|_\infty = \left(\left|\frac{-2\alpha}{h}\right| + \left|\frac{\alpha}{2h}\right|\right)\left|\varepsilon + (\omega_1+\omega_2)p_0\right| + \left(\left|\frac{-2\alpha}{h}\right| + \left|\frac{\alpha}{2h}\right|\right)\left|\varepsilon - (\omega_1+\omega_2)p_N\right|$ and from (6.19), it can been seen that eigenvalues of A are non-zero when α is sufficiently small.

- The row sums of the matrix A are

$$\mathcal{S}_i = \mathcal{R}_i - \frac{3\alpha\varepsilon}{h} - \frac{3\alpha(\omega_1+\omega_2)p_0}{2h} + \frac{3\alpha(\omega_1+\omega_2)p_N}{2h} > 0,\ i = 1, 2, \ldots, N-1$$

because $\mathcal{R}_i > 0$, $i = 1, 2, \ldots, N-1$ and we can choose α sufficiently small so that $\mathcal{S}_i > 0$, $i = 1, 2, \ldots, N-1$.

Hence by choosing sufficiently small h (which gives B is monotone), choosing α in $(-h^6, h^6)$ (which gives A is invertible and row sums of A are positive) we get the following:

It can be seen that $AE = T(h)$ where $T(h) = (T_1(h), T_2(h), \ldots, T_{N-1}(h))^T$, $E = (E_1, E_2, \ldots, E_{N-1})^T$, $E_i = u(x_i) - U_i$.

We get $E = A^{-1}T(h)$ and hence we have

$$E_j = \sum_{i=1}^{N-1} A^{-1}_{j,i} T_i(h),\ j = 1, 2, \ldots, N-1.$$

From the theory of matrices, we have

$$\sum_{i=1}^{N-1} A^{-1}_{k,i} \mathcal{S}_i = 1,\ k = 1, 2, \ldots, N-1,$$

where $A^{-1}_{k,i}$ is the (k,i)-th element of the matrix A^{-1}. Note that $\min_{1\le i\le N-1} \mathcal{S}_i = C_{i_0} h^2$, where C_{i_0} is a constant. Hence, we have

$$\sum_{i=1}^{N-1} A^{-1}_{k,i} \le 1/\min_{1\le i\le N-1} \mathcal{S}_i = 1/C_{i_0} h^2.$$

We get $|E_j| \le \frac{Kh^6}{C_{i_0}h^2}$, where K is a constant. Therefore, as $h \to 0$,

$$\|E\|_\infty = \mathcal{O}(h^4).$$

6.2.2 Numerical examples

To see the efficiency of the proposed method, numerical experiments are done in this section.

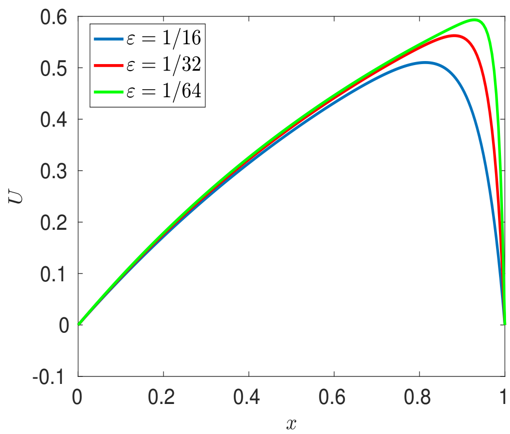
Example 6.2.1. *Consider the BVP [87]*

$$\begin{cases} \varepsilon u''(x) = u'(x) + u(x) - 1, \ x \in (0,1), \\ u(0) = 0, \quad u(1) = 0. \end{cases}$$

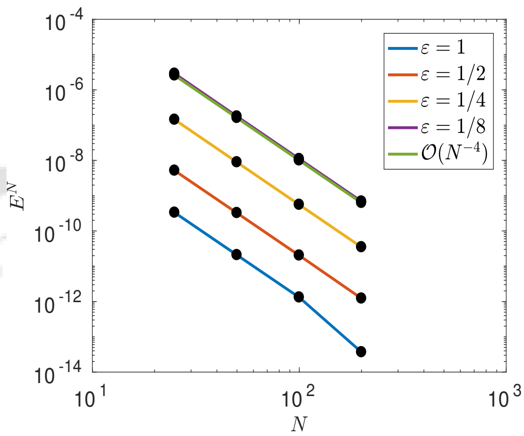
The exact solution is $u(x) = \frac{(e^{\lambda_2}-1)e^{\lambda_1 x}}{e^{\lambda_1}-e^{\lambda_2}} + \frac{(1-e^{\lambda_1})e^{\lambda_2 x}}{e^{\lambda_1}-e^{\lambda_2}} + 1$, where $\lambda_1 = \frac{1+\sqrt{1+4\varepsilon}}{2\varepsilon}$, $\lambda_2 = \frac{1-\sqrt{1+4\varepsilon}}{2\varepsilon}$. The scaling factors which are used to compute the numerical solutions are given in Table 6.7. Table 6.8 represents the maximum point-wise error and the order of convergence. Table 6.9 represents the maximum point-wise error and the order of convergence for the non-polynomial cubic spline method. Figure 6.3 represents the numerical solutions and Log-log plots of the maximum errors for the Example 6.2.1.

Table 6.7: Scaling factor for Example 6.2.1.

ε	1	1/2	1/4	1/8	1/16	1/32	1/64
α	$-0.01h^6$	$-0.1h^6$	$-0.3h^6$	$-0.9999h^6$	$-0.9999h^6$	$-0.9999h^6$	$-0.9999h^6$



(a) Numerical solutions of Example 6.2.1.



(b) Log-log plots of Example 6.2.1.

Figure 6.3: Numerical results of Example 6.2.1.

Table 6.8: Maximum point-wise error and order of convergence corresponding to Example 6.2.1.

$\downarrow \varepsilon / N \rightarrow$	25	50	100	200
1	3.3063e-10	2.0751e-11	1.3070e-12	3.6610e-14
	3.9940	3.9889	5.1579	
1/2	5.1332e-09	3.2294e-10	2.0195e-11	1.2187e-12
	3.9905	3.9992	4.0506	
1/4	1.4261e-07	8.8653e-09	5.5361e-10	3.4688e-11
	4.0078	4.0012	3.9964	
1/8	2.8749e-06	1.7614e-07	1.0970e-08	6.8549e-10
	4.0287	4.0051	4.0003	
1/16	5.9115e-05	3.8044e-06	2.3612e-07	1.4749e-08
	3.9578	4.0101	4.0009	
1/32	1.0044e-03	5.8134e-05	3.7214e-06	2.3141e-07
	4.1108	3.9655	4.0073	
1/64	1.1048e-02	9.6743e-04	5.6226e-05	3.5805e-06
	3.5135	4.1048	3.9730	

Table 6.9: Maximum point-wise error and order of convergence corresponding to Example 6.2.1 with $\alpha = 0$.

$\downarrow \varepsilon / N \rightarrow$	25	50	100	200
1	7.1291e-10	4.4623e-11	2.7997e-12	1.0859e-13
	3.9978	3.9944	4.6883	
1/2	1.9875e-08	1.2437e-09	7.7718e-11	4.9231e-12
	3.9983	4.0003	3.9806	
1/4	3.3081e-07	2.0641e-08	1.2896e-09	8.0731e-11
	4.0024	4.0005	3.9977	
1/8	4.6351e-06	2.8807e-07	1.8020e-08	1.1261e-09
	4.0081	3.9987	4.0002	
1/16	6.3291e-05	4.0807e-06	2.5375e-07	1.5839e-08
	3.9551	4.0073	4.0018	
1/32	1.0129e-03	5.8770e-05	3.7579e-06	2.3374e-07
	4.1073	3.9671	4.0070	
1/64	1.1063e-02	9.6853e-04	5.6308e-05	3.5852e-06
	3.5138	4.1044	3.9732	

6.3 Conclusion

In this chapter, fractal non-polynomial cubic spline methods are developed to solve the both self-adjoint and nonself-adjoint singularly perturbed BVPs. Convergence analysis of the developed methods are discussed and it is shown that the developed numerical methods have fourth-order convergent. Test examples are solved to show the efficiency of the developed method. The important feature of the proposed scheme is that, it has fourth-order convergence for nonself-adjoint BVPs. We obtained the fourth-order method for the nonself-adjoint BVPs by modifying the derivative approximation.

Chapter 7

Fractal Quintic Spline Solutions for Boundary-Value Problems

In this chapter, we have developed the numerical methods for solving self-adjoint singularly perturbed BVPs and the non-linear BVPs. Aziz and Khan [8] developed the fourth-order convergent numerical method using quintic spline and Khan et al. [81] proposed the fifth-order convergent numerical method using sextic spline to get the numerical approximations for self-adjoint singularly perturbed BVPs. To solve the non-linear BVPs, Chawla and Subramanian [43] presented the fourth-order numerical method using the cubic spline and Chawla [39] developed the eight-order numerical scheme using the finite difference method. Tirmizi and Twizell [135] constructed the eight-order convergent numerical method using the finite difference method to get the numerical solutions for non-linear BVPs. Bhatta and Sastri [19] proposed the sixth-order convergence numerical schemes with the help of quintic and sextic splines to solve the non-linear BVPs. Also one can see in References [20,42,44,46,76], different numerical methods were constructed to get the numerical approximations for non-linear BVPs.

In Section 7.1, we have proposed a numerical method using the fractal quintic spline to solve the self-adjoint singularly perturbed BVPs. Continuity conditions of the fractal quintic spline are used to get the numerical scheme and the developed numerical scheme has fourth-order convergence. Also, computational efficiency and the order of convergence are verified through numerical examples.

In Section 7.2, a numerical method is developed for non-linear BVPs. Using quasi-linearization technique, we reduced the non-linear BVP into a sequence of linear BVPs. Then each of these linear BVPs have been solved using the method obtained from fractal quintic spline. The developed numerical method has fourth-order convergence and the numerical examples are presented to illustrate the applicability of the proposed method.

7.1 Self-Adjoint Boundary-Value Problems

We consider the BVP of the form

$$\begin{cases} -\varepsilon u''(x) + q\, u(x) = f(x), \quad x \in (0,1), \\ u(0) = \eta_0, \quad u(1) = \eta_1, \end{cases} \tag{7.1}$$

where $0 < \varepsilon \leq 1$, $q > 0$, $f(x)$ is sufficiently smooth function in $[0,1]$.

Let us consider a uniform mesh on the interval $I = [0,1]$ such that $0 = x_0 < x_1 < \cdots < x_N = 1$, where $x_i = ih$, $i = 0, 1, \ldots, N$ and $h = 1/N$. Let U_i be the approximate solution to $u(x_i)$. Let M_i and S_i be the approximations to $u^{(2)}(x_i)$ and $u^{(4)}(x_i)$ respectively.

Consider the IFS $\Big\{I \times \mathbb{R}; w_i(x,y) = (L_i(x), F_i(x,y)) : i = 1, 2, \ldots, N\Big\}$, where $L_i : I \to I_i = [x_{i-1}, x_i]$ such that $L_i(x) = hx + x_{i-1}$, $x \in I$, $F_i : I \times \mathbb{R} \to \mathbb{R}$ such that $F_i(x,y) = \alpha y + r_i(x)$, for $(x,y) \in I \times \mathbb{R}$, with $r_i(x) = \mathcal{A}_i(x - x_0)^5 + \mathcal{B}_i(x - x_0)^4 + \mathcal{C}_i(x - x_0)^3 + \mathcal{D}_i(x - x_0)^2 + \mathcal{E}_i(x - x_0) + \mathcal{F}_i$ and α is the scaling factor such that $|\alpha| < h^4$.

We assume the following conditions on the IFS:

$$F_i(x_0, U_0) = U_{i-1}, \; F_i(x_N, U_N) = U_i, \quad i = 1, 2, \ldots, N,$$

$$F_{i,1}(x_N, U_{N,1}) = F_{i+1,1}(x_0, U_{0,1}), \; i = 1, 2, \ldots, N-1,$$

$$F_{i,2}(x_0, M_0) = M_{i-1}, \; F_{i,2}(x_N, M_N) = M_i, \quad i = 1, 2, \ldots, N,$$

$$F_{i,3}(x_N, U_{N,3}) = F_{i+1,3}(x_0, U_{0,3}), \; i = 1, 2, \ldots, N-1,$$

$$F_{i,4}(x_0, S_0) = S_{i-1}, \; F_{i,4}(x_N, S_N) = S_i, \quad i = 1, 2, \ldots, N,$$

where $F_{i,k}(x,y) = \frac{\alpha y + r_i^{(k)}(x)}{h^k}$, $k = 1,2,3,4$ and $U_{0,1} = \frac{r_1^{(1)}(x_0)}{h-\alpha}$, $U_{N,1} = \frac{r_N^{(1)}(x_N)}{h-\alpha}$, $U_{0,3} = \frac{r_1^{(3)}(x_0)}{h^3-\alpha}$, $U_{N,3} = \frac{r_N^{(3)}(x_N)}{h^3-\alpha}$. It can be seen that the IFS $\Big\{I \times \mathbb{R}; w_i(x,y) = (L_i(x),\ F_i(x,y)) : i = 1,2,\dots,N\Big\}$ satisfies the conditions for $\mathcal{C}^4$-FIFs. [11, 30, 36].

Let $\mathscr{F} = \{\phi \in \mathcal{C}^4(I,\mathbb{R}) \mid \phi(x_0) = U_0, \phi(x_N) = U_N, \phi^{(2)}(x_0) = M_0, \phi^{(2)}(x_N) = M_N, \phi^{(4)}(x_0) = S_0, \phi^{(4)}(x_N) = S_N\}$. Then $(\mathscr{F},\ \rho)$ is a complete metric space where ρ is the metric induced by $\mathcal{C}^4$-norm on $\mathscr{F}$. Define the Read-Bajraktarević operator T on $(\mathscr{F},\ \rho)$ as

$$\begin{aligned}T\phi(L_i(x)) = \alpha\phi(x) + \mathcal{A}_i(x-x_0)^5 + \mathcal{B}_i(x-x_0)^4 + \mathcal{C}_i(x-x_0)^3 + \mathcal{D}_i(x-x_0)^2 \\ + \mathcal{E}_i(x-x_0) + \mathcal{F}_i, \quad x \in [x_0, x_N],\ i = 1,2,\dots,N.\end{aligned}$$

The constants $\mathcal{A}_i$, $\mathcal{B}_i$, $\mathcal{C}_i$, $\mathcal{D}_i$, $\mathcal{E}_i$ and $\mathcal{F}_i$ are to be evaluated. It can be seen that T is a contraction operator, hence it has a unique fixed-point $\varPhi$ (say). The fixed-point $\varPhi$ satisfies the functional equation

$$\begin{aligned}\varPhi(L_i(x)) = \alpha\varPhi(x) + \mathcal{A}_i(x-x_0)^5 + \mathcal{B}_i(x-x_0)^4 + \mathcal{C}_i(x-x_0)^3 + \mathcal{D}_i(x-x_0)^2 \\ + \mathcal{E}_i(x-x_0) + \mathcal{F}_i, \quad x \in [x_0, x_N],\ i = 1,2,\dots,N. \end{aligned} \tag{7.2}$$

From [36], it can be observed that the conditions $F_i(x_0, U_0) = U_{i-1}, F_i(x_N, U_N) = U_i$, $F_{i,2}(x_0, M_0) = M_{i-1}, F_{i,2}(x_N, M_N) = M_i$, $F_{i,4}(x_0, S_0) = S_{i-1}, F_{i,4}(x_N, S_N) = S_i$ are equivalent to

$$\begin{cases} \varPhi(x_{i-1}) = U_{i-1},\ \varPhi(x_i) = U_i,\ \varPhi^{(2)}(x_{i-1}) = M_{i-1},\ \varPhi^{(2)}(x_i) = M_i, \\ \varPhi^{(4)}(x_{i-1}) = S_{i-1},\ \varPhi^{(4)}(x_i) = S_i, \end{cases} \tag{7.3}$$

the condition $F_{i,1}(x_N, U_{N,1}) = F_{i+1,1}(x_0, U_{0,1})$ can be reformulated as $\varPhi^{(1)}(L_i(x_N)) = \varPhi^{(1)}(L_{i+1}(x_0))$ and the condition $F_{i,3}(x_N, U_{N,3}) = F_{i+1,3}(x_0, U_{0,3})$ can be reformulated as $\varPhi^{(3)}(L_i(x_N)) = \varPhi^{(3)}(L_{i+1}(x_0))$.

For a fractal quintic spline $\varPhi$, the constants $\mathcal{A}_i$, $\mathcal{B}_i$, $\mathcal{C}_i$, $\mathcal{D}_i$, $\mathcal{E}_i$, $\mathcal{F}_i$ are evaluated using (7.3) and it is given by

$$\begin{aligned}\mathcal{A}_i &= \frac{h^4}{120}\Big[\Big(S_i - \frac{\alpha}{h^4}S_N\Big) - \Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big)\Big], \\ \mathcal{B}_i &= \frac{h^4}{24}\Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big),\end{aligned}$$

$$\mathcal{C}_i = -\frac{h^2}{6}\Big(M_{i-1} - \frac{\alpha}{h^2}M_0\Big) + \frac{h^2}{6}\Big(M_i - \frac{\alpha}{h^2}M_N\Big) - \frac{h^4}{12}\Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big)$$
$$- \frac{h^4}{36}\Big[\Big(S_i - \frac{\alpha}{h^4}S_N\Big) - \Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big)\Big],$$
$$\mathcal{D}_i = \frac{h^2}{2}\Big(M_{i-1} - \frac{\alpha}{h^2}M_0\Big),$$
$$\mathcal{E}_i = \Big(U_i - \alpha U_N\Big) - \Big(U_{i-1} - \alpha U_0\Big) + \frac{8h^4}{360}\Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big) + \frac{7h^4}{360}\Big(S_i - \frac{\alpha}{h^4}S_N\Big)$$
$$- \frac{2h^2}{6}\Big(M_{i-1} - \frac{\alpha}{h^2}M_0\Big) - \frac{h^2}{6}\Big(M_i - \frac{\alpha}{h^2}M_N\Big),$$
$$\mathcal{F}_i = U_{i-1} - \alpha U_0.$$

For continuity of $\Phi^{(1)}$ at x_i, $i = 1, 2, \ldots, N-1$, we need $\Phi^{(1)}(x_i^-) = \Phi^{(1)}(x_i^+)$, i.e., $\Phi^{(1)}(L_i(x_N)) = \Phi^{(1)}(L_{i+1}(x_0))$ and it leads to the following:

$$\alpha\Phi^{(1)}(x_N) + 5\mathcal{A}_i + 4\mathcal{B}_i + 3\mathcal{C}_i + 2\mathcal{D}_i + \mathcal{E}_i = \alpha\Phi^{(1)}(x_0) + \mathcal{E}_{i+1}.$$

Substituting the values of $\mathcal{A}_i$, $\mathcal{B}_i$, $\mathcal{C}_i$, $\mathcal{D}_i$, $\mathcal{E}_i$ and $\mathcal{E}_{i+1}$, we get

$$-\frac{15\alpha}{h^4}S_0 + 7S_{i-1} + 16S_i + 7S_{i+1} - \frac{15\alpha}{h^4}S_N = -\frac{360\alpha}{h^4}\Phi^{(1)}(x_0) + \frac{360\alpha}{h^4}\Phi^{(1)}(x_N)$$
$$-\frac{180\alpha}{h^4}M_0 + \frac{60}{h^2}(M_{i+1} + 4M_i + M_{i-1}) - \frac{180\alpha}{h^4}M_N - \frac{360}{h^4}(U_{i+1} - 2U_i + U_{i-1}). \quad (7.4)$$

For continuity of $\Phi^{(3)}$ at x_i, $i = 1, 2, \ldots, N-1,$, we need $\Phi^{(3)}(x_i^-) = \Phi^{(3)}(x_i^+)$, i.e., $\Phi^{(3)}(L_i(x_N)) = \Phi^{(3)}(L_{i+1}(x_0))$ and which leads to the following:

$$-\frac{3\alpha}{h^4}S_0 + S_{i-1} + 4S_i + S_{i+1} - \frac{3\alpha}{h^4}S_N$$
$$= \frac{6\alpha}{h^4}\Phi^{(3)}(x_0) - \frac{6\alpha}{h^4}\Phi^{(3)}(x_N) + \frac{6}{h^2}(M_{i-1} - 2M_i + M_{i+1}). \quad (7.5)$$

Multiplying (7.5) by 7 and subtract the resulting equation from (7.4), we get

$$S_i = \frac{\alpha}{2h^4}(S_0 + S_N) + \frac{30\alpha}{h^4}(\Phi^{(1)}(x_0) - \Phi^{(1)}(x_N)) + \frac{7\alpha}{2h^4}(\Phi^{(3)}(x_0) - \Phi^{(3)}(x_N))$$
$$+ \frac{15\alpha}{h^4}(M_0 + M_N) - \frac{3}{2h^2}(M_{i-1} + 18M_i + M_{i+1}) + \frac{30}{h^4}(U_{i+1} - 2U_i + U_{i-1}).$$

Substitute S_i in (7.5) to obtain

$$U_{i-2} + 2U_{i-1} - 6U_i + 2U_{i+1} + U_{i+2} = -\frac{\alpha}{2}(\Phi^{(3)}(x_0) - \Phi^{(3)}(x_N)) - 3\alpha(M_0 + M_N)$$
$$+ \frac{h^2}{20}[M_{i-2} + 26M_{i-1} + 66M_i + 26M_{i+1} + M_{i+2}] - 6\alpha(\Phi^{(1)}(x_0) - \Phi^{(1)}(x_N)). \quad (7.6)$$

Since (7.6) gives the relation between the U_i and M_i, (7.6) can be used to solve the differential equation (7.1). At the grid point x_i, the differential equation (7.1) is discretized as $-\varepsilon M_i + qU_i = f_i$, where $f_i = f(x_i)$. Substituting

$$\Phi^{(1)}(x_0) = \frac{-3U_0 + 4U_1 - U_2}{2h},\ \Phi^{(1)}(x_N) = \frac{3U_N - 4U_{N-1} + U_{N-2}}{2h},\ M_i = \frac{pU_i - f_i}{\varepsilon}$$
$$\Phi^{(3)}(x_0) = \frac{-5U_0 + 18U_1 - 24U_2 + 14U_3 - 3U_4}{2h^3},$$
$$\Phi^{(3)}(x_N) = \frac{5U_N - 18U_{N-1} + 24U_{N-2} - 14U_{N-3} + 3U_{N-4}}{2h^3}$$

in (7.6), we get the following system:

$$\begin{aligned}
&\Big[-\frac{240\alpha\varepsilon}{h} - \frac{90\alpha\varepsilon}{h^3}\Big]U_1 + \Big[\frac{60\alpha\varepsilon}{h} + \frac{120\alpha\varepsilon}{h^3}\Big]U_2 - \frac{70\alpha\varepsilon}{h^3}U_3 + \frac{15\alpha\varepsilon}{h^3}U_4\\
&+ (qh^2 - 20\varepsilon)U_{i-2} + 2(13qh^2 - 20\varepsilon)U_{i-1} + 6(11qh^2 + 20\varepsilon)U_i\\
&+ 2(13qh^2 - 20\varepsilon)U_{i+1} + (qh^2 - 20\varepsilon)U_{i+2} + \frac{15\alpha\varepsilon}{h^3}U_{N-4} - \frac{70\alpha\varepsilon}{h^3}U_{N-3}\\
&+ \Big[\frac{60\alpha\varepsilon}{h} + \frac{120\alpha\varepsilon}{h^3}\Big]U_{N-2} + \Big[-\frac{240\alpha\varepsilon}{h} - \frac{90\alpha\varepsilon}{h^3}\Big]U_{N-1} = -60\alpha f_0\\
&- 60\alpha f_N + h^2\Big[f_{i-2} + 26f_{i-1} + 66f_i + 26f_{i+1} + f_{i+2}\Big] - \Big[\frac{180\alpha\varepsilon}{h}\\
&+ \frac{25\alpha\varepsilon}{h^3} - 60\alpha q\Big](\eta_0 + \eta_1),\ i = 2, 3, \ldots, N-2, \qquad (7.7)
\end{aligned}$$

where $U_0 = \eta_0$ and $U_N = \eta_1$ are the discretizations for the boundary conditions.

Thus, we have $(N-3)$ equations with $(N-1)$ unknowns $U_1, U_2, \ldots, U_{N-1}$. We need two more equations to get the numerical approximations of the BVP given in (7.1). In order to get another two equations, the following procedure is used:

Our aim is to develop the difference equation whose truncation error is $\mathcal{O}(h^6)$. Put $i = 1$ in (7.7), we get

$$\begin{aligned}
&\Big[-\frac{240\alpha\varepsilon}{h} - \frac{90\alpha\varepsilon}{h^3}\Big]U_1 + \Big[\frac{60\alpha\varepsilon}{h} + \frac{120\alpha\varepsilon}{h^3}\Big]U_2 - \frac{70\alpha\varepsilon}{h^3}U_3 + \frac{15\alpha\varepsilon}{h^3}U_4\\
&+ (qh^2 - 20\varepsilon)U_{-1} + 2(13qh^2 - 20\varepsilon)U_0 + 6(11qh^2 + 20\varepsilon)U_1 + 2(13qh^2 - 20\varepsilon)U_2\\
&+ (qh^2 - 20\varepsilon)U_3 + \frac{15\alpha\varepsilon}{h^3}U_{N-4} - \frac{70\alpha\varepsilon}{h^3}U_{N-3} + \Big[\frac{60\alpha\varepsilon}{h} + \frac{120\alpha\varepsilon}{h^3}\Big]U_{N-2}\\
&+ \Big[-\frac{240\alpha\varepsilon}{h} - \frac{90\alpha\varepsilon}{h^3}\Big]U_{N-1} = -60\alpha f_0 - 60\alpha f_N + h^2\Big[f_{-1} + 26f_0 + 66f_1\\
&+ 26f_2 + f_3\Big] - \Big[\frac{180\alpha\varepsilon}{h} + \frac{25\alpha\varepsilon}{h^3} - 60\alpha q\Big](\eta_0 + \eta_1). \qquad (7.8)
\end{aligned}$$

We replace the fictitious values U_{-1} and f_{-1} by $4U_0-6U_1+4U_2-U_3$ and $4f_0-6f_1+4f_2-f_3$ respectively and adding $\varepsilon h^2\Big[\lambda_0 M_0+\lambda_1 M_1+\lambda_2 M_2+\lambda_3 M_3\Big]$ in (7.8), we get

$$\begin{aligned}
&\Big[-\frac{240\alpha\varepsilon}{h}-\frac{90\alpha\varepsilon}{h^3}\Big]U_1+\Big[\frac{60\alpha\varepsilon}{h}+\frac{120\alpha\varepsilon}{h^3}\Big]U_2-\frac{70\alpha\varepsilon}{h^3}U_3+\frac{15\alpha\varepsilon}{h^3}U_4\\
&+\Big[30qh^2-120\varepsilon\Big]U_0+\Big[60qh^2+240\varepsilon\Big]U_1+\Big[30qh^2-120\varepsilon\Big]U_2\\
&+\frac{15\alpha\varepsilon}{h^3}U_{N-4}-\frac{70\alpha\varepsilon}{h^3}U_{N-3}+\Big[\frac{60\alpha\varepsilon}{h}+\frac{120\alpha\varepsilon}{h^3}\Big]U_{N-2}+\Big[-\frac{240\alpha\varepsilon}{h}\\
&-\frac{90\alpha\varepsilon}{h^3}\Big]U_{N-1}+\varepsilon h^2\Big[\lambda_0 M_0+\lambda_1 M_1+\lambda_2 M_2+\lambda_3 M_3\Big]=-60\alpha f_0\\
&-60\alpha f_N+h^2\Big[30f_0+60f_1+30f_2\Big]-\Big[\frac{180\alpha\varepsilon}{h}+\frac{25\alpha\varepsilon}{h^3}-60\alpha q\Big](\eta_0+\eta_1), \qquad (7.9)
\end{aligned}$$

where $\lambda_0,\lambda_1,\lambda_2,\lambda_3$ are constants that will be evaluated. After replacing $M_i=\frac{qU_i-f_i}{\varepsilon}$ in (7.9), we get

$$\begin{aligned}
&\Big[-\frac{240\alpha\varepsilon}{h}-\frac{90\alpha\varepsilon}{h^3}\Big]U_1+\Big[\frac{60\alpha\varepsilon}{h}+\frac{120\alpha\varepsilon}{h^3}\Big]U_2-\frac{70\alpha\varepsilon}{h^3}U_3+\frac{15\alpha\varepsilon}{h^3}U_4\\
&+\Big[(30+\lambda_0)qh^2-120\varepsilon\Big]U_0+\Big[(60+\lambda_1)qh^2+240\varepsilon\Big]U_1+\Big[(30+\lambda_2)qh^2-120\varepsilon\Big]U_2\\
&+\Big[\lambda_3h^2q\Big]U_3+\frac{15\alpha\varepsilon}{h^3}U_{N-4}-\frac{70\alpha\varepsilon}{h^3}U_{N-3}+\Big[\frac{60\alpha\varepsilon}{h}+\frac{120\alpha\varepsilon}{h^3}\Big]U_{N-2}+\Big[-\frac{240\alpha\varepsilon}{h}\\
&-\frac{90\alpha\varepsilon}{h^3}\Big]U_{N-1}=-60\alpha f_0-60\alpha f_N+h^2\Big[(30+\lambda_0)f_0+(60+\lambda_1)f_1+(30+\lambda_2)f_2\Big]\\
&+\lambda_3h^2f_3-\Big[\frac{180\alpha\varepsilon}{h}+\frac{25\alpha\varepsilon}{h^3}-60\alpha q\Big](\eta_0+\eta_1). \qquad (7.10)
\end{aligned}$$

Now, the values $\lambda_0,\lambda_1,\lambda_2,\lambda_3$ are going to be evaluated so that the truncation error associated with the equation given in (7.10) is $\mathcal{O}(h^6)$. After replacing the approximate value U_i by $u(x_i)$ in (7.10), we define the truncation error $T_1(h)$ associated with the equation given in (7.10) as

$$\begin{aligned}
T_1(h)=&-120\alpha\varepsilon\Big[\frac{-3u(x_0)+4(x_1)-u(x_2)}{2h}\Big]-60\alpha\Big[qu(x_0)-f_0\Big]\\
&-10\alpha\varepsilon\Big[\frac{-5u(x_0)+18u(x_1)-24u(x_2)+14u(x_3)-3u(x_4)}{2h^3}\Big]\\
&+120\alpha\varepsilon\Big[\frac{3u(x_N)-4(x_{N-1})+u(x_{N-2})}{2h}\Big]-60\alpha\Big[qu(x_N)-f_N\Big]\\
&+10\alpha\varepsilon\Big[\frac{5u(x_N)-18u(x_{N-1})+24u(x_{N-2})-14u(x_{N-3})+3u(x_{N-4})}{2h^3}\Big]\\
&+\Big[(30+\lambda_0)qh^2-120\varepsilon\Big]u(x_0)+\Big[(60+\lambda_1)qh^2+240\varepsilon\Big]u(x_1)\\
&+\Big[(30+\lambda_2)qh^2-120\varepsilon\Big]u(x_2)+\Big[\lambda_3qh^2\Big]u(x_3)-h^2\Big[(30+\lambda_0)f_0
\end{aligned}$$

$$+ (60+\lambda_1)f_1 + (30+\lambda_2)f_2\Big] - \lambda_3 h^2 f_3. \tag{7.11}$$

Substitute $f_i = -\varepsilon u''(x_i) + q\, u(x_i)$ in (7.11), we get

$$\begin{aligned} T_1(h) =& -120\alpha\varepsilon\Big[u'(x_0) + \mathcal{O}(h^2)\Big] - 60\alpha\Big[\varepsilon u''(x_0)\Big] - 10\alpha\varepsilon\Big[u'''(x_0) + \mathcal{O}(h^2)\Big] \\ &+ 120\alpha\varepsilon\Big[u'(x_N) + \mathcal{O}(h^2)\Big] - 60\alpha\Big[\varepsilon u''(x_N)\Big] + 10\alpha\varepsilon\Big[u'''(x_0) + \mathcal{O}(h^2)\Big] \\ &- 120\varepsilon u(x_0) + 240\varepsilon u(x_1) - 120\varepsilon u(x_2) + \varepsilon h^2\lambda_3 u''(x_3) \\ &+ \varepsilon h^2\Big[(\lambda_0+30)u''(x_0) + (\lambda_1+60)u''(x_1) + (\lambda_2+30)u''(x_2)\Big]. \end{aligned}$$

Now $T_1(h)$ can be written as $T_1(h) = T_\alpha(h) + T_*(h)$, where

$$\begin{aligned} T_\alpha(h) =& -120\alpha\varepsilon\Big[u'(x_0) + \mathcal{O}(h^2)\Big] - 60\alpha\Big[\varepsilon u''(x_0)\Big] - 10\alpha\varepsilon\Big[u'''(x_0) + \mathcal{O}(h^2)\Big] \\ &+ 120\alpha\varepsilon\Big[u'(x_N) + \mathcal{O}(h^2)\Big] - 60\alpha\Big[\varepsilon u''(x_N)\Big] + 10\alpha\varepsilon\Big[u'''(x_0) + \mathcal{O}(h^2)\Big] \end{aligned}$$

and

$$\begin{aligned} T_*(h) =& -120\varepsilon u(x_0) + 240\varepsilon u(x_1) - 120\varepsilon u(x_2) + \varepsilon h^2\lambda_3 u''(x_3) \\ &+ \varepsilon h^2\Big[(\lambda_0+30)u''(x_0) + (\lambda_1+60)u''(x_1) + (\lambda_2+30)u''(x_2)\Big]. \end{aligned}$$

In order to prove $T_1(h) = \mathcal{O}(h^6)$, we will show that $T_\alpha(h) = \mathcal{O}(h^6)$ and $T_*(h) = \mathcal{O}(h^6)$. Since each term in $T_\alpha(h)$ contains the parameter α, we can choose $|\alpha| < h^6$ and hence we get $T_\alpha(h) = \mathcal{O}(h^6)$.

Now consider $T_*(h) = -120\varepsilon u(x_0) + 240\varepsilon u(x_1) - 120\varepsilon u(x_2) + \varepsilon h^2\lambda_3 u''(x_3) + \varepsilon h^2\Big[(\lambda_0 + 30)u''(x_0) + (\lambda_1+60)u''(x_1) + (\lambda_2+30)u''(x_2)\Big]$. Using the Taylor series expansion for $u(x_0)$, $u(x_2)$, $u''(x_0)$, $u''(x_2)$, $u''(x_3)$ about the point x_1 and comparing the coefficients, we get

$$\begin{aligned} T_*(h) =& \Big[\varepsilon h^2(\lambda_0+\lambda_1+\lambda_2+\lambda_3)\Big]u''(x_1) + \Big[\varepsilon h^3(-\lambda_0+\lambda_2+2\lambda_3)\Big]u^{(3)}(x_1) + \Big[\varepsilon h^4(20+ \\ &\frac{\lambda_0}{2} + \frac{\lambda_2}{2} + 2\lambda_3)\Big]u^{(4)}(x_1) + \Big[\varepsilon h^5(-\frac{\lambda_0}{6} + \frac{\lambda_2}{6} + \frac{4\lambda_3}{3})\Big]u^{(5)}(x_1) + \mathcal{O}(h^6). \end{aligned}$$

For $\lambda_0 = -20$, $\lambda_1 = 40$, $\lambda_2 = -20$ and $\lambda_3 = 0$, we get $T_*(h) = \mathcal{O}(h^6)$. Hence when $|\alpha| < h^6$, $\lambda_0 = -20$, $\lambda_1 = 40$, $\lambda_2 = -20$ and $\lambda_3 = 0$, we get $T_1(h) = \mathcal{O}(h^6)$.

After substituting $\lambda_0 = -20$, $\lambda_1 = 40$, $\lambda_2 = -20$ and $\lambda_3 = 0$ in (7.10), we get an equation as

$$\begin{aligned}
&\Big[-\frac{24\alpha\varepsilon}{h} - \frac{9\alpha\varepsilon}{h^3} + 10qh^2 + 24\varepsilon\Big]U_1 + \Big[\frac{6\alpha\varepsilon}{h} + \frac{12\alpha\varepsilon}{h^3} + qh^2 - 12\varepsilon\Big]U_2 \\
&- \frac{7\alpha\varepsilon}{h^3}U_3 + \frac{3\alpha\varepsilon}{2h^3}U_4 + \frac{3\alpha\varepsilon}{2h^3}U_{N-4} - \frac{7\alpha\varepsilon}{h^3}U_{N-3} + \Big[\frac{6\alpha\varepsilon}{h} + \frac{12\alpha\varepsilon}{h^3}\Big]U_{N-2} \\
&+ \Big[-\frac{24\alpha\varepsilon}{h} - \frac{9\alpha\varepsilon}{h^3}\Big]U_{N-1} = -6\alpha f_0 - 6\alpha f_N + h^2\Big[f_0 + 10f_1 + f_2\Big] \\
&- \Big[qh^2 - 12\varepsilon\Big]U_0 - \Big[\frac{18\alpha\varepsilon}{h} + \frac{5\alpha\varepsilon}{2h^3} - 6\alpha q\Big](\eta_0 + \eta_1). \qquad (7.12)
\end{aligned}$$

Similarly, put $i = N - 1$ in (7.7), using similar procedure, we can get an equation as

$$\begin{aligned}
&\Big[-\frac{24\alpha\varepsilon}{h} - \frac{9\alpha\varepsilon}{h^3}\Big]U_1 + \Big[\frac{6\alpha\varepsilon}{h} + \frac{12\alpha\varepsilon}{h^3}\Big]U_2 - \frac{7\alpha\varepsilon}{h^3}U_3 + \frac{3\alpha\varepsilon}{2h^3}U_4 \\
&+ \frac{3\alpha\varepsilon}{2h^3}U_{N-4} - \frac{7\alpha\varepsilon}{h^3}U_{N-3} + \Big[\frac{6\alpha\varepsilon}{h} + \frac{12\alpha\varepsilon}{h^3} + qh^2 - 12\varepsilon\Big]U_{N-2} \\
&+ \Big[-\frac{24\alpha\varepsilon}{h} - \frac{9\alpha\varepsilon}{h^3} + 10qh^2 + 24\varepsilon\Big]U_{N-1} = -6\alpha f_0 - 6\alpha f_N + h^2\Big[f_N \\
&+ 10f_{N-1} + f_{N-2}\Big] - \Big[qh^2 - 12\varepsilon\Big]U_N - \Big[\frac{18\alpha\varepsilon}{h} + \frac{5\alpha\varepsilon}{2h^3} - 6\alpha q\Big](\eta_0 + \eta_1). \qquad (7.13)
\end{aligned}$$

The system given in (7.12), (7.7) and (7.13) provides the approximations U_1, U_2, $\ldots$, U_{N-1}.

Remark 7.1.1. *If $\alpha = 0$, then the system given in (7.12), (7.7) and (7.13) reduces into the system that corresponding to the quintic spline.*

7.1.1 Convergence analysis

In this section, the convergence analysis of the proposed method is carried out. Let $|\alpha| < h^6$. We get the truncation error $T_1(h)$ associated with (7.12) as

$$\begin{aligned}
T_1(h) =& \Big[\frac{18\alpha\varepsilon}{h} + \frac{5\alpha\varepsilon}{2h^3} - 6\alpha q + qh^2 - 12\varepsilon\Big]u(x_0) + \Big[-\frac{24\alpha\varepsilon}{h} - \frac{9\alpha\varepsilon}{h^3} \\
&+ 10qh^2 + 24\varepsilon\Big]u(x_1) + \Big[\frac{6\alpha\varepsilon}{h} + \frac{12\alpha\varepsilon}{h^3} + qh^2 - 12\varepsilon\Big]u(x_2) - \frac{7\alpha\varepsilon}{h^3}u(x_3) \\
&+ \frac{3\alpha\varepsilon}{2h^3}u(x_4) + \frac{3\alpha\varepsilon}{2h^3}u(x_{N-4}) - \frac{7\alpha\varepsilon}{h^3}u(x_{N-3}) + \Big[\frac{6\alpha\varepsilon}{h} + \frac{12\alpha\varepsilon}{h^3}\Big]u(x_{N-2}) \\
&+ \Big[-\frac{24\alpha\varepsilon}{h} - \frac{9\alpha\varepsilon}{h^3}\Big]u(x_{N-1}) + \Big[\frac{18\alpha\varepsilon}{h} + \frac{5\alpha\varepsilon}{2h^3} - 6\alpha q\Big]u(x_N) + 6\alpha f_0 \\
&+ 6\alpha f_N - h^2\Big[f_0 + 10f_1 + f_2\Big]. \qquad (7.14)
\end{aligned}$$

Substituting $f_i = -\varepsilon u''(x_i) + q\, u(x_i)$ in (7.14) and rearranging the terms, we get

$$\begin{aligned}
T_1(h) =& -12\alpha\varepsilon\Big[\frac{-3u(x_0)+4(x_1)-u(x_2)}{2h}\Big] - 6\alpha\varepsilon\Big[u''(x_0)\Big] \\
& - \alpha\varepsilon\Big[\frac{-5u(x_0)+18u(x_1)-24u(x_2)+14u(x_3)-3u(x_4)}{2h^3}\Big] \\
& + 12\alpha\varepsilon\Big[\frac{3u(x_N)-4(x_{N-1})+u(x_{N-2})}{2h}\Big] - 6\alpha\varepsilon\Big[u''(x_N)\Big] \\
& + \alpha\varepsilon\Big[\frac{5u(x_N)-18u(x_{N-1})+24u(x_{N-2})-14u(x_{N-3})+3u(x_{N-4})}{2h^3}\Big] \\
& - 12\varepsilon u(x_0) + 24\varepsilon u(x_1) - 12\varepsilon u(x_2) + \varepsilon h^2\Big[u''(x_0)+10u''(x_1)+u''(x_2)\Big] \\
=& -12\alpha\varepsilon\Big[u'(x_0)+\mathcal{O}(h^2)\Big] - 6\alpha\varepsilon\Big[u''(x_0)\Big] - \alpha\varepsilon\Big[u'''(x_0)+\mathcal{O}(h^2)\Big] \\
& + 12\alpha\varepsilon\Big[u'(x_N)+\mathcal{O}(h^2)\Big] - 6\alpha\varepsilon\Big[u''(x_N)\Big] + \alpha\varepsilon\Big[u'''(x_N)+\mathcal{O}(h^2)\Big] \\
& - 12\varepsilon u(x_0) + 24\varepsilon u(x_1) - 12\varepsilon u(x_2) + \varepsilon h^2\Big[u''(x_0)+10u''(x_1)+u''(x_2)\Big]. \quad (7.15)
\end{aligned}$$

Using the Taylor expansions for $u(x_0),\ u(x_2), u''(x_0),\ u''(x_2)$ about the point x_1 in (7.15), we get

$$\begin{aligned}
T_1(h) =& -12\alpha\varepsilon\Big[u'(x_0)+\mathcal{O}(h^2)\Big] - 6\alpha\varepsilon\Big[u''(x_0)\Big] - \alpha\varepsilon\Big[u'''(x_0)+\mathcal{O}(h^2)\Big] \\
& + 12\alpha\varepsilon\Big[u'(x_N)+\mathcal{O}(h^2)\Big] - 6\alpha\varepsilon\Big[u''(x_N)\Big] + \alpha\varepsilon\Big[u'''(x_N)+\mathcal{O}(h^2)\Big] \\
& + \frac{\varepsilon h^6}{20}u^{(6)}(x_1) + \mathcal{O}(h^7),
\end{aligned}$$

since $|\alpha| < h^6$, it can be seen that, when $h \to 0$, we get $|T_1(h)| \leq K_1 h^6$ where K_1 is a constant. Thus $T_1(h) = \mathcal{O}(h^6)$ as $h \to 0$. Using the similar procedure followed to find $T_1(h)$, we obtain the truncation error $T_i(h)$, $i = 2, 3, \ldots, N-2$ associated with the equations given in (7.7) as

$$\begin{aligned}
T_i(h) =& -120\alpha\varepsilon\Big[u'(x_0)+\mathcal{O}(h^2)\Big] - 60\alpha\varepsilon\Big[u''(x_0)\Big] - 10\alpha\varepsilon\Big[u'''(x_0)+\mathcal{O}(h^2)\Big] \\
& + 120\alpha\varepsilon\Big[u'(x_N)+\mathcal{O}(h^2)\Big] - 60\alpha\varepsilon\Big[u''(x_N)\Big] + 10\alpha\varepsilon\Big[u'''(x_N)+\mathcal{O}(h^2)\Big] \\
& - \frac{\varepsilon h^6}{6}u^{(6)}(x_i) + \mathcal{O}(h^7), \quad i = 2, 3, \ldots, N-2,
\end{aligned}$$

and the truncation error $T_{N-1}(h)$ associated with the equation (7.13) as

$$T_{N-1}(h) = -12\alpha\varepsilon\Big[u'(x_0)+\mathcal{O}(h^2)\Big] - 6\alpha\varepsilon\Big[u''(x_0)\Big] - \alpha\varepsilon\Big[u'''(x_0)+\mathcal{O}(h^2)\Big]$$

$$+ 12\alpha\varepsilon\Big[u'(x_N) + \mathcal{O}(h^2)\Big] - 6\alpha\varepsilon\Big[u''(x_N)\Big] + \alpha\varepsilon\Big[u'''(x_N) + \mathcal{O}(h^2)\Big]$$
$$+ \frac{\varepsilon h^6}{20} u^{(6)}(x_{N-1}) + \mathcal{O}(h^7).$$

As $h \to 0$, we obtain $|T_i(h)| \leq K_i h^6$, where K_i is a constant and hence $T_i(h) = \mathcal{O}(h^6)$ as $h \to 0$ for $i = 1, 2, \ldots, N-1$. The system given in (7.12), (7.7) and (7.13) can be written in the following form $AU = d$, where

$$A = \begin{bmatrix}
A_{1,1} & A_{1,2} & A_{1,3} & A_{1,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{1,N-4} & A_{1,N-3} & A_{1,N-2} & A_{1,N-1} \\
A_{2,1} & A_{2,2} & A_{2,3} & A_{2,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{2,N-4} & A_{2,N-3} & A_{2,N-2} & A_{2,N-1} \\
A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} & A_{3,5} & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{3,N-4} & A_{3,N-3} & A_{3,N-2} & A_{3,N-1} \\
A_{4,1} & A_{4,2} & A_{4,3} & A_{4,4} & A_{4,5} & A_{4,6} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{4,N-4} & A_{4,N-3} & A_{4,N-2} & A_{4,N-1} \\
A_{5,1} & A_{5,2} & A_{5,3} & A_{5,4} & A_{5,5} & A_{5,6} & A_{5,7} & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{5,N-4} & A_{5,N-3} & A_{5,N-2} & A_{5,N-1} \\
A_{6,1} & A_{6,2} & A_{6,3} & A_{6,4} & A_{6,5} & A_{6,6} & A_{6,7} & A_{6,8} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{6,N-4} & A_{6,N-3} & A_{6,N-2} & A_{6,N-1} \\
A_{7,1} & A_{7,2} & A_{7,3} & A_{7,4} & A_{7,5} & A_{7,6} & A_{7,7} & A_{7,8} & A_{7,9} & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{7,N-4} & A_{7,N-3} & A_{7,N-2} & A_{7,N-1} \\
A_{8,1} & A_{8,2} & A_{8,3} & A_{8,4} & 0 & A_{8,6} & A_{8,7} & A_{8,8} & A_{8,9} & A_{8,10} & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{8,N-4} & A_{8,N-3} & A_{8,N-2} & A_{8,N-1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
A_{N-8,1} & A_{N-8,2} & A_{N-8,3} & A_{N-8,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & A_{N-8,N-10} & A_{N-8,N-9} & A_{N-8,N-8} & A_{N-8,N-7} & A_{N-8,N-6} & 0 & A_{N-8,N-4} & A_{N-8,N-3} & A_{N-8,N-2} & A_{N-8,N-1} \\
A_{N-7,1} & A_{N-7,2} & A_{N-7,3} & A_{N-7,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & A_{N-7,N-9} & A_{N-7,N-8} & A_{N-7,N-7} & A_{N-7,N-6} & A_{N-7,N-5} & A_{N-7,N-4} & A_{N-7,N-3} & A_{N-7,N-2} & A_{N-7,N-1} \\
A_{N-6,1} & A_{N-6,2} & A_{N-6,3} & A_{N-6,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & A_{N-6,N-8} & A_{N-6,N-7} & A_{N-6,N-6} & A_{N-6,N-5} & A_{N-6,N-4} & A_{N-6,N-3} & A_{N-6,N-2} & A_{N-6,N-1} \\
A_{N-5,1} & A_{N-5,2} & A_{N-5,3} & A_{N-5,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & A_{N-5,N-7} & A_{N-5,N-6} & A_{N-5,N-5} & A_{N-5,N-4} & A_{N-5,N-3} & A_{N-5,N-2} & A_{N-5,N-1} \\
A_{N-4,1} & A_{N-4,2} & A_{N-4,3} & A_{N-4,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & A_{N-4,N-6} & A_{N-4,N-5} & A_{N-4,N-4} & A_{N-4,N-3} & A_{N-4,N-2} & A_{N-4,N-1} \\
A_{N-3,1} & A_{N-3,2} & A_{N-3,3} & A_{N-3,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & A_{N-3,N-5} & A_{N-3,N-4} & A_{N-3,N-3} & A_{N-3,N-2} & A_{N-3,N-1} \\
A_{N-2,1} & A_{N-2,2} & A_{N-2,3} & A_{N-2,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{N-2,N-4} & A_{N-2,N-3} & A_{N-2,N-2} & A_{N-2,N-1} \\
A_{N-1,1} & A_{N-1,2} & A_{N-1,3} & A_{N-1,4} & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & A_{N-1,N-4} & A_{N-1,N-3} & A_{N-1,N-2} & A_{N-1,N-1}
\end{bmatrix}$$

$A_{i,j}$ is the coefficient of U_j, $U = (U_1, U_2, \ldots, U_{N-1})^T$, $d = (d_1,\ d_2,\ \ldots,\ d_{N-1})^T$, d_i, $i = 1, 2, \ldots, N-1$ are the right hand side of the system.

Let B be the matrix corresponding to $\alpha = 0$. We take the matrices A, B, $n = N-1$, $\|.\|_\infty$ norm in the Theorem 1.7.4 and hence we obtain

$$\max_i |\lambda_i' - \mu_i'| \leq 2^{\frac{2N-3}{N-1}} (N-1)^{\frac{1}{N-1}} (2Q)^{\frac{N-2}{N-1}} \|A - B\|_\infty^{\frac{1}{N-1}} \tag{7.16}$$

where $Q = \max\{\|A\|_\infty,\ \|B\|_\infty\}$, λ_i' and μ_i', $i = 1, 2, \ldots, N-1$ are the eigenvalues of A and B respectively. Let $\mathcal{R}_i = \displaystyle\sum_{j=1}^{N-1} B_{i,j}$ be the row sum of B. Note that, when h is sufficiently small, we obtain B is irreducible matrix,

- $B_{i,i} > 0$, $B_{i,j} \leq 0$, $i \neq j$; $i, j = 1, 2, \ldots, N-1$,

- $\mathcal{R}_1 = 11h^2 q + 12\varepsilon > 0$,

- $\mathcal{R}_2 = 119h^2 q + 20\varepsilon > 0$,

- $\mathcal{R}_i = 120h^2 q > 0$, $i = 3, 4, \ldots, N-3$,

- $\mathcal{R}_{N-2} = 119h^2 q + 20\varepsilon > 0$,

- $\mathcal{R}_{N-1} = 11h^2q + 12\varepsilon > 0$.

Hence, B is monotone matrix [67]. Thus B^{-1} exist and eigenvalues μ_i', $i = 1, 2, \ldots, N-1$ of B are non-zero. Thus, when h is sufficiently small, B is invertible and its row sums are positive. Once we fix h (which gives B is monotone), then the eigenvalues of B are non-zero. Now, the α can vary in the region $(-h^6, h^6)$. We can choose sufficiently small α in the region $(-h^6, h^6)$ to satisfy the following conditions:

- A is invertible because

$$\|A - B\|_\infty = 2\left| -\frac{240\alpha\varepsilon}{h} - \frac{90\alpha\varepsilon}{h^3}\right| + 2\left|\frac{60\alpha\varepsilon}{h} + \frac{120\alpha\varepsilon}{h^3}\right| + 2\left| -\frac{70\alpha\varepsilon}{h^3}\right| + 2\left|\frac{15\alpha\varepsilon}{h^3}\right|$$

  and from (7.16), it can been seen that, when α is sufficiently small, eigenvalues of A are non-zero.

- The row sum $\mathcal{S}_i$ of the matrix A,

$$\begin{aligned} \mathcal{S}_1 &= \mathcal{R}_1 - \frac{36\alpha\varepsilon}{h} - \frac{5\alpha\varepsilon}{h^3} > 0, \\ \mathcal{S}_i &= \mathcal{R}_i - \frac{360\alpha\varepsilon}{h} - \frac{50\alpha\varepsilon}{h^3} > 0,\ i = 2, 3, \ldots, N-2, \\ \mathcal{S}_{N-1} &= \mathcal{R}_i - \frac{36\alpha\varepsilon}{h} - \frac{5\alpha\varepsilon}{h^3} > 0. \end{aligned}$$

  when α is sufficiently small.

Hence by choosing sufficiently small h (which gives B is monotone), choosing α in $(-h^6, h^6)$ (which gives A is invertible and row sums of A are positive) we get the following:

The system given in (7.12), (7.7) and (7.13) with exact solutions can be written as

$$A\bar{u} = d + T(h),$$

where $\bar{u} = (u(x_1), u(x_2), \ldots, u(x_{N-1}))^T$, $T(h) = (T_1(h), T_2(h), \ldots, T_{N-1}(h))^T$. Since $AU = d$, we have

$$AE = T(h)$$

$E = (E_1, E_2, \ldots, E_{N-1})^T$, $E_i = u(x_i) - U_i$. Consequently, we have $E = A^{-1}T(h)$. By the definition of multiplication of a matrix by its inverse, we have

$$\sum_{i=1}^{N-1} A_{k,i}^{-1}\mathcal{S}_i = 1,\ k = 1, 2, \ldots, N-1,$$

where $A_{k,i}^{-1}$ is the (k,i)-th element of the matrix A^{-1}. Hence, we have

$$\sum_{i=1}^{N-1} A_{k,i}^{-1} \leq \frac{1}{\min_{1\leq i\leq N-1}\mathcal{S}_i} = \frac{1}{C_{i_0}h^2}.$$

We have

$$E_j = \sum_{i=1}^{N-1} A_{j,i}^{-1}T_i(h),\ j = 1, 2, \ldots, N-1,$$

and hence $|E_j| \leq \frac{Kh^6}{C_{i_0}h^2}$, where K is a constant. Therefore, as $h \to 0$, we have

$$\|E\|_\infty = \mathcal{O}(h^4).$$

7.1.2 Numerical examples

In this section, two numerical examples are provided to test the proposed method. Once we fix sufficiently small h (which gives B is monotone), varying α in $(-h^6, h^6)$ (which gives A is invertible and row sums of A are positive), we can compute the numerical solutions of the BVPs.

Example 7.1.1. *Let us consider the BVP [82]*

$$\begin{cases} -\varepsilon u''(x) + u(x) = 1 + 2\sqrt{\varepsilon}\Big[\exp((x-1)/\sqrt{\varepsilon}) + \exp(-x/\sqrt{\varepsilon})\Big],\ x \in (0,1), \\ u(0) = 0, \quad u(1) = 0. \end{cases}$$

The exact solution of BVP is $u(x) = 1-(1-x)\exp(-x/\sqrt{\varepsilon}) - x\exp((x-1)/\sqrt{\varepsilon})$. For each fixed $\varepsilon = 2^{1-i}$, $i = 1, 2, \ldots, 8$, we take $N = 16, 32, 64, 128, 256$. We take the scaling factor as $\alpha = 0.999h^6$. The maximum point-wise errors and the order of convergence are given in Table 7.1. Table 7.2 represents the maximum point-wise errors and the order of convergence corresponding to the quintic spline. Figure 7.1 represents the Log-log plot of the maximum errors.

Table 7.1: Maximum point-wise error and order of convergence corresponding to Example 7.1.1.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256
1	1.6714e-08	2.8739e-10	4.9246e-12	1.0159e-13	4.8850e-15
	5.8619	5.8669	5.5992	4.3782	
2^{-1}	4.9276e-08	8.8010e-10	1.6358e-11	5.3169e-13	3.4472e-14
	5.8071	5.7496	4.9433	3.9471	
2^{-2}	1.4396e-07	2.7213e-09	7.5220e-11	3.7104e-12	2.1194e-13
	5.7252	5.1771	4.3415	4.1298	
2^{-3}	3.9049e-07	9.3472e-09	4.3909e-10	2.5866e-11	1.5943e-12
	5.3846	4.4119	4.0854	4.0200	
2^{-4}	9.1270e-07	4.1168e-08	2.5969e-09	1.6459e-10	1.0336e-11
	4.4705	3.9867	3.9799	3.9931	
2^{-5}	5.0350e-06	1.8980e-07	1.3963e-08	9.1616e-10	5.8036e-11
	4.7294	3.7648	3.9299	3.9806	
2^{-6}	2.8981e-05	7.8873e-07	6.4927e-08	4.3949e-09	2.8058e-10
	5.1994	3.6026	3.8849	3.9693	
2^{-7}	1.5345e-04	3.2639e-06	2.5980e-07	1.8432e-08	1.1944e-09
	5.5550	3.6511	3.8171	3.9479	

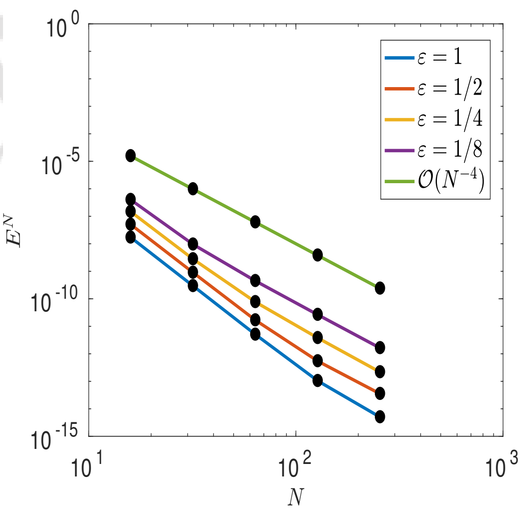


Figure 7.1: Log-log plots of Example 7.1.1.

Table 7.2: Maximum point-wise error and order of convergence corresponding to Example 7.1.1 with $\alpha = 0$.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256
1	1.7782e-08	1.1880e-09	7.5474e-11	4.7376e-12	3.3867e-13
	3.9038	3.9765	3.9938	3.8062	
2^{-1}	7.8498e-08	5.2713e-09	3.3537e-10	2.1056e-11	1.3234e-12
	3.8964	3.9743	3.9935	3.9919	
2^{-2}	3.0485e-07	2.0668e-08	1.3186e-09	8.2860e-11	5.1729e-12
	3.8826	3.9703	3.9922	4.0016	
2^{-3}	9.8871e-07	6.8191e-08	4.3724e-09	2.7510e-10	1.7216e-11
	3.8579	3.9631	3.9904	3.9981	
2^{-4}	2.5407e-06	1.8033e-07	1.1665e-08	7.3596e-10	4.6130e-11
	3.8165	3.9504	3.9864	3.9959	
2^{-5}	5.2594e-06	4.4669e-07	3.0720e-08	1.9735e-09	1.2424e-10
	3.5575	3.8620	3.9603	3.9896	
2^{-6}	2.6680e-05	1.2394e-06	9.4684e-08	6.2762e-09	3.9836e-10
	4.4281	3.7103	3.9151	3.9777	
2^{-7}	1.4974e-04	3.4897e-06	3.0982e-07	2.1596e-08	1.3918e-09
	5.4232	3.4936	3.8426	3.9558	

Example 7.1.2. *Let us consider the BVP [82]*

$$\begin{cases} -\varepsilon u''(x) + 4u(x) = 4 + 2\sqrt{\varepsilon}\Big[\exp((x-1)/\sqrt{\varepsilon}) + \exp(-x/\sqrt{\varepsilon})\Big] \\ \qquad\qquad -3(1-x)\exp(-x/\sqrt{\varepsilon}) - 3x\exp((x-1)/\sqrt{\varepsilon}), \; x \in (0,1), \\ u(0) = 0, \quad u(1) = 0. \end{cases}$$

The exact solution to the above BVP is $u(x) = 1-(1-x)\exp(-x/\sqrt{\varepsilon}) - x\exp((x-1)/\sqrt{\varepsilon})$. For each fixed $\varepsilon = 2^{1-i}$, $i = 1, 2, \ldots, 8$, we take $N = 16, 32, 64, 128, 256$. We taken the scaling factor $\alpha = 0.99h^6$. Table 7.3 represents the maximum point-wise errors and the order of convergence. The maximum point-wise errors and the order of convergence corresponding to the quintic spline are given in 7.4. Figure 7.2 represents the Log-log plot of the maximum errors.

Table 7.3: Maximum point-wise error and order of convergence corresponding to Example 7.1.2.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256
1	1.3064e-08	2.3003e-10	4.2858e-12	1.0414e-13	5.2736e-15
	5.8276	5.7461	5.3630	4.3036	
2^{-1}	3.2794e-08	5.9832e-10	1.2517e-11	4.9383e-13	3.1308e-14
	5.7764	5.5790	4.6637	3.9794	
2^{-2}	7.7419e-08	1.4931e-09	5.4112e-11	2.9698e-12	1.7547e-13
	5.6963	4.7862	4.1875	4.0811	
2^{-3}	1.6400e-07	5.2198e-09	2.9830e-10	1.8628e-11	1.1701e-12
	4.9736	4.1291	4.0012	3.9928	
2^{-4}	8.4670e-07	2.3285e-08	1.6353e-09	1.0673e-10	6.7529e-12
	5.1844	3.8317	3.9375	3.9824	
2^{-5}	4.5007e-06	1.0240e-07	7.8881e-09	5.2830e-10	3.3673e-11
	5.4579	3.6984	3.9003	3.9717	
2^{-6}	2.2285e-05	5.1080e-07	3.2737e-08	2.2710e-09	1.4612e-10
	5.4472	3.9637	3.8495	3.9581	
2^{-7}	9.9061e-05	3.1174e-06	1.1998e-07	8.8238e-09	5.7851e-10
	4.9899	4.6995	3.7652	3.9310	

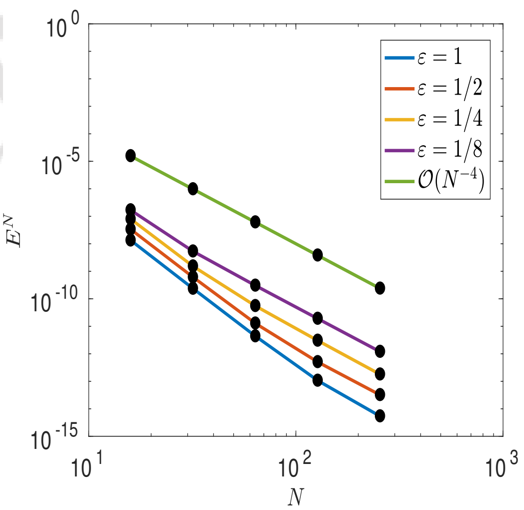


Figure 7.2: Log-log plots of Example 7.1.2.

Table 7.4: Maximum point-wise error and order of convergence corresponding to Example 7.1.2 with $\alpha = 0$.

$\downarrow \varepsilon/N \rightarrow$	16	32	64	128	256
1	1.3887e-08	9.2389e-10	5.8637e-11	3.6867e-12	2.2277e-13
	3.9099	3.9778	3.9914	4.0487	
2^{-1}	5.1738e-08	3.4449e-09	2.1874e-10	1.3731e-11	8.5465e-13
	3.9087	3.9772	3.9937	4.0060	
2^{-2}	1.6057e-07	1.0701e-08	6.7993e-10	4.2675e-11	2.6852e-12
	3.9074	3.9762	3.9939	3.9903	
2^{-3}	4.0483e-07	2.6984e-08	1.7180e-09	1.0794e-10	6.7564e-12
	3.9072	3.9733	3.9924	3.9978	
2^{-4}	8.7991e-07	6.5777e-08	4.3883e-09	2.7954e-10	1.7566e-11
	3.7417	3.9059	3.9725	3.9922	
2^{-5}	4.0079e-06	1.8208e-07	1.3098e-08	8.5444e-10	5.4025e-11
	4.4602	3.7971	3.9382	3.9833	
2^{-6}	2.1369e-05	5.3589e-07	4.2063e-08	2.8483e-09	1.8210e-10
	5.3174	3.6713	3.8844	3.9673	
2^{-7}	9.7688e-05	2.9932e-06	1.3570e-07	9.7786e-09	6.3772e-10
	5.0284	4.4632	3.7946	3.9386	

7.2 Nonlinear Boundary-Value Problem

We consider the nonlinear BVP of the form

$$\begin{cases} u_{xx}(x) + F(x, u(x)) = 0, & x \in (0,1), \\ u(0) = \eta_0, & u(1) = \eta_1. \end{cases} \tag{7.17}$$

We assume that for $(x, u(x)) \in D = \{0 \leq x \leq 1,\ -\infty < u(x) < \infty\}$, the functions F and $\partial F/\partial u$ are continuous. It is known that problem (7.17) possess unique solution provided $\sup\limits_{(x,u)\in D} \partial F/\partial u < \pi^2$ [42]. We assume that $\partial F/\partial u \leq 0$ on D and $\partial F/\partial u < 0$ on $D^\circ = \{0 < x < 1,\ -\infty < u(x) < \infty\}$. Here the notation u_{xx} is used instead of u'' mainly for the sake of convenience in the future.

Let $0 = x_0 < x_1 \cdots < x_N = 1$ be the uniform partition of the interval $I = [0, 1]$. Let $u(x)$ be the solution of the differential equation given in (7.17) and U_i be the approximate solution to $u(x_i)$.

We use the quasi-linearization technique to reduce the non-linear problem into sequence of linear problems. We choose the reasonable initial approximation for the function $u(x)$ in $F(x, u(x))$, call it $u^{(0)}(x)$ and expand $F(x, u(x))$ around the function $u^{(0)}(x)$ to obtain

$$F(x, u^{(1)}(x)) = F(x, u^{(0)}(x)) + (u^{(1)}(x) - u^{(0)}(x))\Big(\frac{\partial F}{\partial u}\Big)_{(x, u^{(0)}(x))} + \ldots,$$

or in general, we can write for $r = 0, 1, 2, \ldots$ (r is the iteration index)

$$F(x, u^{(r+1)}(x)) = F(x, u^{(r)}(x)) + (u^{(r+1)}(x) - u^{(r)}(x))\Big(\frac{\partial F}{\partial u}\Big)_{(x, u^{(r)}(x))} + \ldots.$$

From 7.17, we have

$$\begin{cases} u_{xx}^{(r+1)}(x) + F(x, u^{(r+1)}(x)) = 0, \quad x \in (0, 1), \\ u^{(r+1)}(0) = \eta_0, \quad u^{(r+1)}(1) = \eta_1. \end{cases} \tag{7.18}$$

By substituting $F(x, u^{(r+1)}(x)) = F(x, u^{(r)}(x)) + (u^{(r+1)}(x) - u^{(r)}(x))\Big(\frac{\partial F}{\partial u}\Big)_{(x, u^{(r)}(x))}$ in (7.18), we get

$$\begin{cases} u_{xx}^{(r+1)}(x) + q^{(r)}(x)u^{(r+1)}(x) = f^{(r)}(x), \quad x \in (0, 1), \\ u^{(r+1)}(0) = \eta_0, \quad u^{(r+1)}(1) = \eta_1, \end{cases} \tag{7.19}$$

where $q^{(r)}(x) = \Big(\frac{\partial F}{\partial u}\Big)_{(x, u^{(r)}(x))}$, $f^{(r)}(x) = u^{(r)}(x)\Big(\frac{\partial F}{\partial u}\Big)_{(x, u^{(r)}(x))} - F(x, u^{(r)}(x))$. Thus, we have reduced the non-linear problem into a sequence of linear problems. Now our aim is to solve these linear problems numerically.

Let $|\alpha| < h^6$. Let $U_i^{(r)}$ be the approximation of $u^{(r)}(x_i)$ and $M_i^{(r)}$ be the approximation of $u_{xx}^{(r)}(x_i)$. Now, at $x = x_i$, the differential equation (7.19) can be discretized as $M_i^{(r+1)} + q_i^{(r)}U_i^{(r+1)} = f_i^{(r)}$ where $q_i^{(r)} = \Big(\frac{\partial F}{\partial u}\Big)_{(x_i, U_i^{(r)})}$, $f_i^{(r)} = U_i^{(r)}\Big(\frac{\partial F}{\partial u}\Big)_{(x_i, U_i^{(r)})} - F(x_i, U_i^{(r)})$. Also, the boundary conditions can be discretized as $U_0^{(r+1)} = \eta_0$, $U_N^{(r+1)} = \eta_1$. Substitute $M_i^{(r+1)} = f_i^{(r)} - q_i^{(r)}U_i^{(r+1)}$, $\Phi^{(3)}(x_0) = \frac{-U_0^{(r+1)} + 3U_1^{(r+1)} - 3U_2^{(r+1)} + U_3^{(r+1)}}{h^3}$, $\Phi^{(1)}(x_0) =$

$\frac{U_1^{(r+1)}-U_0^{(r+1)}}{h}$, $\Phi^{(3)}(x_N) = \frac{U_N^{(r+1)}-3U_{N-1}^{(r+1)}+3U_{N-2}^{(r+1)}-U_{N-3}^{(r+1)}}{h^3}$, $\Phi^{(1)}(x_N) = \frac{U_N^{(r+1)}-U_{N-1}^{(r+1)}}{h}$ in (7.6), we get the following system:

$$\begin{aligned}
&\Big[-\frac{3\alpha}{2h^3}-\frac{6\alpha}{h}\Big]U_1^{(r+1)}+\frac{3\alpha}{2h^3}U_2^{(r+1)}-\frac{\alpha}{2h^3}U_3^{(r+1)}+\Big[-1-\frac{h^2}{20}q_{i-2}^{(r)}\Big]U_{i-2}^{(r+1)}\\
&+\Big[-2-\frac{13h^2}{10}q_{i-1}^{(r)}\Big]U_{i-1}^{(r+1)}+\Big[6-\frac{33h^2}{10}q_i^{(r)}\Big]U_i^{(r+1)}+\Big[-2-\frac{13h^2}{10}q_{i+1}^{(r)}\Big]U_{i+1}^{(r+1)}\\
&+\Big[-1-\frac{h^2}{20}q_{i+2}^{(r)}\Big]U_{i+2}^{(r+1)}-\frac{\alpha}{2h^3}U_{N-3}^{(r+1)}+\frac{3\alpha}{2h^3}U_{N-2}^{(r+1)}+\Big[-\frac{3\alpha}{2h^3}-\frac{6\alpha}{h}\Big]U_{N-1}^{(r+1)}\\
&=3\alpha f_0^{(r)}-\frac{h^2}{20}f_{i-2}^{(r)}-\frac{13h^2}{10}f_{i-1}^{(r)}-\frac{33h^2}{10}f_i^{(r)}-\frac{13h^2}{10}f_{i+1}^{(r)}-\frac{h^2}{20}f_{i+2}^{(r)}+3\alpha f_N^{(r)}\\
&-\Big[\frac{\alpha}{2h^3}+3\alpha q_0^{(r)}+\frac{6\alpha}{h}\Big]\eta_0-\Big[\frac{\alpha}{2h^3}+3\alpha q_N^{(r)}+\frac{6\alpha}{h}\Big]\eta_1,\ i=2,3,\ldots,N-2. \qquad (7.20)
\end{aligned}$$

Difference equations: We define

$$\begin{aligned}
&T_1^{(r)}(h)=\\
&\Big[\frac{\alpha}{2h^3}+3\alpha q^{(r)}(x_0)+\frac{6\alpha}{h}\Big]u^{(r+1)}(x_0)+\Big[-\frac{3\alpha}{2h^3}-\frac{6\alpha}{h}\Big]u^{(r+1)}(x_1)+\frac{3\alpha}{2h^3}u^{(r+1)}(x_2)\\
&-\frac{\alpha}{2h^3}u^{(r+1)}(x_3)-\frac{\alpha}{2h^3}u^{(r+1)}(x_{N-3})+\frac{3\alpha}{2h^3}u^{(r+1)}(x_{N-2})+\Big[-\frac{3\alpha}{2h^3}\\
&-\frac{6\alpha}{h}\Big]u^{(r+1)}(x_{N-1})+\Big[\frac{\alpha}{2h^3}+3\alpha q^{(r)}(x_N)+\frac{6\alpha}{h}\Big]u^{(r+1)}(x_N)-3\alpha f^{(r)}(x_0)\\
&-3\alpha f^{(r)}(x_N)+\sum_{i=0}^{3}\Big(c_iu^{(r+1)}(x_i)+d_ih^2u_{xx}^{(r+1)}(x_i)\Big). \qquad (7.21)
\end{aligned}$$

Substituting $f^{(r)}(x_i)=u_{xx}^{(r+1)}(x_i)+q^{(r)}(x_i)u^{(r+1)}(x_i)$ in (7.21), we get

$$\begin{aligned}
T_1^{(r)}(h)=&-\frac{\alpha}{2}\Big[\frac{-u^{(r+1)}(x_0)+3u^{(r+1)}(x_1)-3u^{(r+1)}(x_2)+u^{(r+1)}(x_3)}{h^3}\Big]\\
&+\frac{\alpha}{2}\Big[\frac{u^{(r+1)}(x_N)-3u^{(r+1)}(x_{N-1})+3u^{(r+1)}(x_{N-2})-u^{(r+1)}(x_{N-3})}{h^3}\Big]\\
&-6\alpha\Big[\frac{-u^{(r+1)}(x_0)+u^{(r+1)}(x_1)}{h}\Big]+6\alpha\Big[\frac{u^{(r+1)}(x_N)-u^{(r+1)}(x_{N-1})}{h}\Big]\\
&-3\alpha u_{xx}^{(r+1)}(x_0)-3\alpha u_{xx}^{(r+1)}(x_N)+\sum_{i=0}^{3}\Big(c_iu^{(r+1)}(x_i)+d_ih^2u_{xx}^{(r+1)}(x_i)\Big).
\end{aligned}$$

It can be seen that

$$\frac{-u^{(r+1)}(x_0)+3u^{(r+1)}(x_1)-3u^{(r+1)}(x_2)+u^{(r+1)}(x_3)}{h^3}=u_{xxx}^{(r+1)}(x_0)+\mathcal{O}(h),$$
$$\frac{u^{(r+1)}(x_N)-3u^{(r+1)}(x_{N-1})+3u^{(r+1)}(x_{N-2})-u^{(r+1)}(x_{N-3})}{h^3}=u_{xxx}^{(r+1)}(x_N)+\mathcal{O}(h),$$

$$\frac{-u^{(r+1)}(x_0)+u^{(r+1)}(x_1)}{h}=u_x^{(r+1)}(x_0)+\mathcal{O}(h),$$
$$\frac{u^{(r+1)}(x_N)-u^{(r+1)}(x_{N-1})}{h}=u_x^{(r+1)}(x_N)+\mathcal{O}(h).$$

Hence, we have

$$\begin{aligned}T_1^{(r)}(h)=&-\frac{\alpha}{2}\Big[u_{xxx}^{(r+1)}(x_0)+\mathcal{O}(h)\Big]+\frac{\alpha}{2}\Big[u_{xxx}^{(r+1)}(x_N)+\mathcal{O}(h)\Big]\\&-6\alpha\Big[u_x^{(r+1)}(x_0)+\mathcal{O}(h)\Big]+6\alpha\Big[u_x^{(r+1)}(x_N)+\mathcal{O}(h)\Big]\\&-3\alpha u_{xx}^{(r+1)}(x_0)-3\alpha u_{xx}^{(r+1)}(x_N)+\sum_{i=0}^{3}\Big(c_iu^{(r+1)}(x_i)+d_ih^2u_{xx}^{(r+1)}(x_i)\Big).\end{aligned}$$

Our aim is to get $T_1^{(r)}(h)=\mathcal{O}(h^6)$. We write $T_1^{(r)}(h)=T_\alpha^{(r)}(h)+T_*^{(r)}(h)$, where $T_\alpha^{(r)}(h)=-\frac{\alpha}{2}\Big[u_{xxx}^{(r+1)}(x_0)+\mathcal{O}(h)\Big]+\frac{\alpha}{2}\Big[u_{xxx}^{(r+1)}(x_N)+\mathcal{O}(h)\Big]-6\alpha\Big[u_x^{(r+1)}(x_0)+\mathcal{O}(h)\Big]+6\alpha\Big[u_x^{(r+1)}(x_N)+\mathcal{O}(h)\Big]-3\alpha u_{xx}^{(r+1)}(x_0)-3\alpha u_{xx}^{(r+1)}(x_N)$ and $T_*^{(r)}(h)=\sum_{i=0}^{3}\Big(c_iu^{(r+1)}(x_i)+d_ih^2u_{xx}^{(r+1)}(x_i)\Big)$.

Since $|\alpha|< h^6$, then we get $T_\alpha^{(r)}(h)=\mathcal{O}(h^6)$. Choosing $c_0=-4$, $c_1=7$, $c_2=-2$, $c_3=-1$, $d_0=\frac{1}{3}$, $d_1=\frac{41}{12}$, $d_2=\frac{7}{6}$, $d_3=\frac{1}{12}$, we get

$$\begin{aligned}T_*^{(r)}(h)=&-4u^{(r+1)}(x_0)+7u^{(r+1)}(x_1)-2u^{(r+1)}(x_2)-u^{(r+1)}(x_3)+\frac{h^2}{3}u_{xx}^{(r+1)}(x_0)\\&+\frac{41h^2}{12}u_{xx}^{(r+1)}(x_1)+\frac{7h^2}{6}u_{xx}^{(r+1)}(x_2)+\frac{h^2}{12}u_{xx}^{(r+1)}(x_3).\end{aligned}$$

Using the Taylor expansions for $u^{(r+1)}(x_0)$, $u^{(r+1)}(x_2)$, $u^{(r+1)}(x_3)$, $u_{xx}^{(r+1)}(x_0)$, $u_{xx}^{(r+1)}(x_2)$, $u_{xx}^{(r+1)}(x_3)$ about the point $x=x_1$, we get $T_*^{(r)}(h)=\frac{h^6}{48}u_{xxxxxx}^{(r+1)}(x_1)+\mathcal{O}(h^7)$. Thus, by the assumption $|\alpha|< h^6$ and $c_0=-4$, $c_1=7$, $c_2=-2$, $c_3=-1$, $d_0=\frac{1}{3}$, $d_1=\frac{41}{12}$, $d_2=\frac{7}{6}$, $d_3=\frac{1}{12}$, we get

$$\begin{aligned}&T_1^{(r)}(h)=\\&\Big[\frac{\alpha}{2h^3}+3\alpha q^{(r)}(x_0)+\frac{6\alpha}{h}\Big]u^{(r+1)}(x_0)+\Big[-\frac{3\alpha}{2h^3}-\frac{6\alpha}{h}\Big]u^{(r+1)}(x_1)+\frac{3\alpha}{2h^3}u^{(r+1)}(x_2)\\&-\frac{\alpha}{2h^3}u^{(r+1)}(x_3)-\frac{\alpha}{2h^3}u^{(r+1)}(x_{N-3})+\frac{3\alpha}{2h^3}u^{(r+1)}(x_{N-2})+\Big[-\frac{3\alpha}{2h^3}\\&-\frac{6\alpha}{h}\Big]u^{(r+1)}(x_{N-1})+\Big[\frac{\alpha}{2h^3}+3\alpha q^{(r)}(x_N)+\frac{6\alpha}{h}\Big]u^{(r+1)}(x_N)-3\alpha f^{(r)}(x_0)\\&-3\alpha f^{(r)}(x_N)-4u^{(r+1)}(x_0)+7u^{(r+1)}(x_1)-2u^{(r+1)}(x_2)-u^{(r+1)}(x_3)\\&+\frac{1}{3}h^2u_{xx}^{(r+1)}(x_0)+\frac{41}{12}h^2u_{xx}^{(r+1)}(x_1)+\frac{7}{6}h^2u_{xx}^{(r+1)}(x_2)+\frac{1}{12}h^2u_{xx}^{(r+1)}(x_3).\end{aligned}\tag{7.22}$$

As $h \to 0$, we have $|T_1^{(r)}(h)| \le K_1^{(r)} h^6$, where $K_1^{(r)}$ is a constant. Thus $T_1^{(r)}(h) = \mathcal{O}(h^6)$ as $h \to 0$. Hence from (7.22), replacing the original values $u^{(r)}(x_i)$ and $u_{xx}^{(r)}(x_i)$ by the approximated values $U_i^{(r)}$ and $M_i^{(r)}$ respectively, we get the difference equation (associated with truncation error is $T_1^{(r)}(h)$) as

$$\Big[7 - \frac{41h^2}{12} q_1^{(r)} - \frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\Big] U_1^{(r+1)} + \Big[-2 - \frac{7h^2}{6} q_2^{(r)} + \frac{3\alpha}{2h^3}\Big] U_2^{(r+1)}$$
$$+ \Big[-1 - \frac{h^2}{12} q_3^{(r)} - \frac{\alpha}{2h^3}\Big] U_3^{(r+1)} - \frac{\alpha}{2h^3} U_{N-3}^{(r+1)} + \frac{3\alpha}{2h^3} U_{N-2}^{(r+1)} + \Big[-\frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\Big] U_{N-1}^{(r+1)}$$
$$= 3\alpha f_0^{(r)} - \frac{h^2}{3} f_0^{(r)} - \frac{41h^2}{12} f_1^{(r)} - \frac{7h^2}{6} f_2^{(r)} - \frac{h^2}{12} f_3^{(r)} + 3\alpha f_N^{(r)} - \Big[-4 - \frac{h^2}{3} q_0^{(r)}$$
$$+ \frac{\alpha}{2h^3} + 3\alpha q_0^{(r)} + \frac{6\alpha}{h}\Big]\eta_0 - \Big[\frac{\alpha}{2h^3} + 3\alpha q_N^{(r)} + \frac{6\alpha}{h}\Big]\eta_1. \qquad (7.23)$$

Similarly, we define

$$T_{N-1}^{(r)}(h) =$$
$$\Big[\frac{\alpha}{2h^3} + 3\alpha q^{(r)}(x_0) + \frac{6\alpha}{h}\Big] u^{(r+1)}(x_0) + \Big[-\frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\Big] u^{(r+1)}(x_1) + \frac{3\alpha}{2h^3} u^{(r+1)}(x_2)$$
$$- \frac{\alpha}{2h^3} u^{(r+1)}(x_3) - \frac{\alpha}{2h^3} u^{(r+1)}(x_{N-3}) + \frac{3\alpha}{2h^3} u^{(r+1)}(x_{N-2}) + \Big[-\frac{3\alpha}{2h^3}$$
$$- \frac{6\alpha}{h}\Big] u^{(r+1)}(x_{N-1}) + \Big[\frac{\alpha}{2h^3} + 3\alpha q^{(r)}(x_N) + \frac{6\alpha}{h}\Big] u^{(r+1)}(x_N) - 3\alpha f^{(r)}(x_0)$$
$$- 3\alpha f^{(r)}(x_N) + \sum_{i=N-3}^{N} \Big(c_i u^{(r+1)}(x_i) + d_i h^2 u_{xx}^{(r+1)}(x_i)\Big).$$

By the assumption $|\alpha| < h^6$ and $c_N = -4$, $c_{N-1} = 7$, $c_{N-2} = -2$, $c_{N-3} = -1$, $d_N = \frac{1}{3}$, $d_{N-1} = \frac{41}{12}$, $d_{N-2} = \frac{7}{6}$, $d_{N-3} = \frac{1}{12}$, we get $T_{N-1}^{(r)}(h) = \mathcal{O}(h^6)$. Hence, we get the difference equation (associated with $T_{N-1}^{(r)}(h)$) as follows:

$$\Big[-\frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\Big] U_1^{(r+1)} + \frac{3\alpha}{2h^3} U_2^{(r+1)} - \frac{\alpha}{2h^3} U_3^{(r+1)} + \Big[-1 - \frac{h^2}{12} q_{N-3}^{(r)} - \frac{\alpha}{2h^3}\Big] U_{N-3}^{(r+1)}$$
$$+ \Big[-2 - \frac{7h^2}{6} q_{N-2}^{(r)} + \frac{3\alpha}{2h^3}\Big] U_{N-2}^{(r+1)} + \Big[7 - \frac{41h^2}{12} q_{N-1}^{(r)} - \frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\Big] U_{N-1}^{(r+1)}$$
$$= 3\alpha f_0^{(r)} - \frac{h^2}{3} f_N^{(r)} - \frac{41h^2}{12} f_{N-1}^{(r)} - \frac{7h^2}{6} f_{N-2}^{(r)} - \frac{h^2}{12} f_{N-3}^{(r)} + 3\alpha f_N^{(r)} - \Big[\frac{\alpha}{2h^3} + 3\alpha q_0^{(r)}$$
$$+ \frac{6\alpha}{h}\Big]\eta_0 - \Big[-4 - \frac{h^2}{3} q_N^{(r)} + \frac{\alpha}{2h^3} + 3\alpha q_N^{(r)} + \frac{6\alpha}{h}\Big]\eta_1. \qquad (7.24)$$

As $h \to 0$, we have $|T_{N-1}^{(r)}(h)| \le K_{N-1}^{(r)} h^6$, where $K_{N-1}^{(r)}$ is a constant. Hence $T_{N-1}^{(r)}(h) = \mathcal{O}(h^6)$ as $h \to 0$. We get the truncation error $T_i^{(r)}(h)$, $i = 2, 3, \ldots, N-2$ associated with the equations (7.20) as follows:

$$\begin{aligned}
&T_i^{(r)}(h) = \\
&\Big[\frac{\alpha}{2h^3} + 3\alpha q^{(r)}(x_0) + \frac{6\alpha}{h}\Big]u^{(r+1)}(x_0) + \Big[-\frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\Big]u^{(r+1)}(x_1) \\
&+ \frac{3\alpha}{2h^3}u^{(r+1)}(x_2) - \frac{\alpha}{2h^3}u^{(r+1)}(x_3) + \Big[-1 - \frac{h^2}{20}q^{(r)}(x_{i-2})\Big]u^{(r+1)}(x_{i-2}) \\
&+ \Big[-2 - \frac{13h^2}{10}q^{(r)}(x_{i-1})\Big]u^{(r+1)}(x_{i-1}) + \Big[6 - \frac{33h^2}{10}q^{(r)}(x_i)\Big]u^{(r+1)}(x_i) \\
&+ \Big[-2 - \frac{13h^2}{10}q^{(r)}(x_{i+1})\Big]u^{(r+1)}(x_{i+1}) + \Big[-1 - \frac{h^2}{20}q^{(r)}(x_{i+2})\Big]u^{(r+1)}(x_{i+2}) \\
&- \frac{\alpha}{2h^3}u^{(r+1)}(x_{N-3}) + \frac{3\alpha}{2h^3}u^{(r+1)}(x_{N-2}) + \Big[-\frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\Big]u^{(r+1)}(x_{N-1}) \\
&+ \Big[\frac{\alpha}{2h^3} + 3\alpha q^{(r)}(x_N) + \frac{6\alpha}{h}\Big]u^{(r+1)}(x_N) - 3\alpha f^{(r)}(x_0) + \frac{h^2}{20}f^{(r)}(x_{i-2}) \\
&+ \frac{13h^2}{10}f^{(r)}(x_{i-1}) + \frac{33h^2}{10}f^{(r)}(x_i) + \frac{13h^2}{10}f^{(r)}(x_{i+1}) + \frac{h^2}{20}f^{(r)}(x_{i+2}) - 3\alpha f^{(r)}(x_N).
\end{aligned} \tag{7.25}$$

Substituting $f^{(r)}(x_i) = u_{xx}^{(r+1)}(x_i) + q^{(r)}(x_i)u^{(r+1)}(x_i)$ in (7.25), we get

$$\begin{aligned}
T_i^{(r)}(h) =& -\frac{\alpha}{2}\Big[\frac{-u^{(r+1)}(x_0) + 3u^{(r+1)}(x_1) - 3u^{(r+1)}(x_2) + u^{(r+1)}(x_3)}{h^3}\Big] \\
&+ \frac{\alpha}{2}\Big[\frac{u^{(r+1)}(x_N) - 3u^{(r+1)}(x_{N-1}) + 3u^{(r+1)}(x_{N-2}) - u^{(r+1)}(x_{N-3})}{h^3}\Big] \\
&- 6\alpha\Big[\frac{-u^{(r+1)}(x_0) + u^{(r+1)}(x_1)}{h}\Big] + 6\alpha\Big[\frac{u^{(r+1)}(x_N) - u^{(r+1)}(x_{N-1})}{h}\Big] \\
&- 3\alpha u_{xx}^{(r+1)}(x_0) - 3\alpha u_{xx}^{(r+1)}(x_N) - u^{(r+1)}(x_{i-2}) - 2u^{(r+1)}(x_{i-1}) + 6u^{(r+1)}(x_i) \\
&- 2u^{(r+1)}(x_{i+1}) - u^{(r+1)}(x_{i+2}) + \frac{h^2}{20}u_{xx}^{(r+1)}(x_{i-2}) + \frac{13h^2}{10}u_{xx}^{(r+1)}(x_{i-1}) \\
&+ \frac{33h^2}{10}u_{xx}^{(r+1)}(x_i) + \frac{13h^2}{10}u_{xx}^{(r+1)}(x_{i+1}) + \frac{h^2}{20}u_{xx}^{(r+1)}(x_{i+2}).
\end{aligned}$$

Further simplify, we obtain

$$\begin{aligned}
T_i^{(r)}(h) =& -\frac{\alpha}{2}\Big[u_{xxx}^{(r+1)}(x_0) + \mathcal{O}(h)\Big] + \frac{\alpha}{2}\Big[u_{xxx}^{(r+1)}(x_N) + \mathcal{O}(h)\Big] \\
&- 6\alpha\Big[u_x^{(r+1)}(x_0) + \mathcal{O}(h)\Big] + 6\alpha\Big[u_x^{(r+1)}(x_N) + \mathcal{O}(h)\Big] \\
&- 3\alpha u_{xx}^{(r+1)}(x_0) - 3\alpha u_{xx}^{(r+1)}(x_N) - \frac{h^6}{120}u_{xxxxxx}^{(r+1)}(x_i) + \mathcal{O}(h^7),
\end{aligned}$$

$i = 2, 3, \ldots, N-2$. As $h \to 0$, we have $|T_i^{(r)}(h)| \leq K_i^{(r)}h^6$, where $K_i^{(r)}$ is a constant. It can be seen that as $h \to 0$, $T_i^{(r)}(h) = \mathcal{O}(h^6)$, $i = 2, 3, \ldots, N-2$.

The system given in (7.23), (7.20) and (7.24) can be written in the following form $A^{(r)}U^{(r+1)} = d^{(r)}$, where

$$A^{(r)} = \begin{bmatrix}
A^{(r)}_{1,1} & A^{(r)}_{1,2} & A^{(r)}_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{1,N-3} & A^{(r)}_{1,N-2} & A^{(r)}_{1,N-1} \\
A^{(r)}_{2,1} & A^{(r)}_{2,2} & A^{(r)}_{2,3} & A^{(r)}_{2,4} & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{2,N-3} & A^{(r)}_{2,N-2} & A^{(r)}_{2,N-1} \\
A^{(r)}_{3,1} & A^{(r)}_{3,2} & A^{(r)}_{3,3} & A^{(r)}_{3,4} & A^{(r)}_{3,5} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{3,N-3} & A^{(r)}_{3,N-2} & A^{(r)}_{3,N-1} \\
A^{(r)}_{4,1} & A^{(r)}_{4,2} & A^{(r)}_{4,3} & A^{(r)}_{4,4} & A^{(r)}_{4,5} & A^{(r)}_{4,6} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{4,N-3} & A^{(r)}_{4,N-2} & A^{(r)}_{4,N-1} \\
A^{(r)}_{5,1} & A^{(r)}_{5,2} & A^{(r)}_{5,3} & A^{(r)}_{5,4} & A^{(r)}_{5,5} & A^{(r)}_{5,6} & A^{(r)}_{5,7} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{5,N-3} & A^{(r)}_{5,N-2} & A^{(r)}_{5,N-1} \\
A^{(r)}_{6,1} & A^{(r)}_{6,2} & A^{(r)}_{6,3} & A^{(r)}_{6,4} & A^{(r)}_{6,5} & A^{(r)}_{6,6} & A^{(r)}_{6,7} & A^{(r)}_{6,8} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{6,N-3} & A^{(r)}_{6,N-2} & A^{(r)}_{6,N-1} \\
A^{(r)}_{7,1} & A^{(r)}_{7,2} & A^{(r)}_{7,3} & 0 & A^{(r)}_{7,5} & A^{(r)}_{7,6} & A^{(r)}_{7,7} & A^{(r)}_{7,8} & A^{(r)}_{7,9} & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{7,N-3} & A^{(r)}_{7,N-2} & A^{(r)}_{7,N-1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
A^{(r)}_{N-7,1} & A^{(r)}_{N-7,2} & A^{(r)}_{N-7,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & A^{(r)}_{N-7,N-9} & A^{(r)}_{N-7,N-8} & A^{(r)}_{N-7,N-7} & A^{(r)}_{N-7,N-6} & A^{(r)}_{N-7,N-5} & 0 & A^{(r)}_{N-7,N-3} & A^{(r)}_{N-7,N-2} & A^{(r)}_{N-7,N-1} \\
A^{(r)}_{N-6,1} & A^{(r)}_{N-6,2} & A^{(r)}_{N-6,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & A^{(r)}_{N-6,N-8} & A^{(r)}_{N-6,N-7} & A^{(r)}_{N-6,N-6} & A^{(r)}_{N-6,N-5} & A^{(r)}_{N-6,N-4} & A^{(r)}_{N-6,N-3} & A^{(r)}_{N-6,N-2} & A^{(r)}_{N-6,N-1} \\
A^{(r)}_{N-5,1} & A^{(r)}_{N-5,2} & A^{(r)}_{N-5,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & A^{(r)}_{N-5,N-7} & A^{(r)}_{N-5,N-6} & A^{(r)}_{N-5,N-5} & A^{(r)}_{N-5,N-4} & A^{(r)}_{N-5,N-3} & A^{(r)}_{N-5,N-2} & A^{(r)}_{N-5,N-1} \\
A^{(r)}_{N-4,1} & A^{(r)}_{N-4,2} & A^{(r)}_{N-4,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & A^{(r)}_{N-4,N-6} & A^{(r)}_{N-4,N-5} & A^{(r)}_{N-4,N-4} & A^{(r)}_{N-4,N-3} & A^{(r)}_{N-4,N-2} & A^{(r)}_{N-4,N-1} \\
A^{(r)}_{N-3,1} & A^{(r)}_{N-3,2} & A^{(r)}_{N-3,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{N-3,N-5} & A^{(r)}_{N-3,N-4} & A^{(r)}_{N-3,N-3} & A^{(r)}_{N-3,N-2} & A^{(r)}_{N-3,N-1} \\
A^{(r)}_{N-2,1} & A^{(r)}_{N-2,2} & A^{(r)}_{N-2,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{N-2,N-4} & A^{(r)}_{N-2,N-3} & A^{(r)}_{N-2,N-2} & A^{(r)}_{N-2,N-1} \\
A^{(r)}_{N-1,1} & A^{(r)}_{N-1,2} & A^{(r)}_{N-1,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A^{(r)}_{N-1,N-3} & A^{(r)}_{N-1,N-2} & A^{(r)}_{N-1,N-1}
\end{bmatrix}$$

$A^{(r)}_{i,j}$ is the coefficient of $U^{(r+1)}_j$, $U^{(r+1)} = (U^{(r+1)}_1, U^{(r+1)}_2, \dots, U^{(r+1)}_{N-1})^T$, $d^{(r)} = (d^{(r)}_1$, $d^{(r)}_2$, $\dots$, $d^{(r)}_{N-1})^T$, $d^{(r)}_i$ is the righthand side of the system. Let $B^{(r)}$ be the matrix corresponding to $\alpha = 0$. We get $\|A^{(r)} - B^{(r)}\|_\infty = \max_i \sum_{j=1}^{N-1} |A^{(r)}_{i,j} - B^{(r)}_{i,j}|$. Thus, we get

$$\|A^{(r)} - B^{(r)}\|_\infty = 2\Big| -\frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\Big| + 2\Big|\frac{3\alpha}{2h^3}\Big| + 2\Big| -\frac{\alpha}{2h^3}\Big|.$$

We take the matrices $A^{(r)}$, $B^{(r)}$, $n = N-1$, $\|.\|_\infty$ norm in the Theorem 1.7.4 and hence, we get

$$\max_i |\lambda^{(r)}_i - \mu^{(r)}_i| \leq 2^{\frac{2N-3}{N-1}} (N-1)^{\frac{1}{N-1}} (2Q^{(r)})^{\frac{N-2}{N-1}} \|A^{(r)} - B^{(r)}\|_\infty^{\frac{1}{N-1}} \tag{7.26}$$

where $Q^{(r)} = \max\{\|A^{(r)}\|_\infty, \|B^{(r)}\|_\infty\}$, $\lambda^{(r)}_i$ and $\mu^{(r)}_i$, $i = 1, 2, \dots, N-1$ are the eigenvalues of $A^{(r)}$ and $B^{(r)}$ respectively.

Note that, when h is sufficiently small, we get $B^{(r)}$ is irreducible, $B^{(r)}_{i,i} > 0$, $B^{(r)}_{i,j} \leq 0$, $i \neq j$, $i, j = 1, 2, \dots, N-1$, row sums

$$\mathcal{R}^{(r)}_1 = 4 - \frac{41h^2}{12} q^{(r)}_1 - \frac{7h^2}{6} q^{(r)}_2 - \frac{h^2}{12} q^{(r)}_3 > 0,$$

$$\mathcal{R}^{(r)}_2 = 1 - \frac{13h^2}{10} q^{(r)}_1 - \frac{33h^2}{10} q^{(r)}_2 - \frac{13h^2}{10} q^{(r)}_3 - \frac{h^2}{20} q^{(r)}_4 > 0,$$

$$\mathcal{R}^{(r)}_i = -\frac{h^2}{20} q^{(r)}_{i-2} - \frac{13h^2}{10} q^{(r)}_{i-1} - \frac{33h^2}{10} q^{(r)}_i - \frac{13h^2}{10} q^{(r)}_{i+1} - \frac{h^2}{20} q^{(r)}_{i+2} > 0,$$

$i = 3, 4, \dots, N-3$,

$$\mathcal{R}^{(r)}_{N-2} = 1 - \frac{13h^2}{10} q^{(r)}_{N-1} - \frac{33h^2}{10} q^{(r)}_{N-2} - \frac{13h^2}{10} q^{(r)}_{N-3} - \frac{h^2}{20} q^{(r)}_{N-4} > 0,$$

$$\mathcal{R}_{N-1}^{(r)} = 4 - \frac{41h^2}{12}q_{N-1}^{(r)} - \frac{7h^2}{6}q_{N-2}^{(r)} - \frac{h^2}{12}q_{N-3}^{(r)} > 0.$$

Hence $B^{(r)}$ is monotone matrix [67]. Thus ${B^{(r)}}^{-1}$ exist and eigenvalues $\mu_i^{(r)}$, $i = 1, 2, \ldots, N-1$ of $B^{(r)}$ are non-zero when h sufficiently small. Once we fix the h, then α can be suitably chosen in the region $(-h^6, h^6)$ to satisfy the following conditions:

- $A^{(r)}$ is invertible because

$$\|A^{(r)} - B^{(r)}\|_\infty = 2\left| -\frac{3\alpha}{2h^3} - \frac{6\alpha}{h}\right| + 2\left|\frac{3\alpha}{2h^3}\right| + 2\left| -\frac{\alpha}{2h^3}\right|$$

  and from (7.26), it can been seen that, eigenvalues of $A^{(r)}$ are non-zero when α is sufficiently small.

- The row sum of $A^{(r)}$,

$$\mathcal{S}_i^{(r)} = \mathcal{R}_i^{(r)} - \frac{12\alpha}{h} - \frac{\alpha}{h^3} > 0, \; i = 1, 2, \ldots, N-1,$$

  when α is sufficiently small.

Once we fix h is sufficiently small (which gives $B^{(r)}$ is monotone) and α is sufficiently small in $(-h^6, h^6)$ (which gives $A^{(r)}$ is invertible and row sums of $A^{(r)}$ are positive), we get the following:

The system (7.23), (7.20) and (7.24) with exact solutions can be written as

$$A^{(r)}\bar{u}^{(r+1)} = d^{(r)} + T^{(r)}(h)$$

where $\bar{u}^{(r+1)} = (u^{(r+1)}(x_1), u^{(r+1)}(x_2), \ldots, u^{(r+1)}(x_{N-1}))^T$ and $T^{(r)}(h) = (T_1^{(r)}(h), T_2^{(r)}(h), \ldots, T_{N-1}^{(r)}(h))^T$. Since

$$A^{(r)}U^{(r+1)} = d^{(r)},$$

we get

$$A^{(r)}(\bar{u}^{(r+1)} - U^{(r+1)}) = T^{(r)}(h),$$

that is,

$$A^{(r)}E^{(r+1)} = T^{(r)}(h),$$

where $E^{(r+1)} = (E_1^{(r+1)}, E_2^{(r+1)}, \ldots, E_{N-1}^{(r+1)})^T$, $E_i^{(r+1)} = u^{(r+1)}(x_i) - U_i^{(r+1)}$. Consequently, we obtain

$$E^{(r+1)} = {A^{(r)}}^{-1} T^{(r)}(h).$$

By the definition of multiplication of a matrix by its inverse, we have

$$\sum_{i=1}^{N-1} {A^{(r)}}_{j,i}^{-1} \mathcal{S}_i^{(r)} = 1,\ j = 1, 2, \ldots, N-1,$$

where ${A^{(r)}}_{j,i}^{-1}$ is the (j,i)-th element of the matrix ${A^{(r)}}^{-1}$. Hence, we have

$$\sum_{i=1}^{N-1} {A^{(r)}}_{j,i}^{-1} \leq 1/ \min_{1\leq i\leq N-1} \mathcal{S}_i^{(r)} = 1/C_{i_0}^{(r)} h^2,$$

where $C_{i_0}^{(r)}$ is a constant. We have

$$E_j^{(r+1)} = \sum_{i=1}^{N-1} {A^{(r)}}_{j,i}^{-1} T_i^{(r)}(h),\ j = 1, 2, \ldots, N-1,$$

and hence $|E_j^{(r+1)}| \leq \frac{K^{(r)} h^6}{C_{i_0}^{(r)} h^2}$, where $K^{(r)}$ is a constant. Therefore, as $h \to 0$, we have

$$\|E^{(r+1)}\|_\infty = \mathcal{O}(h^4).$$

7.2.1 Numerical examples

The numerical solutions of the non-linear BVPs are computed as follows: For each fixed N, by taking initial approximation $U_0^{(0)}, U_1^{(0)}, \ldots, U_N^{(0)}$, we compute the numerical solutions $U^{(r+1)} = (U_1^{(r+1)}, U_2^{(r+1)}, \ldots, U_{N-1}^{(r+1)})$, $r = 0, 1, \ldots$. We can take sufficient number of iterations so that maximum error between two successive iterations $\max_i |U_i^{(r+1)} - U_i^{(r)}| < \delta$, where δ is the small tolerance value that is prescribed. Once the criterion is satisfied, we consider $U^{(r+1)}$ as the numerical solution U to non-linear BVP. To compute the numerical solutions of each of the following non-linear BVPs, we take $\delta = 10^{-15}$ and 8 iterations.

Example 7.2.1. *Consider the BVP [43]*

$$\begin{cases} u_{xx}(x) + \exp(-2 * u(x)) = 0, & x \in (0, 1), \\ u(0) = 0, \quad u(1) = \log(2). \end{cases}$$

The exact solution to the above given BVP is $u(x) = \log(1+x)$. We calculate the numerical solutions for $N = 16, 32, 64, 128$. The scaling factors $\alpha = 0.9999h^6$ are used to calculate the numerical solutions. The maximum point-wise error and the order of convergence are given in Table 7.5. Table 7.6 represents the maximum point-wise error and the order of convergence corresponding to $\alpha = 0$. We take $U_0^{(0)} = 0$, $U_i^{(0)} = 1$, $i = 1, 2, \ldots, N-1$, $U_N^{(0)} = \log(2)$ to compute the numerical solutions.

Table 7.5: Maximum point-wise error and order of convergence of Example 7.2.1.

N	16	32	64	128
E^N	1.3093e-07	4.6024e-09	1.6823e-10	6.9627e-12
p^N	4.8303	4.7739	4.5946	

Table 7.6: Maximum point-wise error an order of convergence of Example 7.2.1 with $\alpha = 0$.

N	16	32	64	128
E^N	2.9388e-08	2.4273e-09	1.6374e-10	1.0454e-11
p^N	3.5978	3.8898	3.9694	

Example 7.2.2. *Consider the BVP [39]*

$$\begin{cases} u_{xx}(x) - \dfrac{(2-x)\exp(2u(x)) + (1/(1+x))}{3} = 0, \quad x \in (0,1), \\ u(0) = 0, \quad u(1) = \log(1/2). \end{cases}$$

The exact solution to the above given BVP is $u(x) = \log(1/(1+x))$. We calculate the numerical solutions for $N = 16$, 32, 64, 128. The scaling factors $\alpha = 0.9999h^6$ are used to calculate the numerical solutions. The maximum point-wise error and the order of convergence are given in Table 7.7. The maximum point-wise error and the order of convergence corresponding to $\alpha = 0$ are given in 7.8. We take $U_0^{(0)} = 0$, $U_i^{(0)} = 1$, $i = 1, 2, \ldots, N-1$, $U_N^{(0)} = \log(1/2)$ to compute the numerical solutions.

Table 7.7: Maximum point-wise error and order of convergence of Example 7.2.2.

N	16	32	64	128
E^N	1.3680e-07	4.8082e-09	1.7523e-10	7.2051e-12
p^N	4.8304	4.7782	4.6041	

Table 7.8: Maximum point-wise error and order of convergence of Example 7.2.2 with $\alpha = 0$.

N	16	32	64	128
E^N	3.0589e-08	2.5274e-09	1.7045e-10	1.0874e-11
p^N	3.5973	3.8903	3.9704	

7.3 Conclusion

In this chapter, we have developed the numerical methods to get the numerical solutions for the singularly perturbed BVPs and the nonlinear BVPs. The convergence analysis of the developed methods are established and the constructed methods have fourth-order convergence. Proposed methods can be implemented easily on a computer and test examples have been solved to demonstrate the efficiency of the proposed methods. The construction of higher-order numerical methods for nonlinear BVPs are challenging tasks. In this chapter, we obtained fourth-order numerical method for nonlinear BVPs using quasilinearization technique and quintic spline.

Chapter 8

Fractal Quintic Spline Solutions for Fourth-Order Boundary-Value Problems

In this chapter, we consider the fourth-order BVPs of the form

$$\begin{cases} u^{(4)}(x) + p(x)u(x) = f(x),\ x \in (0,1), \\ u(0) = \eta_0,\ u(1) = \eta_1,\ u''(0) = \hat{\eta_0},\ u''(1) = \hat{\eta_1}, \end{cases} \tag{8.1}$$

and

$$\begin{cases} u^{(4)}(x) + p(x)u(x) = f(x),\ x \in (0,1), \\ u(0) = \eta_0,\ u(1) = \eta_1,\ u'(0) = \hat{\eta_0},\ u'(1) = \hat{\eta_1}, \end{cases} \tag{8.2}$$

where the functions p and f are continuous in $I = [0,1]$. The problems of these kind arise in the plate deflection theory. The analytical solution of these BVPs can not be obtained for the arbitrary choices of the functions p and f.

Usmani [137] proposed the second-order convergent method using quartic splines to compute the approximation to the solution of the fourth-order BVP given in (8.1). Usmani and Warsi [139] developed the second-order convergent method using quintic spline and the fourth-order convergent method using sextic spline to get the numerical approximation to the solution of the fourth-order BVP of the type given in (8.1). To get the numerical solutions of the fourth-order BVPs of the type given in (8.2), Usmani [136] established the second-order and the fourth-order convergent methods using quintic and sextic splines respectively. Siddiqi and Akram [130] developed the second-order convergent numerical method using quintic spline for fourth-order BVPs. To get the numerical solutions for fourth-order BVPs, Ramadan, Lashien and zahra [107] constructed the fourth-order convergent method using the non-polynomial quintic spline. Also, one can see in References [4,5,88,109,129,140], where variety of numerical methods have been proposed by different authors for fourth-order BVPs.

In Section 8.1, a numerical method for BVP given in (8.1) is established. The truncation error of the developed method is derived. Numerical examples are provided to support the theoretical results. In Section 8.2, a numerical method is derived for the BVP given in (8.2). The truncation error of the established method is derived and numerical examples are provided to test the proposed method.

8.1 Numerical Method

Let us consider the fourth-order BVP (8.1)

$$\begin{cases} u^{(4)}(x) + p(x)u(x) = f(x), \ x \in (0,1), \\ u(0) = \eta_0, \ u(1) = \eta_1, \ u''(0) = \hat{\eta}_0, \ u''(1) = \hat{\eta}_1. \end{cases}$$

Let $0 = x_0 < x_1 < \cdots < x_N = 1$ be the uniform partition of the interval $I = [0,1]$. Let $u(x)$ be the solution of the differential equation given in (8.1). Let U_i, M_i and S_i be the approximations to $u(x_i)$, $u^{(2)}(x_i)$ and $u^{(4)}(x_i)$ respectively. From (7.4), we get

$$\begin{aligned} M_{i+1} + 4M_i + M_{i-1} = & -\frac{\alpha}{4h^2}S_0 - \frac{\alpha}{4h^2}S_N + \frac{7h^2}{60}S_{i-1} + \frac{16h^2}{60}S_i + \frac{7h^2}{60}S_{i+1} \\ & + \frac{6\alpha}{h^2}\Phi^{(1)}(x_0) - \frac{6\alpha}{h^2}\Phi^{(1)}(x_N) + \frac{3\alpha}{h^2}M_0 + \frac{3\alpha}{h^2}M_N \\ & + \frac{6}{h^2}(U_{i+1} - 2U_i + U_{i-1}). \end{aligned} \tag{8.3}$$

From (7.5), we have

$$\begin{aligned} M_{i+1} - 2M_i + M_{i-1} = & -\frac{\alpha}{2h^2}S_0 - \frac{\alpha}{2h^2}S_N + \frac{h^2}{6}(S_{i+1} + 4S_i + S_{i-1}) \\ & - \frac{\alpha}{h^2}\Phi^{(3)}(x_0) + \frac{\alpha}{h^2}\Phi^{(3)}(x_N). \end{aligned} \tag{8.4}$$

Subtract (8.4) from (8.3) to get

$$M_i = \frac{\alpha}{24h^2}(S_0+S_N) + \frac{\alpha}{h^2}(\Phi^{(1)}(x_0) - \Phi^{(1)}(x_N)) + \frac{\alpha}{2h^2}(M_0+M_N) + \frac{\alpha}{6h^2}(\Phi^{(3)}(x_0) - \Phi^{(3)}(x_N)) + \frac{1}{h^2}(U_{i+1} - 2U_i + U_{i-1}) - \frac{h^2}{120}(S_{i+1} + 8S_i + S_{i-1}). \tag{8.5}$$

By substituting (8.5) in (8.4) and after some calculation, we obtain

$$U_{i-2} - 4U_{i-1} + 6U_i - 4U_{i+1} + U_{i+2} = -\frac{\alpha}{2}(S_0+S_N) - \alpha\Phi^{(3)}(x_0) + \alpha\Phi^{(3)}(x_N) + \frac{h^4}{120}(S_{i+2} + 26S_{i+1} + 66S_i + 26S_{i-1} + S_{i-2}). \tag{8.6}$$

The relation given in (8.6) gives the relation between U_i's and S_i's and hence it can be used to solve the BVP (8.1). The differential equation (8.1) is discretized at $x = x_i$ as $S_i + p_iU_i = f_i$, where $p_i = p(x_i)$, $f_i = f(x_i)$. The boundary conditions are discretized as $U_0 = \eta_0$, $U_N = \eta_1$, $M_0 = \hat{\eta_0}$, $M_N = \hat{\eta_1}$.

Substitute

$$\Phi^{(3)}(x_0) = \frac{-5U_0 + 18U_1 - 24U_2 + 14U_3 - 3U_4}{2h^3},\ S_i = f_i - p_iU_i,$$
$$\Phi^{(3)}(x_N) = \frac{5U_N - 18U_{N-1} + 24U_{N-2} - 14U_{N-3} + 3U_{N-4}}{2h^3}$$

in (8.6) and after some calculation we get,

$$\begin{cases} \frac{9\alpha}{h^3}U_1 - \frac{12\alpha}{h^3}U_2 + \frac{7\alpha}{h^3}U_3 - \frac{3\alpha}{2h^3}U_4 + \Big[1 + \frac{h^4p_{i-2}}{120}\Big]U_{i-2} + \Big[-4 \\ + \frac{13h^4p_{i-1}}{60}\Big]U_{i-1} + \Big[6 + \frac{11h^4p_i}{20}\Big]U_i + \Big[-4 + \frac{13h^4p_{i+1}}{60}\Big]U_{i+1} \\ + \Big[1 + \frac{h^4p_{i+2}}{120}\Big]U_{i+2} - \frac{3\alpha}{2h^3}U_{N-4} + \frac{7\alpha}{h^3}U_{N-3} - \frac{12\alpha}{h^3}U_{N-2} + \frac{9\alpha}{h^3}U_{N-1} \\ = \frac{h^4}{120}\Big[f_{i-2} + 26f_{i-1} + 66f_i + 26f_{i+1} + f_{i+2}\Big] - \frac{\alpha}{2}(f_0 + f_N) \\ + \frac{\alpha}{2}(p_0\eta_0 + p_N\eta_1) + \frac{5\alpha}{2h^3}\eta_0 + \frac{5\alpha}{2h^3}\eta_1,\ i = 2, 3, \ldots, N-2. \end{cases} \tag{8.7}$$

In the system (8.7), there are $(N-3)$ equations with $(N-1)$ unknowns U_1, U_2, $\ldots$, U_{N-1}. Therefore, we need another two equations to compute the numerical solutions.

Construction of Difference Equations: Substitute $i = 1, 2$ in (8.3) and (8.4), we get

$$M_2 + 4M_1 + M_0 = -\frac{\alpha}{4h^2}S_0 - \frac{\alpha}{4h^2}S_N + \frac{7h^2}{60}S_0 + \frac{16h^2}{60}S_1 + \frac{7h^2}{60}S_2$$

$$+\frac{6\alpha}{h^2}\Phi^{(1)}(x_0)-\frac{6\alpha}{h^2}\Phi^{(1)}(x_N)+\frac{3\alpha}{h^2}M_0+\frac{3\alpha}{h^2}M_N$$
$$+\frac{6}{h^2}(U_2-2U_1+U_0), \tag{8.8}$$
$$M_3+4M_2+M_1=-\frac{\alpha}{4h^2}S_0-\frac{\alpha}{4h^2}S_N+\frac{7h^2}{60}S_1+\frac{16h^2}{60}S_2+\frac{7h^2}{60}S_3$$
$$+\frac{6\alpha}{h^2}\Phi^{(1)}(x_0)-\frac{6\alpha}{h^2}\Phi^{(1)}(x_N)+\frac{3\alpha}{h^2}M_0+\frac{3\alpha}{h^2}M_N$$
$$+\frac{6}{h^2}(U_3-2U_2+U_1), \tag{8.9}$$
$$M_2-2M_1+M_0=-\frac{\alpha}{2h^2}S_0-\frac{\alpha}{2h^2}S_N+\frac{h^2}{6}(S_2+4S_1+S_0)$$
$$-\frac{\alpha}{h^2}\Phi^{(3)}(x_0)+\frac{\alpha}{h^2}\Phi^{(3)}(x_N), \tag{8.10}$$
$$M_3-2M_2+M_1=-\frac{\alpha}{2h^2}S_0-\frac{\alpha}{2h^2}S_N+\frac{h^2}{6}(S_3+4S_2+S_1)$$
$$-\frac{\alpha}{h^2}\Phi^{(3)}(x_0)+\frac{\alpha}{h^2}\Phi^{(3)}(x_N). \tag{8.11}$$

From (8.8) and (8.10), also from (8.9) and (8.11), we get

$$M_1=\frac{\alpha}{24h^2}(S_0+S_N)+\frac{\alpha}{h^2}(\Phi^{(1)}(x_0)-\Phi^{(1)}(x_N))+\frac{\alpha}{2h^2}(M_0+M_N)+\frac{\alpha}{6h^2}(\Phi^{(3)}(x_0)$$
$$-\Phi^{(3)}(x_N))+\frac{1}{h^2}(U_2-2U_1+U_0)-\frac{h^2}{120}(S_2+8S_1+S_0), \tag{8.12}$$
$$M_2=\frac{\alpha}{24h^2}(S_0+S_N)+\frac{\alpha}{h^2}(\Phi^{(1)}(x_0)-\Phi^{(1)}(x_N))+\frac{\alpha}{2h^2}(M_0+M_N)+\frac{\alpha}{6h^2}(\Phi^{(3)}(x_0)$$
$$-\Phi^{(3)}(x_N))+\frac{1}{h^2}(U_3-2U_2+U_1)-\frac{h^2}{120}(S_3+8S_2+S_1). \tag{8.13}$$

Substitute (8.12) and (8.13) in (8.10) to get

$$-2U_0+5U_1-4U_2+U_3=-M_0h^2-\frac{11\alpha}{24}(S_0+S_N)+\alpha(\Phi^{(1)}(x_0)-\Phi^{(1)}(x_N))$$
$$+\frac{\alpha}{2}(M_0+M_N)+\frac{5\alpha}{6}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0))+\frac{h^4}{120}(18S_0+65S_1+26S_2+S_3). \tag{8.14}$$

Substitute

$$\Phi^{(1)}(x_0)=\frac{-3U_0+4U_1-U_2}{2h},\ \Phi^{(1)}(x_N)=\frac{3U_N-4U_{N-1}+U_{N-2}}{2h},\ S_i=f_i-p_iU_i,$$

approximation for $\Phi^{(3)}(x_0)$ and $\Phi^{(3)}(x_N)$ in (8.14), we obtain the difference equation as

$$\Big[5+\frac{13h^4p_1}{24}-\frac{2\alpha}{h}+\frac{15\alpha}{2h^3}\Big]U_1+\Big[-4+\frac{13h^4p_2}{60}+\frac{\alpha}{2h}-\frac{10\alpha}{h^3}\Big]U_2$$
$$+\Big[1+\frac{h^4p_3}{120}+\frac{35\alpha}{6h^3}\Big]U_3-\frac{5\alpha}{4h^3}U_4-\frac{5\alpha}{4h^3}U_{N-4}+\frac{35\alpha}{6h^3}U_{N-3}+\Big[-\frac{10\alpha}{h^3}$$

$$+\frac{\alpha}{2h}\Big]U_{N-2}+\Big[-\frac{2\alpha}{h}+\frac{15\alpha}{2h^3}\Big]U_{N-1}=-\hat{\eta_0}h^2+\frac{\alpha}{2}(\hat{\eta_0}+\hat{\eta_1})$$
$$-\frac{11\alpha}{24}(f_0+f_N)+\frac{11\alpha}{24}(p_0\eta_0+p_N\eta_1)+\frac{h^4}{120}(18f_0+65f_1+26f_2+f_3)$$
$$-\Big[-2+\frac{3h^4p_0}{20}+\frac{3\alpha}{2h}-\frac{25\alpha}{12h^3}\Big]\eta_0-\Big[-\frac{25\alpha}{12h^3}+\frac{3\alpha}{2h}\Big]\eta_1. \tag{8.15}$$

Similarly, we can get the difference equation as follows: Substitute $i = N-1, N-2$ in (8.3) and (8.4), we get

$$M_{N-2}+4M_{N-1}+M_N=-\frac{\alpha}{4h^2}S_0-\frac{\alpha}{4h^2}S_N+\frac{7h^2}{60}S_N+\frac{16h^2}{60}S_{N-1}+\frac{7h^2}{60}S_{N-2}$$
$$+\frac{6\alpha}{h^2}\Phi^{(1)}(x_0)-\frac{6\alpha}{h^2}\Phi^{(1)}(x_N)+\frac{3\alpha}{h^2}M_0+\frac{3\alpha}{h^2}M_N$$
$$+\frac{6}{h^2}(U_{N-2}-2U_{N-1}+U_N), \tag{8.16}$$

$$M_{N-3}+4M_{N-2}+M_{N-1}=-\frac{\alpha}{4h^2}S_0-\frac{\alpha}{4h^2}S_N+\frac{7h^2}{60}S_{N-1}+\frac{16h^2}{60}S_{N-2}+\frac{7h^2}{60}S_{N-3}$$
$$+\frac{6\alpha}{h^2}\Phi^{(1)}(x_0)-\frac{6\alpha}{h^2}\Phi^{(1)}(x_N)+\frac{3\alpha}{h^2}M_0+\frac{3\alpha}{h^2}M_N$$
$$+\frac{6}{h^2}(U_{N-3}-2U_{N-2}+U_{N-1}), \tag{8.17}$$

$$M_{N-2}-2M_{N-1}+M_N=-\frac{\alpha}{2h^2}S_0-\frac{\alpha}{2h^2}S_N+\frac{h^2}{6}(S_{N-2}+4S_{N-1}+S_N)$$
$$-\frac{\alpha}{h^2}\Phi^{(3)}(x_0)+\frac{\alpha}{h^2}\Phi^{(3)}(x_N), \tag{8.18}$$

$$M_{N-3}-2M_{N-2}+M_{N-1}=-\frac{\alpha}{2h^2}S_0-\frac{\alpha}{2h^2}S_N+\frac{h^2}{6}(S_{N-3}+4S_{N-2}+S_{N-1})$$
$$-\frac{\alpha}{h^2}\Phi^{(3)}(x_0)+\frac{\alpha}{h^2}\Phi^{(3)}(x_N). \tag{8.19}$$

From (8.16) and (8.18), also from (8.17) and (8.19), we get

$$M_{N-1}=\frac{\alpha}{24h^2}(S_0+S_N)+\frac{\alpha}{h^2}(\Phi^{(1)}(x_0)-\Phi^{(1)}(x_N))+\frac{\alpha}{2h^2}(M_0+M_N)+\frac{\alpha}{6h^2}(\Phi^{(3)}(x_0)$$
$$-\Phi^{(3)}(x_N))+\frac{1}{h^2}(U_{N-2}-2U_{N-1}+U_N)-\frac{h^2}{120}(S_{N-2}+8S_{N-1}+S_N), \tag{8.20}$$

$$M_{N-2}=\frac{\alpha}{24h^2}(S_0+S_N)+\frac{\alpha}{h^2}(\Phi^{(1)}(x_0)-\Phi^{(1)}(x_N))+\frac{\alpha}{2h^2}(M_0+M_N)+\frac{\alpha}{6h^2}(\Phi^{(3)}(x_0)$$
$$-\Phi^{(3)}(x_N))+\frac{1}{h^2}(U_{N-3}-2U_{N-2}+U_{N-1})-\frac{h^2}{120}(S_{N-3}+8S_{N-2}+S_{N-1}). \tag{8.21}$$

Substitute (8.20) and (8.21) in (8.18), we get

$$-2U_N+5U_{N-1}-4U_{N-2}+U_{N-3}=-M_Nh^2-\frac{11\alpha}{24}(S_0+S_N)+\alpha(\Phi^{(1)}(x_0)-\Phi^{(1)}(x_N))$$

$$+\frac{\alpha}{2}(M_0+M_N)+\frac{5\alpha}{6}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0))$$
$$+\frac{h^4}{120}(18S_N+65S_{N-1}+26S_{N-2}+S_{N-3}). \quad (8.22)$$

Substituting $S_i = f_i - p_iU_i$, approximations of $\Phi^{(1)}(x_0)$, $\Phi^{(1)}(x_N)$, $\Phi^{(3)}(x_0)$, $\Phi^{(3)}(x_N)$ in (8.22) we get the difference equation as

$$\Big[-\frac{2\alpha}{h}+\frac{15\alpha}{2h^3}\Big]U_1+\Big[-\frac{10\alpha}{h^3}+\frac{\alpha}{2h}\Big]U_2+\frac{35\alpha}{6h^3}U_3-\frac{5\alpha}{4h^3}U_4-\frac{5\alpha}{4h^3}U_{N-4}$$
$$+\Big[1+\frac{h^4p_{N-3}}{120}+\frac{35\alpha}{6h^3}\Big]U_{N-3}+\Big[-4+\frac{13h^4p_{N-2}}{60}+\frac{\alpha}{2h}-\frac{10\alpha}{h^3}\Big]U_{N-2}$$
$$+\Big[5+\frac{13h^4p_{N-1}}{24}-\frac{2\alpha}{h}+\frac{15\alpha}{2h^3}\Big]U_{N-1}=-\hat{\eta}_1h^2-\frac{11\alpha}{24}(f_0+f_N)$$
$$+\frac{11\alpha}{24}(p_0\eta_0+p_N\eta_1)+\frac{\alpha}{2}(\hat{\eta}_0+\hat{\eta}_1)+\frac{h^4}{120}(18f_N+65f_{N-1}+26f_{N-2}$$
$$+f_{N-3})-\Big[-2+\frac{3h^4p_N}{20}+\frac{3\alpha}{2h}-\frac{25\alpha}{12h^3}\Big]\eta_1-\Big[-\frac{25\alpha}{12h^3}+\frac{3\alpha}{2h}\Big]\eta_0. \quad (8.23)$$

The system (8.15), (8.7) and (8.23) gives the approximate solution U_i, $i = 1, 2, \ldots, N-1$.

Remark 8.1.1. *If $\alpha = 0$, the system given in (8.15), (8.7) and (8.23) reduces into the system corresponding to the quintic spline.*

8.1.1 Truncation error

Let $|\alpha| < h^6$. The truncation error τ_1 is associated with the difference equation given in (8.15) is defined as

$$\tau_1=\Big[-2+\frac{3h^4p_0}{20}+\frac{3\alpha}{2h}-\frac{25\alpha}{12h^3}\Big]u(x_0)+\Big[5+\frac{13h^4p_1}{24}-\frac{2\alpha}{h}+\frac{15\alpha}{2h^3}\Big]u(x_1)$$
$$+\Big[-4+\frac{13h^4p_2}{60}+\frac{\alpha}{2h}-\frac{10\alpha}{h^3}\Big]u(x_2)+\Big[1+\frac{h^4p_3}{120}+\frac{35\alpha}{6h^3}\Big]u(x_3)$$
$$-\frac{5\alpha}{4h^3}u(x_4)-\frac{5\alpha}{4h^3}u(x_{N-4})+\frac{35\alpha}{6h^3}u(x_{N-3})+\Big[-\frac{10\alpha}{h^3}+\frac{\alpha}{2h}\Big]u(x_{N-2})$$
$$+\Big[-\frac{2\alpha}{h}+\frac{15\alpha}{2h^3}\Big]u(x_{N-1})+\Big[-\frac{25\alpha}{12h^3}+\frac{3\alpha}{2h}\Big]u(x_N)+h^2u''(x_0)$$
$$-\frac{\alpha}{2}(u''(x_0)+u''(x_N))+\frac{11\alpha}{24}(f_0+f_N)-\frac{11\alpha}{24}(p_0u(x_0)$$
$$+p_Nu(x_N))-\frac{h^4}{120}(18f_0+65f_1+26f_2+f_3).$$

After re-arranging the terms, we get

$$\tau_1=\frac{5\alpha}{6}\Big[\frac{-5u(x_0)+18u(x_1)-24u(x_2)+14u(x_3)-3u(x_4)}{2h^3}\Big]$$

$$
\begin{aligned}
&- \alpha\Big[\frac{-3u(x_0)+4u(x_1)-u(x_2)}{2h}\Big] + \alpha\Big[\frac{3u(x_N)-4u(x_{N-1})+u(x_{N-2})}{2h}\Big] \\
&- \frac{5\alpha}{6}\Big[\frac{5u(x_N)-18u(x_{N-1})+24u(x_{N-2})-14u(x_{N-3})+3u(x_{N-4})}{2h^3}\Big] \\
&- \frac{\alpha}{2}(u''(x_0)+u''(x_N)) + \frac{11\alpha}{24}(f_0+f_N) - \frac{11\alpha}{24}(p_0u(x_0) \\
&+ p_Nu(x_N)) + h^2u''(x_0) - 2u(x_0) + 5u(x_1) - 4u(x_2) + u(x_3) \\
&+ \frac{3h^4p_0}{20}u(x_0) + \frac{13h^4p_1}{21}u(x_1) + \frac{13h^4p_2}{60}u(x_2) + \frac{h^4p_3}{120}u(x_3) \\
&- \frac{h^4}{120}(18f_0+65f_1+26f_2+f_3).
\end{aligned}
$$

It can be seen that

$$
\begin{aligned}
\tau_1 = &\frac{5\alpha}{6}\Big[u'''(x_0)+\mathcal{O}(h^2)\Big] - \alpha\Big[u'(x_0)+\mathcal{O}(h^2)\Big] - \frac{5\alpha}{6}\Big[u'''(x_N)+\mathcal{O}(h^2)\Big] \\
&+ \alpha\Big[u'(x_N)+\mathcal{O}(h^2)\Big] - \frac{\alpha}{2}(u''(x_0)+u''(x_N)) + \frac{11\alpha}{24}(f_0+f_N) \\
&- \frac{11\alpha}{24}(p_0u(x_0)+p_Nu(x_N)) + h^2u''(x_0) - 2u(x_0) + 5u(x_1) - 4u(x_2) \\
&+ u(x_3) + \frac{3h^4p_0}{20}u(x_0) + \frac{13h^4p_1}{21}u(x_1) + \frac{13h^4p_2}{60}u(x_2) \\
&+ \frac{h^4p_3}{120}u(x_3) - \frac{h^4}{120}(18f_0+65f_1+26f_2+f_3).
\end{aligned}
\tag{8.24}
$$

Substitute $f_i = u^{(4)}(x_i) + p_iu(x_i)$ in (8.24), we obtain

$$
\begin{aligned}
\tau_1 = &\frac{5\alpha}{6}\Big[u'''(x_0)+\mathcal{O}(h^2)\Big] - \alpha\Big[u'(x_0)+\mathcal{O}(h^2)\Big] - \frac{5\alpha}{6}\Big[u'''(x_N)+\mathcal{O}(h^2)\Big] \\
&+ \alpha\Big[u'(x_N)+\mathcal{O}(h^2)\Big] - \frac{\alpha}{2}(u''(x_0)+u''(x_N)) + \frac{11\alpha}{24}(u^{(4)}(x_0)+u^{(4)}(x_N)) \\
&+ h^2u''(x_0) - 2u(x_0) + 5u(x_1) - 4u(x_2) + u(x_3) - \frac{h^4}{120}(18u^{(4)}(x_0) \\
&+ 65u^{(4)}(x_1) + 26u^{(4)}(x_2) + u^{(4)}(x_3)).
\end{aligned}
$$

Using the Taylor expansions about the point x_1, we get

$$
\begin{aligned}
\tau_1 = &\frac{5\alpha}{6}\Big[u'''(x_0)+\mathcal{O}(h^2)\Big] - \alpha\Big[u'(x_0)+\mathcal{O}(h^2)\Big] - \frac{5\alpha}{6}\Big[u'''(x_N)+\mathcal{O}(h^2)\Big] \\
&+ \alpha\Big[u'(x_N)+\mathcal{O}(h^2)\Big] - \frac{\alpha}{2}(u''(x_0)+u''(x_N)) + \frac{11\alpha}{24}(u^{(4)}(x_0)+u^{(4)}(x_N)) \\
&- \frac{7h^6}{90}u^{(6)}(x_1) + \mathcal{O}(h^7).
\end{aligned}
$$

Using the similar procedure followed to get τ_1, we obtain the truncation error τ_i, $i = 2, 3, \ldots, N-2$ associated with the difference equations given in (8.7) as

$$\tau_i = \alpha\Big[u'''(x_0) + \mathcal{O}(h^2)\Big] - \alpha\Big[u'''(x_N) + \mathcal{O}(h^2)\Big] + \frac{\alpha}{2}(u^{(4)}(x_0) + u^{(4)}(x_N)) \\ - \frac{h^6}{12}u^{(6)}(x_i) + \mathcal{O}(h^7),\ i = 2, 3, \ldots, N-2.$$

Similarly, we obtain the truncation error τ_{N-1} associated with the difference equation given in (8.23) as

$$\tau_{N-1} = \frac{5\alpha}{6}\Big[u'''(x_0) + \mathcal{O}(h^2)\Big] - \alpha\Big[u'(x_0) + \mathcal{O}(h^2)\Big] - \frac{5\alpha}{6}\Big[u'''(x_N) + \mathcal{O}(h^2)\Big] \\ + \alpha\Big[u'(x_N) + \mathcal{O}(h^2)\Big] - \frac{\alpha}{2}(u''(x_0) + u''(x_N)) + \frac{11\alpha}{24}(u^{(4)}(x_0) + u^{(4)}(x_N)) \\ - \frac{7h^6}{90}u^{(6)}(x_{N-1}) + \mathcal{O}(h^7).$$

Since $|\alpha| < h^6$, as $h \to 0$, we have $\tau_i = \mathcal{O}(h^6)$, $i = 1, 2, \ldots, N-1$.

Remark 8.1.2. *From the approximations U_i, $i = 0, 1, \ldots, N$, we get the approximations of S_i as $S_i = f_i - p_iU_i$, $i = 0, 1, \ldots, N$. Since $M_0 = \hat{\eta}_0$ and $M_N = \hat{\eta}_1$, we can get the approximations M_i, $i = 1, 2, \ldots, N-1$ from (8.5). With the help of approximations U_i, M_i and S_i, $i = 0, 1, \ldots, N$, we can calculate the coefficients $\mathcal{A}_i$, $\mathcal{B}_i$, $\mathcal{C}_i$, $\mathcal{D}_i$, $\mathcal{E}_i$, $\mathcal{F}_i$ of (7.2). Hence, we get the fractal quintic spline from the solutions of the differential equation (8.1). Hence, the information required in between mesh grids can get from (7.2).*

8.1.2 Numerical examples

In this section, two numerical examples are provided. For fixed h, by choosing α such that $|\alpha| < h^6$, we can compute the numerical solution. For any value of N, let

$$E_2^N = \max_i |u^{(2)}(x_i) - M_i|,\ E_4^N = \max_i |u^{(4)}(x_i) - S_i|, \\ p_2^N = E_2^N / E_2^{2N}, \quad p_4^N = E_4^N / E_4^{2N}.$$

Example 8.1.1. *Consider the fourth-order BVP [139]*

$$\begin{cases} u^{(4)}(x) + xu(x) = -(8 + 7x + x^3)\exp(x),\ x \in (0, 1), \\ u(0) = u(1) = 0, \quad u''(0) = 0,\ u''(1) = -4\exp(1). \end{cases}$$

The exact solution is $u(x) = x(1-x)\exp(x)$. The scaling factor $\alpha = 0.9999h^6$ is used to compute the numerical solution. The maximum point-wise error and the order of convergence are given in Table 8.1. Table 8.2 represents the maximum point-wise error and the order of convergence corresponding to quintic spline.

Table 8.1: Maximum point-wise error and order of convergence corresponding to Example 8.1.1.

N	16	32	64	128
E^N	1.7217e-05	2.0166e-06	3.5914e-07	8.0523e-08
p^N	3.0939	2.4893	2.1571	
E_2^N	2.8241e-04	5.1557e-05	1.1630e-05	2.9860e-06
p_2^N	2.4536	2.1483	1.9616	
E_4^N	7.6085e-06	7.1011e-07	1.8758e-07	5.2749e-08
p_4^N	3.4215	1.9205	1.8303	

Table 8.2: Maximum point-wise error and order of convergence corresponding to Example 8.1.1 with $\alpha = 0$.

N	16	32	64	128
E^N	2.1621e-04	5.4012e-05	1.3512e-05	3.3778e-06
p^N	2.0011	1.9991	2.0001	
E_2^N	2.0918e-03	5.2281e-04	1.3069e-04	3.2682e-05
p_2^N	2.0004	2.0001	1.9996	
E_4^N	1.2715e-04	3.1993e-05	7.9966e-06	1.9991e-06
p_4^N	1.9907	2.0003	2.0001	

Example 8.1.2. *Consider the fourth-order BVP*

$$\begin{cases} u^{(4)}(x) - xu(x) = -(11 + 9x + x^2 - x^3)\exp(x), \ x \in (0, 1), \\ u(0) = 1, \ u(1) = 0, \quad u''(0) = -1, \ u''(1) = -6\exp(1). \end{cases}$$

The exact solution of the BVP is $u(x) = (1 - x^2)\exp(x)$. The scaling factor $\alpha = 0.9999h^6$ is used to compute the numerical solution. The maximum point-wise error and the order of convergence are given in Table 8.3. The maximum point-wise error and the order of convergence corresponding to the quintic spline are given in Table 8.4.

Table 8.3: Maximum point-wise error and order of convergence corresponding to Example 8.1.2.

N	16	32	64	128
E^N	2.0391e-05	2.3565e-06	4.1576e-07	9.2846e-08
p^N	3.1132	2.5028	2.1628	
E_2^N	3.3020e-04	5.9829e-05	1.3419e-05	3.4487e-06
p_2^N	2.4644	2.1566	1.9601	
E_4^N	9.0609e-06	8.3312e-07	2.1611e-07	6.1050e-08
p_4^N	3.4431	1.9467	1.8237	

Table 8.4: Maximum point-wise error and order of convergence corresponding to Example 8.1.2 with $\alpha = 0$.

N	16	32	64	128
E^N	2.5837e-04	6.4547e-05	1.6147e-05	4.0365e-06
p^N	2.0010	1.9991	2.0001	
E_2^N	2.4995e-03	6.2471e-04	1.5617e-04	3.9065e-05
p_2^N	2.0004	2.0001	1.9992	
E_4^N	1.5187e-04	3.8215e-05	9.5519e-06	2.3878e-06
p_4^N	1.9906	2.0003	2.0001	

8.2 Numerical Scheme

Let us consider the BVP given in (8.2)

$$\begin{cases} u^{(4)}(x) + p(x)u(x) = f(x), \; x \in (0,1), \\ u(0) = \eta_0, \; u(1) = \eta_1, \; u'(0) = \hat{\eta}_0, \; u'(1) = \hat{\eta}_1. \end{cases}$$

Let $0 = x_0 < x_1 < \cdots < x_N = 1$ be the uniform partition of the interval $I = [0,1]$. Let $u(x)$ be the solution of the differential equation given in (8.2). Let U_i, U_i' and S_i be the approximations to $u(x_i)$, $u^{(1)}(x_i)$ and $u^{(4)}(x_i)$ respectively. Consider the IFS the form $\Big\{I \times \mathbb{R}; w_i(x,y) = (L_i(x),\; F_i(x,y)) : i = 1,2,\dots,N\Big\}$, where $L_i : I \to I_i = [x_{i-1}, x_i]$ such that $L_i(x) = hx + x_{i-1}, \; x \in I$, $F_i : I \times \mathbb{R} \to \mathbb{R}$ such that $F_i(x,y) = \alpha y + r_i(x)$, $(x,y) \in I \times \mathbb{R}$, with $r_i(x) = \mathcal{A}_i(x-x_0)^5 + \mathcal{B}_i(x-x_0)^4 + \mathcal{C}_i(x-x_0)^3 + \mathcal{D}_i(x-x_0)^2 + \mathcal{E}_i(x-x_0) + \mathcal{F}_i$, α is the scaling factor such that $|\alpha| < h^4$.

We assume the following conditions on the IFS:

$$F_i(x_0, U_0) = U_{i-1}, \; F_i(x_N, U_N) = U_i, \quad i = 1,2,\dots,N,$$

$$F_{i,1}(x_0, U_0') = U_{i-1}', \; F_{i,1}(x_0, U_N') = U_i', \; i = 1,2,\dots,N,$$

$$F_{i,2}(x_N, U_{N,2}) = F_{i+1,2}(x_0, U_{0,2}), \; i = 1,2,\dots,N-1,$$

$$F_{i,3}(x_N, U_{N,3}) = F_{i+1,3}(x_0, U_{0,3}), \; i = 1,2,\dots,N-1,$$

$$F_{i,4}(x_0, S_0) = S_{i-1}, \; F_{i,4}(x_N, S_N) = S_i, \quad i = 1,2,\dots,N,$$

where $F_{i,k}(x,y) = \frac{\alpha y + r_i^{(k)}(x)}{h^k}$, $k = 1,2,3,4$ and $U_{0,2} = \frac{r_1^{(2)}(x_0)}{h^2-\alpha}$, $U_{N,2} = \frac{r_N^{(2)}(x_N)}{h^2-\alpha}$, $U_{0,3} = \frac{r_1^{(3)}(x_0)}{h^3-\alpha}$, $U_{N,3} = \frac{r_N^{(3)}(x_N)}{h^3-\alpha}$.

Let $\mathscr{F} = \{\varphi \in \mathcal{C}^4(I,\mathbb{R}) \mid \varphi(x_0) = U_0, \varphi(x_N) = U_N, \varphi^{(1)}(x_0) = U_0', \varphi^{(1)}(x_N) = U_N', \varphi^{(4)}(x_0) = S_0, \varphi^{(4)}(x_N) = S_N\}$ be endowed with the metric ρ induced by the $\mathcal{C}^4$-norm. $(\mathscr{F}, \rho)$ is a complete metric space. Define the Read-Bajraktarević operator T on $(\mathscr{F}, \rho)$ as

$$T\varphi(L_i(x)) = \alpha\varphi(x) + \mathcal{A}_i(x-x_0)^5 + \mathcal{B}_i(x-x_0)^4 + \mathcal{C}_i(x-x_0)^3 + \mathcal{D}_i(x-x_0)^2 + \mathcal{E}_i(x-x_0) + \mathcal{F}_i, \quad x \in [x_0, x_N], \; i = 1,2,\dots,N,$$

where $|\alpha| < h^4$. The constants $\mathcal{A}_i$, $\mathcal{B}_i$, $\mathcal{C}_i$, $\mathcal{D}_i$, $\mathcal{E}_i$ and $\mathcal{F}_i$ are to be evaluated. The operator T is contraction map. So, it has a unique fixed-point Φ (say). The fixed-point Φ satisfies the following functional equation:

$$\Phi(L_i(x)) = \alpha\Phi(x) + \mathcal{A}_i(x-x_0)^5 + \mathcal{B}_i(x-x_0)^4 + \mathcal{C}_i(x-x_0)^3 + \mathcal{D}_i(x-x_0)^2 + \mathcal{E}_i(x-x_0) + \mathcal{F}_i, \quad x \in [x_0, x_N],\ i = 1,2,\ldots,N. \tag{8.25}$$

The conditions $F_i(x_0, U_0) = U_{i-1},\ F_i(x_N, U_N) = U_i,\ F_{i,1}(x_0, U_0') = U_{i-1}',\ F_{i,1}(x_N, U_N') = U_i',\ F_{i,4}(x_0, S_0) = S_{i-1},\ F_{i,4}(x_N, S_N) = S_i$ are equivalent to

$$\begin{cases} \Phi(x_{i-1}) = U_{i-1},\ \Phi(x_i) = U_i,\ \Phi^{(1)}(x_{i-1}) = U_{i-1}',\ \Phi^{(1)}(x_i) = U_i', \\ \Phi^{(4)}(x_{i-1}) = S_{i-1},\ \Phi^{(4)}(x_i) = S_i, \end{cases} \tag{8.26}$$

$F_{i,2}(x_N, U_{N,2}) = F_{i+1,2}(x_0, U_{0,2})$ can be reformulated as $\Phi^{(2)}(L_i(x_N)) = \Phi^{(2)}(L_{i+1}(x_0))$, $F_{i,3}(x_N, U_{N,3}) = F_{i+1,3}(x_0, U_{0,3})$ can be reformulated as $\Phi^{(3)}(L_i(x_N)) = \Phi^{(3)}(L_{i+1}(x_0))$. For fractal quintic spline Φ, the constants $\mathcal{A}_i$, $\mathcal{B}_i$, $\mathcal{C}_i$, $\mathcal{D}_i$, $\mathcal{E}_i$ and $\mathcal{F}_i$ are evaluated using the conditions (8.26) and hence we get

$$\begin{aligned}
\mathcal{A}_i &= \frac{h^4}{120}\Big[\Big(S_i - \frac{\alpha}{h^4}S_N\Big) - \Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big)\Big],\\
\mathcal{B}_i &= \frac{h^4}{24}\Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big),\\
\mathcal{C}_i &= 2\Big[\Big(U_{i-1} - \alpha U_0\Big) - \Big(U_i - \alpha U_N\Big)\Big] + h\Big[\Big(U_{i-1}' - \frac{\alpha}{h}U_0'\Big) + \Big(U_i' - \frac{\alpha}{h}U_N'\Big)\Big]\\
&\quad - \frac{7h^4}{120}\Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big) - \frac{3h^4}{120}\Big(S_i - \frac{\alpha}{h^4}S_N\Big),\\
\mathcal{D}_i &= 3\Big[\Big(U_i - \alpha U_N\Big) - \Big(U_{i-1} - \alpha U_0\Big)\Big] - h\Big[2\Big(U_{i-1}' - \frac{\alpha}{h}U_0'\Big) + \Big(U_i' - \frac{\alpha}{h}U_N'\Big)\Big]\\
&\quad + \frac{3h^4}{120}\Big(S_{i-1} - \frac{\alpha}{h^4}S_0\Big) + \frac{2h^4}{120}\Big(S_i - \frac{\alpha}{h^4}S_N\Big),\\
\mathcal{E}_i &= h\Big(U_{i-1}' - \frac{\alpha}{h}U_0'\Big),\\
\mathcal{F}_i &= U_{i-1} - \alpha U_0.
\end{aligned}$$

For continuity of $\Phi^{(2)}$, at interior nodes x_i, $i = 1,2,\ldots,N-1$, we need $\Phi^{(2)}(x_i^-) = \Phi^{(2)}(x_i^+)$ i.e., $\Phi^{(2)}(L_i(x_N)) = \Phi^{(2)}(L_{i+1}(x_0))$ and which leads to the following:

$$\alpha\Phi^{(2)}(x_N) + 20\mathcal{A}_i + 12\mathcal{B}_i + 6\mathcal{C}_i + 2\mathcal{D}_i = \alpha\Phi^{(2)}(x_0) + 2\mathcal{D}_{i+1}.$$

Substituting the values of $\mathcal{A}_i$, $\mathcal{B}_i$, $\mathcal{C}_i$, $\mathcal{D}_i$ and $\mathcal{D}_{i+1}$, we get

$$-\frac{3\alpha}{h}U_0' + U_{i-1}' + 4U_i' + U_{i+1}' - \frac{3\alpha}{h}U_N' = \frac{h^3}{60}(S_{i+1} - S_{i-1}) + \frac{\alpha}{120h}(S_N - S_0)$$

$$+\frac{3}{h}(U_{i+1}-U_{i-1})+\frac{6\alpha}{h}(U_0-U_N)+\frac{\alpha}{2h}(\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N)). \quad (8.27)$$

At interior nodes $x_i,\ i=1,2,\dots,N-1$, for continuity of $\Phi^{(3)}$, we need $\Phi^{(3)}(x_i^-) = \Phi^{(3)}(x_i^+)$ i.e., $\Phi^{(3)}(L_i(x_N)) = \Phi^{(3)}(L_{i+1}(x_0))$ and which leads to the following:

$$\begin{aligned} U'_{i+1}-U'_{i-1} &= \frac{h^3}{120}(3S_{i-1}+14S_i+3S_{i+1})+\frac{2}{h}(U_{i+1}-2U_i+U_{i-1}) \\ &-\frac{10\alpha}{120h}(S_0+S_N)+\frac{\alpha}{6h}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0)). \end{aligned} \quad (8.28)$$

In order to get the relation between U_i's and S_i's, we use the following procedure: The continuity condition of $\Phi^{(2)}(x)$ at $x=x_{i-1}$ and $x=x_{i+1}$ are

$$\begin{aligned} &-\frac{3\alpha}{h}U'_0+U'_{i-2}+4U'_{i-1}+U'_i-\frac{3\alpha}{h}U'_N = \frac{h^3}{60}(S_i-S_{i-2})+\frac{\alpha}{120h}(S_N-S_0) \\ &+\frac{3}{h}(U_i-U_{i-2})+\frac{6\alpha}{h}(U_0-U_N)+\frac{\alpha}{2h}(\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N)) \end{aligned} \quad (8.29)$$

and

$$\begin{aligned} &-\frac{3\alpha}{h}U'_0+U'_i+4U'_{i+1}+U'_{i+2}-\frac{3\alpha}{h}U'_N = \frac{h^3}{60}(S_{i+2}-S_i)+\frac{\alpha}{120h}(S_N-S_0) \\ &+\frac{3}{h}(U_{i+2}-U_i)+\frac{6\alpha}{h}(U_0-U_N)+\frac{\alpha}{2h}(\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N)) \end{aligned} \quad (8.30)$$

respectively. Also, the continuity condition of $\Phi^{(3)}(x)$ at $x=x_{i-1}$ and $x=x_{i+1}$ are

$$\begin{aligned} U'_i-U'_{i-2} &= \frac{h^3}{120}(3S_{i-2}+14S_{i-1}+3S_i)+\frac{2}{h}(U_i-2U_{i-1}+U_{i-2}) \\ &-\frac{10\alpha}{120h}(S_0+S_N)+\frac{\alpha}{6h}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0)) \end{aligned} \quad (8.31)$$

and

$$\begin{aligned} U'_{i+2}-U'_i &= \frac{h^3}{120}(3S_i+14S_{i+1}+3S_{i+2})+\frac{2}{h}(U_{i+2}-2U_{i+1}+U_i) \\ &-\frac{10\alpha}{120h}(S_0+S_N)+\frac{\alpha}{6h}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0)). \end{aligned} \quad (8.32)$$

respectively. Our aim is to find the scalar combination $s_1 \times (8.29) + s_2 \times (8.27) + s_3 \times (8.30) + s_4 \times (8.31) + s_5 \times (8.28) + s_6 \times (8.32)$ which gives the relation between U_i's and S_i's where $s_i,\ i=1,2,\dots,6$ are the scalars. If $s_1=1,\ s_2=0,\ s_3=-1,\ s_4=1,\ s_5=4,\ s_6=1$ then the corresponding scalar combination yields

$$U_{i-2}-4U_{i-1}+6U_i-4U_{i+1}+U_{i+2} = -\frac{\alpha}{2}(S_0+S_N)-\alpha\Phi^{(3)}(x_0)+\alpha\Phi^{(3)}(x_N)$$

$$+\frac{h^4}{120}(S_{i+2}+26S_{i+1}+66S_i+26S_{i-1}+S_{i-2}). \tag{8.33}$$

Notice that (8.33) gives the relation between U_i's and S_i's. The differential equation (8.2) is discretized as $S_i + p_iU_i = f_i$, where $p_i = p(x_i)$, $f_i = f(x_i)$. The boundary conditions are discretized as $U_0 = \eta_0$, $U_N = \eta_1$, $U_0' = \hat{\eta}_0$, $U_N' = \hat{\eta}_1$.

Substituting $S_i = f_i - p_iU_i$, approximations for $\Phi^{(3)}(x_0)$ and $\Phi^{(3)}(x_N)$ in (8.33), and get

$$\begin{cases} \frac{9\alpha}{h^3}U_1 - \frac{12\alpha}{h^3}U_2 + \frac{7\alpha}{h^3}U_3 - \frac{3\alpha}{2h^3}U_4 + \Big[1+\frac{h^4p_{i-2}}{120}\Big]U_{i-2} + \Big[-4 \\ +\frac{13h^4p_{i-1}}{60}\Big]U_{i-1} + \Big[6+\frac{11h^4p_i}{20}\Big]U_i + \Big[-4+\frac{13h^4p_{i+1}}{60}\Big]U_{i+1} \\ +\Big[1+\frac{h^4p_{i+2}}{120}\Big]U_{i+2} - \frac{3\alpha}{2h^3}U_{N-4} + \frac{7\alpha}{h^3}U_{N-3} - \frac{12\alpha}{h^3}U_{N-2} + \frac{9\alpha}{h^3}U_{N-1} \\ = \frac{h^4}{120}\Big[f_{i-2}+26f_{i-1}+66f_i+26f_{i+1}+f_{i+2}\Big] - \frac{\alpha}{2}(f_0+f_N) \\ +\frac{\alpha}{2}(p_0\eta_0+p_N\eta_N) + \frac{5\alpha}{2h^3}\eta_0 + \frac{5\alpha}{2h^3}\eta_N,\ i=2,3,\dots,N-2. \end{cases} \tag{8.34}$$

Difference Equations: Put $i = 1, 2$ in (8.27) and (8.28), we get

$$\begin{aligned} -\frac{3\alpha}{h}U_0' + U_0' + 4U_1' + U_2' - \frac{3\alpha}{h}U_N' &= \frac{h^3}{60}(S_2-S_0) + \frac{\alpha}{120h}(S_N-S_0) \\ +\frac{3}{h}(U_2-U_0) + \frac{6\alpha}{h}(U_0-U_N) &+ \frac{\alpha}{2h}(\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N)), \end{aligned} \tag{8.35}$$

$$\begin{aligned} -\frac{3\alpha}{h}U_0' + U_1' + 4U_2' + U_3' - \frac{3\alpha}{h}U_N' &= \frac{h^3}{60}(S_3-S_1) + \frac{\alpha}{120h}(S_N-S_0) \\ +\frac{3}{h}(U_3-U_1) + \frac{6\alpha}{h}(U_0-U_N) &+ \frac{\alpha}{2h}(\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N)), \end{aligned} \tag{8.36}$$

$$\begin{aligned} U_2' - U_0' &= \frac{h^3}{120}(3S_0+14S_1+3S_2) + \frac{2}{h}(U_2-2U_1+U_0) \\ &- \frac{10\alpha}{120h}(S_0+S_N) + \frac{\alpha}{6h}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0)), \end{aligned} \tag{8.37}$$

$$\begin{aligned} U_3' - U_1' &= \frac{h^3}{120}(3S_1+14S_2+3S_3) + \frac{2}{h}(U_3-2U_2+U_1) \\ &- \frac{10\alpha}{120h}(S_0+S_N) + \frac{\alpha}{6h}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0)). \end{aligned} \tag{8.38}$$

The scalar combination $(8.35) - 2\times(8.36) + 7\times(8.37) + 2\times(8.38)$ yields the equation as

$$-11U_0 + 18U_1 - 9U_2 + 2U_3 = 6hU_0' + \frac{h^4}{120}[19S_0+108S_1+51S_2+2S_3]$$

$$
\begin{aligned}
&-3\alpha(U_0'+U_N')-\frac{\alpha}{2}(\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N))+\frac{3\alpha}{2}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0))\\
&-6\alpha(U_0-U_N)-\frac{89\alpha}{120}S_0-\frac{91\alpha}{120}S_N.
\end{aligned}
\tag{8.39}
$$

By substituting $\Phi^{(2)}(x_0)=\frac{2U_0-5U_1+4U_2-U_3}{h^2}$, $\Phi^{(2)}(x_N)=\frac{2U_N-5U_{N-1}+4U_{N-2}-U_{N-3}}{h^2}$, $S_i=f_i-p_iU_i$, approximations for $\Phi^{(3)}(x_0)$, $\Phi^{(3)}(x_N)$ in (8.39), we get the difference equation as

$$
\begin{aligned}
&\Big[18-\frac{5\alpha}{2h^2}+\frac{27\alpha}{2h^3}+\frac{9h^4p_1}{10}\Big]U_1+\Big[-9+\frac{2\alpha}{h^2}-\frac{18\alpha}{h^3}+\frac{17h^4p_2}{40}\Big]U_2+\Big[2-\frac{\alpha}{2h^2}\\
&+\frac{21\alpha}{2h^3}+\frac{h^4p_3}{60}\Big]U_3-\Big[\frac{9\alpha}{4h^3}\Big]U_4-\Big[\frac{9\alpha}{4h^3}\Big]U_{N-4}+\Big[\frac{\alpha}{2h^2}+\frac{21\alpha}{2h^3}\Big]U_{N-3}+\Big[-\frac{2\alpha}{h^2}\\
&-\frac{18\alpha}{h^3}\Big]U_{N-2}+\Big[\frac{5\alpha}{2h^2}+\frac{27\alpha}{2h^3}\Big]U_{N-1}=6h\hat{\eta}_0+\frac{19h^4}{120}f_0+\frac{9h^4}{10}f_1+\frac{17h^4}{40}f_2+\frac{h^4}{60}f_3\\
&-3\alpha(\hat{\eta}_0+\hat{\eta}_1)-\frac{89\alpha}{120}f_0-\frac{91\alpha}{120}f_N-\Big[-11+\frac{\alpha}{h^2}-\frac{15\alpha}{4h^3}+6\alpha+\frac{19h^4}{120}p_0\\
&-\frac{89\alpha}{120}p_0\Big]\eta_0-\Big[-\frac{\alpha}{h^2}-\frac{15\alpha}{4h^3}-6\alpha-\frac{91\alpha}{120}p_N\Big]\eta_1.
\end{aligned}
\tag{8.40}
$$

Put $i=N-2,N-1$ in (8.27) and (8.28), and get

$$
\left\{
\begin{aligned}
&-\frac{3\alpha}{h}U_0'+U_{N-3}'+4U_{N-2}'+U_{N-1}'-\frac{3\alpha}{h}U_N'=\frac{h^3}{60}(S_{N-1}-S_{N-3})\\
&+\frac{\alpha}{120h}(S_N-S_0)+\frac{3}{h}(U_{N-1}-U_{N-3})+\frac{6\alpha}{h}(U_0-U_N)\\
&+\frac{\alpha}{2h}(\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N)),
\end{aligned}
\right.
\tag{8.41}
$$

$$
\begin{aligned}
&-\frac{3\alpha}{h}U_0'+U_{N-2}'+4U_{N-1}'+U_N'-\frac{3\alpha}{h}U_N'=\frac{h^3}{60}(S_N-S_{N-2})+\frac{\alpha}{120h}(S_N-S_0)\\
&+\frac{3}{h}(U_N-U_{N-2})+\frac{6\alpha}{h}(U_0-U_N)+\frac{\alpha}{2h}(\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N)),
\end{aligned}
\tag{8.42}
$$

$$
\begin{aligned}
&U_{N-1}'-U_{N-3}'=\frac{h^3}{120}(3S_{N-3}+14S_{N-2}+3S_{N-1})+\frac{2}{h}(U_{N-1}-2U_{N-2}+U_{N-3})\\
&-\frac{10\alpha}{120h}(S_0+S_N)+\frac{\alpha}{6h}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0)),
\end{aligned}
\tag{8.43}
$$

$$
\begin{aligned}
&U_N'-U_{N-2}'=\frac{h^3}{120}(3S_{N-2}+14S_{N-1}+3S_N)+\frac{2}{h}(U_N-2U_{N-1}+U_{N-2})\\
&-\frac{10\alpha}{120h}(S_0+S_N)+\frac{\alpha}{6h}(\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0)).
\end{aligned}
\tag{8.44}
$$

The scalar combination $2 \times (8.41) - (8.42) + 2 \times (8.43) + 7 \times (8.44)$ gives the equation as

$$2U_{N-3} - 9U_{N-2} + 18U_{N-1} - 11U_N = -6hU'_N + \frac{h^4}{120}[2S_{N-3} + 51S_{N-2} + 108S_{N-1} + 19S_N] + 3\alpha(U'_0 + U'_N) + \frac{\alpha}{2}(\Phi^{(2)}(x_0) - \Phi^{(2)}(x_N)) - 6\alpha(U_N - U_0) + \frac{3\alpha}{2}(\Phi^{(3)}(x_N) - \Phi^{(3)}(x_0)) - \frac{91\alpha}{120}S_0 - \frac{89\alpha}{120}S_N. \quad (8.45)$$

By substituting $S_i = f_i - p_iU_i$ and approximations of $\Phi^{(2)}(x_0)$, $\Phi^{(2)}(x_N)$, $\Phi^{(3)}(x_0)$, $\Phi^{(3)}(x_N)$ in (8.45), we get

$$\begin{cases} \Big[\frac{5\alpha}{2h^2} + \frac{27\alpha}{2h^3}\Big]U_1 + \Big[-\frac{2\alpha}{h^2} - \frac{18\alpha}{h^3}\Big]U_2 + \Big[\frac{\alpha}{2h^2} + \frac{21\alpha}{2h^3}\Big]U_3 - \frac{9\alpha}{4h^3}U_4 \\ - \frac{9\alpha}{4h^3}U_{N-4} + \Big[2 - \frac{\alpha}{2h^2} + \frac{21\alpha}{2h^3} + \frac{h^4}{60}p_{N-3}\Big]U_{N-3} + \Big[-9 + \frac{2\alpha}{h^2} - \frac{18\alpha}{h^3} \\ + \frac{17h^4}{40}p_{N-2}\Big]U_{N-2} + \Big[18 - \frac{5\alpha}{2h^2} + \frac{27\alpha}{2h^3} + \frac{9h^4}{10}p_{N-1}\Big]U_{N-1} = -6h\hat{\eta}_1 \\ + \frac{h^4}{60}f_{N-3} + \frac{17h^4}{40}f_{N-2} + \frac{9h^4}{10}f_{N-1} + \frac{19h^4}{120}f_N + 3\alpha(\hat{\eta}_0 + \hat{\eta}_1) - \frac{91\alpha}{120}f_0 \\ - \frac{89\alpha}{120}f_N - \Big[-\frac{\alpha}{h^2} - \frac{15\alpha}{4h^3} - 6\alpha - \frac{91\alpha}{120}p_0\Big]\eta_0 - \Big[-11 + \frac{\alpha}{h^2} \\ - \frac{15\alpha}{4h^3} + \frac{19h^4}{120}p_N + 6\alpha - \frac{89\alpha}{120}p_N\Big]\eta_1. \end{cases} \quad (8.46)$$

The system (8.40), (8.34) and (8.46) provides the approximations U_i, $i = 1, 2, \ldots, N-1$.

8.2.1 Truncation error

Let $|\alpha| < h^6$. The truncation error corresponding to the equation (8.40) is defined as

$$\tau_1 = \Big[-11 + \frac{\alpha}{h^2} - \frac{15\alpha}{4h^3} + 6\alpha + \frac{19h^4}{120}p_0 - \frac{89\alpha}{120}p_0\Big]u(x_0) + \Big[18 - \frac{5\alpha}{2h^2} + \frac{27\alpha}{2h^3} + \frac{9h^4p_1}{10}\Big]u(x_1) + \Big[-9 + \frac{2\alpha}{h^2} - \frac{18\alpha}{h^3} + \frac{17h^4p_2}{40}\Big]u(x_2) + \Big[2 - \frac{\alpha}{2h^2} + \frac{21\alpha}{2h^3} + \frac{h^4p_3}{60}\Big]u(x_3) - \Big[\frac{9\alpha}{4h^3}\Big]u(x_4) - \Big[\frac{9\alpha}{4h^3}\Big]u(x_{N-4}) + \Big[\frac{\alpha}{2h^2} + \frac{21\alpha}{2h^3}\Big]u(x_{N-3}) + \Big[-\frac{2\alpha}{h^2} - \frac{18\alpha}{h^3}\Big]u(x_{N-2}) + \Big[\frac{5\alpha}{2h^2} + \frac{27\alpha}{2h^3}\Big]u(x_{N-1}) + \Big[-\frac{\alpha}{h^2} - \frac{15\alpha}{4h^3} - 6\alpha - \frac{91\alpha}{120}p_N\Big]u(x_N) - 6hu'(x_0) - \frac{19h^4}{120}f_0 - \frac{9h^4}{10}f_1 - \frac{17h^4}{40}f_2 - \frac{h^4}{60}f_3 + 3\alpha(u'(x_0) + u'(x_N)) + \frac{89\alpha}{120}f_0 + \frac{91\alpha}{120}f_N.$$

After simplification, we obtain

$$\tau_1 = \frac{3\alpha}{2}\Big[\frac{-5u(x_0)+18u(x_1)-24u(x_2)+14u(x_3)-3u(x_4)}{2h^3}\Big]$$
$$-\frac{3\alpha}{2}\Big[\frac{5u(x_N)-18u(x_{N-1})+24u(x_{N-2})-14u(x_{N-3})+3u(x_{N-4})}{2h^3}\Big]$$
$$+\frac{\alpha}{2}\Big[\frac{2u(x_0)-5u(x_1)+4u(x_2)-u(x_3)}{h^2}\Big]+3\alpha(u'(x_0)+u'(x_N))$$
$$-\frac{\alpha}{2}\Big[\frac{2u(x_N)-5u(x_{N-1})+4u(x_{N-2})-u(x_{N-3})}{h^2}\Big]+6\alpha(u(x_0)-u(x_N))$$
$$-\frac{89\alpha}{120}p_0u(x_0)-\frac{91\alpha}{120}p_Nu(x_N)+\frac{89\alpha}{120}f_0+\frac{91\alpha}{120}f_N$$
$$+\Big[-11+\frac{19h^4}{120}p_0\Big]u(x_0)+\Big[18+\frac{9h^4p_1}{10}\Big]u(x_1)+\Big[-9$$
$$+\frac{17h^4p_2}{40}\Big]u(x_2)+\Big[2+\frac{h^4p_3}{60}\Big]u(x_3)-6hu'(x_0)-\frac{19h^4}{120}f_0$$
$$-\frac{9h^4}{10}f_1-\frac{17h^4}{40}f_2-\frac{h^4}{60}f_3.$$

Further, τ_1 can be written as

$$\tau_1 = \frac{3\alpha}{2}\Big[u^{(3)}(x_0)+\mathcal{O}(h^2)\Big]-\frac{3\alpha}{2}\Big[u^{(3)}(x_N)+\mathcal{O}(h^2)\Big]+\frac{\alpha}{2}\Big[u^{(2)}(x_0)+\mathcal{O}(h^2)\Big]$$
$$+3\alpha(u'(x_0)+u'(x_N))-\frac{\alpha}{2}\Big[u^{(2)}(x_N)+\mathcal{O}(h^2)\Big]+6\alpha(u(x_0)-u(x_N))$$
$$-\frac{89\alpha}{120}p_0u(x_0)-\frac{91\alpha}{120}p_Nu(x_N)+\frac{89\alpha}{120}f_0+\frac{91\alpha}{120}f_N$$
$$+\Big[-11+\frac{19h^4}{120}p_0\Big]u(x_0)+\Big[18+\frac{9h^4p_1}{10}\Big]u(x_1)+\Big[-9$$
$$+\frac{17h^4p_2}{40}\Big]u(x_2)+\Big[2+\frac{h^4p_3}{60}\Big]u(x_3)-6hu'(x_0)-\frac{19h^4}{120}f_0$$
$$-\frac{9h^4}{10}f_1-\frac{17h^4}{40}f_2-\frac{h^4}{60}f_3.$$

After substituting $f_i = u^{(4)}(x_i)+p_iu(x_i)$ in τ_1, we get

$$\tau_1 = \frac{3\alpha}{2}\Big[u^{(3)}(x_0)+\mathcal{O}(h^2)\Big]-\frac{3\alpha}{2}\Big[u^{(3)}(x_N)+\mathcal{O}(h^2)\Big]+\frac{\alpha}{2}\Big[u^{(2)}(x_0)+\mathcal{O}(h^2)\Big]$$
$$+3\alpha(u'(x_0)+u'(x_N))-\frac{\alpha}{2}\Big[u^{(2)}(x_N)+\mathcal{O}(h^2)\Big]+6\alpha(u(x_0)-u(x_N))$$
$$+\frac{89\alpha}{120}u^{(4)}(x_0)+\frac{91\alpha}{120}u^{(4)}(x_N)-11u(x_0)+18u(x_1)-9u(x_2)+2u(x_3)$$
$$-6hu'(x_0)-\frac{19h^4}{120}u^{(4)}(x_0)-\frac{9h^4}{10}u^{(4)}(x_1)-\frac{17h^4}{40}u^{(4)}(x_2)-\frac{h^4}{60}u^{(4)}(x_3). \tag{8.47}$$

Using the Taylor expansions for $u(x_0)$, $u(x_2)$, $u(x_3)$, $u'(x_0)$,$u^{(4)}(x_0)$, $u^{(4)}(x_2)$, $u^{(4)}(x_3)$ about the point x_1 in (8.47), we get

$$\begin{aligned}\tau_1 =& \frac{3\alpha}{2}\Big[u^{(3)}(x_0)+\mathcal{O}(h^2)\Big]-\frac{3\alpha}{2}\Big[u^{(3)}(x_N)+\mathcal{O}(h^2)\Big]+\frac{\alpha}{2}\Big[u^{(2)}(x_0)+\mathcal{O}(h^2)\Big]\\ &+3\alpha(u'(x_0)+u'(x_N))-\frac{\alpha}{2}\Big[u^{(2)}(x_N)+\mathcal{O}(h^2)\Big]+6\alpha(u(x_0)-u(x_N))\\ &+\frac{89\alpha}{120}u^{(4)}(x_0)+\frac{91\alpha}{120}u^{(4)}(x_N)-\frac{1}{8}u^{(6)}(x_1)+\mathcal{O}(h^7).\end{aligned}$$

Using the procedure followed to get τ_1, we get the truncation errors τ_i, $i = 2, 3, \ldots, N-2$ corresponding to the equations given in (8.34) as

$$\begin{aligned}\tau_i =& \alpha\Big[u'''(x_0)+\mathcal{O}(h^2)\Big]-\alpha\Big[u'''(x_N)+\mathcal{O}(h^2)\Big]+\frac{\alpha}{2}(u^{(4)}(x_0)+u^{(4)}(x_N))\\ &-\frac{h^6}{12}u^{(6)}(x_i)+\mathcal{O}(h^7),\ i=2,3,\ldots,N-2,\end{aligned}$$

and truncation error τ_{N-1} corresponding to the equations given in (8.46) as

$$\begin{aligned}\tau_{N-1} =& \frac{3\alpha}{2}\Big[u^{(3)}(x_0)+\mathcal{O}(h^2)\Big]-\frac{3\alpha}{2}\Big[u^{(3)}(x_N)+\mathcal{O}(h^2)\Big]-\frac{\alpha}{2}\Big[u^{(2)}(x_0)+\mathcal{O}(h^2)\Big]\\ &-3\alpha(u'(x_0)+u'(x_N))+\frac{\alpha}{2}\Big[u^{(2)}(x_N)+\mathcal{O}(h^2)\Big]-6\alpha(u(x_0)-u(x_N))\\ &+\frac{91\alpha}{120}u^{(4)}(x_0)+\frac{89\alpha}{120}u^{(4)}(x_N)-\frac{1}{8}u^{(6)}(x_{N-1})+\mathcal{O}(h^7).\end{aligned}$$

Since $|\alpha|< h^6$, as $h \to 0$, we have $\tau_i = \mathcal{O}(h^6)$, $i = 1, 2, \ldots, N-1$.

Remark 8.2.1. *Once we get the the approximations U_i, $i = 0, 1, \ldots, N$, we can calculate the approximation S_i as $S_i = f_i - p_iU_i$, $i = 0, 1, \ldots, N$. We have $U'_0 = \hat{\eta}_0$ and $U'_N = \hat{\eta}_1$. In order to get the approximations U'_i, $i = 1, 2, \ldots, N-1$, we derive the following relation: Subtracting (8.28) with (8.27), we get*

$$\begin{aligned}U'_i =& -\frac{1}{2}U'_{i-1}+\frac{3\alpha}{4h}(U'_0+U'_N)-\frac{h^3}{480}[S_{i+1}+14S_i+5S_{i-1}]+\frac{9\alpha}{480h}S_0+\frac{11\alpha}{480h}S_N\\ &+\frac{1}{4h}[U_{i+1}+4U_i-5U_{i-1}]+\frac{3\alpha}{2h}[U_0-U_N]+\frac{\alpha}{8h}[\Phi^{(2)}(x_0)-\Phi^{(2)}(x_N)]\\ &-\frac{\alpha}{24h}[\Phi^{(3)}(x_N)-\Phi^{(3)}(x_0)],\ i=1,2,\ldots,N-1.\end{aligned}$$

Hence, we can compute the coefficients given in (8.25) and hence we get the fractal quintic spline that passes through the approximate solution of the differential equation (8.2). Also, the information required at off-nodal points can be obtained from (8.25).

8.2.2 Numerical examples

For any value of N, let

$$E_1^N = \max_i |u^{(1)}(x_i) - U_i'|, \quad p_1^N = E_1^N / E_1^{2N}.$$

Example 8.2.1. *Consider the BVP [136]*

$$\begin{cases} u^{(4)}(x) + xu(x) = -(8 + 7x + x^3)\exp(x), \ x \in (0,1), \\ u(0) = u(1) = 0, \quad u'(0) = 1, \ u'(1) = -\exp(1). \end{cases}$$

The exact solution is $u(x) = x(1-x)\exp(x)$. The scaling factor $\alpha = 0.9999h^6$ is used to compute the numerical solution. The maximum point-wise error and the order of convergence are given in Table 8.5. Table 8.6 represents the maximum point-wise error and the order of convergence corresponding to the quintic spline.

Table 8.5: Maximum point-wise error and order of convergence corresponding to Example 8.2.1.

N	16	32	64	128
E^N	4.4461e-06	6.6963e-07	1.3939e-07	3.3023e-08
p^N	2.7311	2.2642	2.0776	
E_1^N	1.8346e-05	3.1926e-06	7.6446e-07	1.8983e-07
p_1^N	2.5227	2.0622	2.0097	

Table 8.6: Maximum point-wise error and order of convergence corresponding to Example 8.2.1 with $\alpha = 0$.

N	16	32	64	128
E^N	4.2859e-05	1.0722e-05	2.6845e-06	6.7112e-07
p^N	1.9990	1.9979	2.0000	
E_1^N	1.3766e-04	3.4409e-05	8.6061e-06	2.1536e-06
p_1^N	2.0003	1.9994	1.9986	

Example 8.2.2. *Consider the BVP*

$$\begin{cases} u^{(4)}(x) - u(x) = -4(2x\cos x + 3\sin x),\ x \in (0,1), \\ u(0) = u(1) = 0, \quad u'(0) = -1,\ u'(1) = 2\sin(1). \end{cases}$$

The exact solution is $u(x) = (x^2 - 1)\sin x$. The maximum point-wise error and order of convergence are given in Table 8.7. The scaling factor $\alpha = 0.9999h^6$ is used to compute the solution. Table 8.8 represents the maximum point-wise error and the order of convergence corresponding to the quintic spline.

Table 8.7: Maximum point-wise error and order of convergence corresponding to Example 8.2.2.

N	16	32	64	128
E^N	1.4437e-06	2.1997e-07	5.5419e-08	1.4486e-08
p^N	2.7144	1.9888	1.9357	
E_1^N	6.9567e-06	1.4185e-06	3.5691e-07	8.9446e-08
p_1^N	2.2941	1.9907	1.9965	

Table 8.8: Maximum point-wise error and order of convergence corresponding to Example 8.2.2 with $\alpha = 0$.

N	16	32	64	128
E^N	1.6512e-05	4.1402e-06	1.0357e-06	2.5899e-07
p^N	1.9957	1.9991	1.9997	
E_1^N	5.3543e-05	1.3388e-05	3.3487e-06	8.3804e-07
p_1^N	1.9997	1.9993	1.9985	

8.3 Conclusion

To get the numerical approximation for the fourth-order BVPs, we have devised the numerical methods using fractal quintic spline. The truncation errors of the proposed methods are derived and the developed methods have second-order convergent. In order to see the applicability of the proposed methods, numerical examples are considered and tabulated the numerical results.

Bibliography

[1] M. Abbas, A. A. Majid, and J. M. Ali. Monotonicity preserving $\mathcal{C}^2$-rational cubic spline for monotone data. *Appl. Math. Comput.*, 219(6):2885–2895, 2012.

[2] M. Abukhaled, S. A. Khuri, and A. Sayfy. A numerical approach for solving a class of singular boundary-value problems arising in physiology. *Int. J. Numer. Anal. Model.*, 8(2):353–363, 2011.

[3] J. Ahlberg, E. Nilson, and J. Walsh. *The Theory of Splines and Their Applications.* Academic Press, New York, 1967.

[4] E. A. Al-Said and M. A. Noor. Quartic spline method for solving fourth-order obstacle boundary-value problems. *J. Comput. Appl. Math.*, 143(1):107–116, 2002.

[5] E. A. Al-Said, M. A. Noor, and T. M. Rassias. Cubic splines method for solving fourth-order obstacle problems. *Appl. Math. Comput.*, 174(1):180–187, 2006.

[6] E. Albasiny and W. Hoskins. Cubic spline solutions to two-point boundary-value problems. *Comput. J.*, 12(2):151–153, 1969.

[7] U. M. Ascher, R. M. M. Mattheij, and R. D. Russell. *Numerical Solution of Boundary-Value Problems for Ordinary Differential Equations.* Prentice-Hall, Englewood Cliffs, New Jersey, 1994.

[8] T. Aziz and A. Khan. Quintic spline approach to the solution of a singularly perturbed boundary-value problem. *J. Optim. Theory Appl.*, 112(3):517–527, 2002.

[9] T. Aziz and A. Khan. A spline method for second-order singularly perturbed boundary-value problems. *J. Comput. Appl. Math.*, 147(2):445–452, 2002.

[10] M. F. Barnsley. Fractal functions and interpolation. *Constr. Approx.*, 2(1):303–329, 1986.

[11] M. F. Barnsley and A. N. Harrington. The calculus of fractal interpolation functions. *J. Approx. Theory*, 57(1):14–34, 1989.

[12] R. K. Bawa. Spline based computational technique for linear singularly perturbed boundary-value problems. *Appl. Math. Comput.*, 167(1):225–236, 2005.

[13] R. K. Bawa and S. Natesan. Uniformly convergent computational technique for singularly perturbed self-adjoint mixed boundary-value problems. In *International Conference on Computational Science*, pages 1104–1107. Springer, 2005.

[14] R. K. Bawa and S. Natesan. Parallel implementation of a spline based computational approach for singular perturbation problems. In *International Conference on High-Performance Computing*, pages 254–262. Springer, 2006.

[15] R. K. Beatson and Z. Ziegler. Monotonicity preserving surface interpolation. *SIAM J. Numer. Anal.*, 22(2):401–411, 1985.

[16] G. Beliakov. Monotonicity preserving approximation of multivariate scattered data. *BIT*, 45(4):653–677, 2005.

[17] R. E. Bellman and R. E. Kalaba. *Quasilinearization and Nonlinear Boundary-Value Problems*. American Elsevier, New York, 1965.

[18] R. Bhatia, L. Elsner, and G. Krause. Bounds for the variation of the roots of a polynomial and the eigenvalues of a matrix. *Linear Algebra Appl.*, 142:195–209, 1990.

[19] S. K. Bhatta and K. S. Sastri. A sixth-order spline procedure for a class of nonlinear boundary-value problems. *Int. J. Comput. Math.*, 49(3-4):255–271, 1993.

[20] S. K. Bhatta and K. S. Sastri. Symmetric spline procedures for boundary-value problems with mixed boundary conditions. *J. Comput. Appl. Math.*, 45(3):237–250, 1993.

[21] W. Bickley. Piecewise cubic interpolation and two-point boundary problems. *Comput. J.*, 11(2):206–208, 1968.

[22] J. Bigge and E. Bohl. Deformations of the bifurcation diagram due to discretization. *Math. Comput.*, 45(172):393–403, 1985.

[23] E. Böhl. *Finite Modelle gewöhnlicher Randwertaufgaben*. Teubner, Stuttgart, 1981.

[24] P. Bouboulis. Construction of fractal surfaces via solutions of partial differential equations. *SRX Mathematics*, 2010, 2010.

[25] P. Bouboulis and L. Dalla. Closed fractal interpolation surfaces. *J. Math. Anal. Appl.*, 327(1):116–126, 2007.

[26] P. Bouboulis and L. Dalla. Fractal interpolation surfaces derived from fractal interpolation functions. *J. Math. Anal. Appl.*, 336(2):919–936, 2007.

[27] R. E. Carlson and F. N. Fritsch. Monotone piecewise bicubic interpolation. *SIAM J. Numer. Anal.*, 22(2):386–400, 1985.

[28] G. Casciola and L. Romani. Rational interpolants with tension parameters. *J. Curves and Surface Design*, pages 41–50, 2003.

[29] A. K. B. Chand. Coalescence cubic spline fractal interpolation surfaces. *Int. J. Numer. Anal. Model. Ser. B*, 3(2):1–16, 2012.

[30] A. K. B. Chand and G. P. Kapoor. Generalized cubic spline fractal interpolation functions. *SIAM J. Numer. Anal.*, 44(2):655–676, 2006.

[31] A. K. B. Chand and N. Vijender. Monotonicity preserving rational quadratic fractal interpolation functions. *Adv. Numer. Anal.*, Art. ID 504825, 17 pp, 2014.

[32] A. K. B. Chand and N. Vijender. Positive blending hermite rational cubic spline fractal interpolation surfaces. *Calcolo*, 52(1):1–24, 2015.

[33] A. K. B. Chand and N. Vijender. Monotonicity/symmetricity preserving rational quadratic fractal interpolation surfaces. *Int. J. Numer. Anal. Model.*, 13(1), 2016.

[34] A. K. B. Chand, N. Vijender, and M. A. Navascués. Shape preservation of scientific data through rational fractal splines. *Calcolo*, 51(2):329–362, 2014.

[35] A. K. B. Chand, N. Vijender, and M. A. Navascués. Convexity/concavity and stability aspects of rational cubic fractal interpolation surfaces. *Comput. Math. Model.*, 28(3):407–430, 2017.

[36] A. K. B. Chand and P. Viswanathan. A constructive approach to cubic hermite fractal interpolation function and its constrained aspects. *BIT*, 53(4):841–865, 2013.

[37] A. K. B. Chand, P. Viswanathan, and N. Vijender. Bivariate shape preserving interpolation: A fractal-classical hybrid approach. *Chaos Solitons Fractals*, 81:330–344, 2015.

[38] A. K. B. Chand, P. Viswanathan, and N. Vijender. Bicubic partially blended rational fractal surface for a constrained interpolation problem. *Comput. Appl. Math.*, 37(1):785–804, 2018.

[39] M. M. Chawla. An eighth-order tridiagonal finite difference method for nonlinear two-point boundary-value problems. *BIT*, 17(3):281–285, 1977.

[40] M. M. Chawla and C. P. Katti. A finite difference method for a class of singular two-point boundary-value problems. *IMA J. Numer. Anal.*, 4(4):457–466, 1984.

[41] M. M. Chawla, S. McKee, and G. Shaw. Order h^2 method for a singular two-point boundary-value problem. *BIT*, 26(3):318–326, 1986.

[42] M. M. Chawla and P. N. Shivakumar. Numerov's method for non-linear two-point boundary-value problems. *Int. J. Comput. Math.*, 17(2):167–176, 1985.

[43] M. M. Chawla and R. Subramanian. A new fourth-order cubic spline method for non-linear two-point boundary-value problems. *Int. J. Comput. Math.*, 22(3-4):321–341, 1987.

[44] M. M. Chawla and R. Subramanian. A new fourth-order cubic spline method for second-order nonlinear two-point boundary-value problems. *J. Comput. Appl. Math.*, 23(1):1–10, 1988.

[45] M. M. Chawla and R. Subramanian. High accuracy quintic spline solution of fourth-order two-point boundary-value problems. *Int. J. Comput. Math.*, 31(1-2):87–94, 1989.

[46] M. M. Chawla, R. Subramanian, and P. N. Shivakumar. Numerov's method for non-linear two-point boundary-value problems ii. monotone approximations. *Int. J. Comput. Math.*, 26(3-4):219–227, 1989.

[47] F. Combes. Astrophysical fractals: Interstellar medium and galaxies. In *The Chaotic Universe*, pages 143–172. World Scientific, 2000.

[48] P. Costantini. On monotone and convex spline interpolation. *Math. Comp.*, 46(173):203–214, 1986.

[49] P. Costantini. Curve and surface construction using variable degree polynomial splines. *Comput. Aided Geom. Design*, 17(5):419–446, 2000.

[50] P. Costantini and F. Fontanella. Shape preserving bivariate interpolation. *SIAM J. Numer. Anal.*, 27(2):488–506, 1990.

[51] L. Dalla. Bivariate fractal interpolation functions on grids. *Fractals*, 10(01):53–58, 2002.

[52] C. De Boor. *The Method of Projection as Applied to the Numerical Solution of Two-Point Boundary-Value Problems using Cubic Splines*. PhD thesis, 1966.

[53] C. De Boor and B. Swartz. Piecewise monotone interpolation. *J. Approx. Theory*, 21(4):411–416, 1977.

[54] R. Delbourgo and J. A. Gregory. The determination of derivative parameters for a monotonic rational quadratic interpolant. *IMA J. Numer. Anal.*, 5(4):397–406, 1985.

[55] R. Delbourgo and J. A. Gregory. Shape preserving piecewise rational interpolation. *SIAM J. Sci. Statist. Comput.*, 6(4):967–976, 1985.

[56] K. J. Falcomer. *Fractal Geometry: Mathematical Foundation and Applications*. Wiley, New York, 1990.

[57] G. E. Farin. *Curves and Surfaces for CAGD.* Morgan Kaufmann, Burlington, 2002.

[58] Z. Feng, Y. Feng, and Z. Yuan. Fractal interpolation surfaces with function vertical scaling factors. *Appl. Math. Lett.*, 25(11):1896–1900, 2012.

[59] A. Fielding. Applications of fractal geometry to biology. *Bioinformatics*, 8(4):359–366, 1992.

[60] F. N. Fritsch and J. Butland. A method for constructing local monotone piecewise cubic interpolants. *SIAM J. Sci. Statist. Comput.*, 5(2):300–304, 1984.

[61] F. N. Fritsch and R. E. Carlson. Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.*, 17(2):238–246, 1980.

[62] J. S. Geronimo and D. Hardin. Fractal interpolation surfaces and a related 2-D multiresolution analysis. *J. Math. Anal. Appl.*, 176(2):561–586, 1993.

[63] J. A. Gregory and R. Delbourgo. Piecewise rational quadratic interpolation to monotonic data. *IMA J. Numer. Anal.*, 2(2):123–130, 1982.

[64] T. N. E. Greville. Introduction to spline functions. In *Theory and Applications of Spline Functions*, pages 1–35. Academic press, New York, 1969.

[65] B. Gustafsson. A numerical method for solving singular boundary-value problems. *Numer. Math.*, 21(4):328–344, 1973.

[66] S. Havlin, S. V. Buldyrev, A. L. Goldberger, R. N. Mantegna, S. M. Ossadnik, C. K. Peng, M. Simons, and H. E. Stanley. Fractals in biology and medicine. *Chaos Solitons Fractals*, 6:171–201, 1995.

[67] P. Henrici. *Discrete Variable Methods in Ordinary Differential Equations.* John Wiley and Sons, New York, 1962.

[68] M. Z. Hussain, M. Hussain, and B. Aqeel. Shape preserving surfaces with constraints on tension parameters. *Appl. Math. Comput.*, 247:442–464, 2014.

[69] J. M. Hutchinson, A. W. Lo, and T. Poggio. A nonparametric approach to pricing and hedging derivative securities via learning networks. *J. Finance*, 49(3):851–889, 1994.

[70] M. Irodotou-Ellina and E. N. Houstis. An $\mathcal{O}(h^6)$ quintic spline collocation method for fourth-order two-point boundary-value problems. *BIT*, 28(2):288–301, 1988.

[71] S. K. Iyengar and P. Jain. Spline finite difference methods for singular two-point boundary-value problems. *Numer. Math.*, 50(3):363–376, 1986.

[72] M. K. Jain and T. Aziz. Cubic spline solution of two-point boundary-value problems with significant first derivatives. *Comput. Methods Appl. Mech. Engrg.*, 39(1):83–91, 1983.

[73] P. Jamet. On the convergence of finite difference approximations to one-dimensional singular boundary-value problems. *Numer. Math.*, 14(4):355–378, 1970.

[74] M. K. Kadalbajoo and R. K. Bawa. Variable-mesh difference scheme for singularly perturbed boundary-value problems using splines. *J. Optim. Theory Appl.*, 90(2):405–416, 1996.

[75] M. K. Kadalbajoo and A. Jha. Exponentially fitted cubic spline for two-parameter singularly perturbed boundary value problems. *Int. J. Comput. Math.*, 89(6):836–850, 2012.

[76] M. K. Kadalbajoo and K. C. Patidar. Spline techniques for solving singularly perturbed nonlinear problems on nonuniform grids. *J. Optim. Theory Appl.*, 114(3):573–591, 2002.

[77] M. K. Kadalbajoo and K. S. Raman. Numerical solution of singular boundary-value problems by invariant imbedding. *J. Comput. Phys.*, 55(2):268–277, 1984.

[78] P. D. Kaklis and D. G. Pandelis. Convexity preserving polynomial splines of non-uniform degree. *IMA J. Numer. Anal.*, 10(2):223–234, 1990.

[79] P. D. Kaklis and N. S. Sapidis. Convexity preserving interpolatory parametric splines of non-uniform polynomial degree. *Comput. Aided Geom. Design*, 12(1):1–26, 1995.

[80] A. Khan. *Spline Solution of Differential Equations.* PhD thesis, 2001.

[81] A. Khan, I. Khan, and T. Aziz. Sextic spline solution of a singularly perturbed boundary-value problems. *Appl. Math. Comput.*, 181(1):432–439, 2006.

[82] A. Khan and P. Khandelwal. Non-polynomial sextic spline solution of singularly perturbed boundary-value problems. *Int. J. Comput. Math.*, 91(5):1122–1135, 2014.

[83] M. A. Khan, S. U. Islam, I. A. Tirmizi, E. H. Twizell, and S. Ashraf. A class of methods based on non-polynomial sextic spline functions for the solution of a special fifth-order boundary-value problems. *J. Math. Anal. Appl.*, 321(2):651–660, 2006.

[84] L. M. Kocić. Fractals and their applications in computer graphics. *Filomat*, pages 207–231, 1995.

[85] A. Kouibia and M. Pasadas. Approximation by interpolating variational splines. *J. Comput. Appl. Math.*, 218(2):342–349, 2008.

[86] P. Lamberti and C. Manni. Shape-preserving $\mathcal{C}^2$-functional interpolation via parametric cubics. *Numer. Algorithms*, 28(1-4):229–254, 2001.

[87] B. Lin, K. Li, and Z. Cheng. B-spline solution of a singularly perturbed boundary-value problem arising in biology. *Chaos Solitons Fractals*, 42(5):2934–2948, 2009.

[88] G. B. Loghmani and S. R. Alavizadeh. Numerical solution of fourth-order problems with separated boundary conditions. *Appl. Math. Comput.*, 191(2):571–581, 2007.

[89] F. R. Loscalzo and T. D. Talbot. Spline function approximations for solutions of ordinary differential equations. *Bull. Amer. Math. Soc.*, 73:438–442, 1967.

[90] B. Mandelbrot. Stochastic models for the earth's relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islands. *Proceedings of the National Academy of Sciences*, 72(10):3825–3828, 1975.

[91] B. Mandelbrot. *Fractals: Form, Chance, and Dimension.* W H Freeman, San Francisco, 1977.

[92] B. Mandelbrot. *The Fractal Geometry of Nature.* W H freeman, New York, 1983.

[93] P. R. Massopust. Fractal surfaces. *J. Math. Anal. Appl.*, 151(1):275–290, 1990.

[94] G. Micula. Numerical integration of differential equation $y^{(n)} = f(x, y)$ by spline functions. *Rev. Roum. Math. Pures et Appl.(Bucarest)*, 17:1385–1390, 1972.

[95] G. Micula and S. Micula. *Handbook of Splines.* Kluwer Academic Publishers, Dordrecht, 1999.

[96] A. D. Miller and R. Vyborny. Some remarks on functions with one-sided derivatives. *Amer. Math. Monthly*, 93(6):471–475, 1986.

[97] R. K. Mohanty and N. Jha. A class of variable mesh spline in compression methods for singularly perturbed two-point singular boundary-value problems. *Appl. Math. Comput.*, 168(1):704–716, 2005.

[98] S. Natesan, R. K. Bawa, and C. Clavero. Uniformly convergent compact numerical scheme for the normalized flux of singularly perturbed reaction–diffusion problems. *Int. J. Inf. Syst. Sci.*, 3(2):207–221, 2007.

[99] M. Navascués. Non-smooth polynomials. *Int. J. Math. Anal*, 1(1-4):159–174, 2007.

[100] M. A. Navascués and M. V. Sebastián. Smooth fractal interpolation. *J. Inequal. Appl.*, 2006(1):78734, 2006.

[101] A. Nikolis. Numerical solutions of ordinary differential equations with quadratic trigonometric splines. *Appl. Math. E-Notes*, 4:142–149, 2004.

[102] A. Nikolis and I. Seimenis. Solving dynamical systems with cubic trigonometric splines. *Appl. Math. E-Notes*, 5:116–123, 2005.

[103] E. Passow and J. A. Roulier. Monotone and convex spline interpolation. *SIAM J. Numer. Anal.*, 14(5):904–909, 1977.

[104] M. J. D. Powell. *Approximation Theory and Methods*. Cambridge university press, Cambridge, 1981.

[105] P. M. Prenter. *Splines and Variational Methods*. Dover, New York, 2008.

[106] M. A. Ramadan, I. F. Lashien, and W. K. Zahra. High-order accuracy nonpolynomial spline solutions for 2μ-th-order two-point boundary-value problems. *Appl. Math. Comput.*, 204(2):920–927, 2008.

[107] M. A. Ramadan, I. F. Lashien, and W. K. Zahra. Quintic nonpolynomial spline solutions for fourth-order two-point boundary-value problem. *Commun. Nonlinear Sci. Numer. Simul.*, 14(4):1105–1114, 2009.

[108] J. Rashidinia, M. Ghasemi, and Z. Mahmoodi. Spline approach to the solution of a singularly perturbed boundary-value problems. *Appl. Math. Comput.*, 189(1):72–78, 2007.

[109] J. Rashidinia and A. Golbabaee. Convergence of numerical solution of a fourth-order boundary-value problem. *Appl. Math. Comput.*, 171(2):1296–1305, 2005.

[110] J. Rashidinia, Z. Mahmoodi, and M. Ghasemi. Parametric spline method for a class of singular two-point boundary-value problems. *Appl. Math. Comput.*, 188(1):58–63, 2007.

[111] J. Rashidinia and R. Mohammadi. Non-polynomial spline approximations for the solution of singularly perturbed boundary-value problems. *TWMS J. Pure Appl. Math.*, 1(2):236–251, 2010.

[112] J. Rashidinia, R. Mohammadi, and R. Jalilian. The numerical solution of non-linear singular boundary-value problems arising in physiology. *Appl. Math. Comput.*, 185(1):360–367, 2007.

[113] J. Rashidinia, R. Mohammadi, and R. Jalilian. Spline solution of non-linear singular boundary-value problems. *Int. J. Comput. Math.*, 85(1):39–52, 2008.

[114] J. Rashidinia, R. Mohammadi, and S. H. Moatamedoshariati. Quintic spline methods for the solution of singularly perturbed boundary-value problems. *Int. J. Comput. Methods Eng. Sci. Mech.*, 11(5):247–257, 2010.

[115] A. S. V. Ravi Kanth and V. Bhattacharya. Cubic spline for a class of non-linear singular boundary-value problems arising in physiology. *Appl. Math. Comput.*, 174(1):768–774, 2006.

[116] A. S. V. Ravi Kanth and Y. N. Reddy. A numerical method for singular two-point boundary-value problems via chebyshev economizition. *Appl. Math. Comput.*, 146(2):691–700, 2003.

[117] A. S. V. Ravi Kanth and Y. N. Reddy. Cubic spline for a class of singular two-point boundary-value problems. *Appl. Math. Comput.*, 170(2):733–740, 2005.

[118] J. N. Reddy. *An Introduction to the Finite Element Method.* McGraw-hill, New York, 1993.

[119] R. J. Renka. Interpolation of scattered data with a $\mathcal{C}^1$ convexity preserving surface. *ACM Transactions on Mathematical Software*, 30(2):200–211, 2004.

[120] R. D. Russell and L. F. Shampine. A collocation method for boundary-value problems. *Numer. Math.*, 19(1):1–28, 1972.

[121] M. Sakai. Piecewise cubic interpolation and two-point boundary-value problems. *Publ. Res. Inst. Math. Sci.*, 7(2):345–362, 1971.

[122] N. S. Sapidis, P. D. Kaklis, and T. A. Loukakis. A method for computing the tension parameters in convexity preserving spline in tension interpolation. *Numer. Math.*, 54(2):179–192, 1989.

[123] M. Sarfraz and M. Z. Hussain. Data visualization using rational spline interpolation. *J. Comput. Appl. Math.*, 189(1-2):513–525, 2006.

[124] M. Sarfraz, M. Z. Hussain, and M. Hussain. Modeling rational spline for visualization of shaped data. *J. Numer. Math.*, 21(1):63–88, 2013.

[125] J. W. Schmidt and W. Heß. Positive interpolation with rational quadratic splines. *Computing*, 38(3):261–267, 1987.

[126] I. J. Schoenberg. Contributions to the problem of approximation of equidistant data by analytic functions. part. A: On the problem of smoothing of graduations. A first class of analytic approximation formulae. *Quart. Appl. Math.*, 4(1):45–99, 1946.

[127] L. I. Schumaker. On shape preserving quadratic spline interpolation. *SIAM J. Numer. Anal.*, 20(4):854–864, 1983.

[128] D. G. Schweikert. An interpolation curve using a spline in tension. *J. Math. and Phys.*, 45:312–317, 1966.

[129] S. S. Siddiqi and G. Akram. Solution of the system of fourth-order boundary-value problems using non-polynomial spline technique. *Appl. Math. Comput.*, 185(1):128–135, 2007.

[130] S. S. Siddiqi and G. Akram. Quintic spline solutions of fourth-order boundary-value problems. *Int. J. Numer. Anal. Model.*, 5(1):101–111, 2008.

[131] H. Späth. *Spline Algorithms for Curves and Surfaces.* Utilitas Mathematica Pub., Winnipeg, 1974.

[132] K. Surla and M. Stojanović. Solving singularly perturbed boundary-value problems by spline in tension. *J. Comput. Appl. Math.*, 24(3):355–363, 1988.

[133] R. P. Tewarson. On the use of splines for the numerical solution of nonlinear two-point boundary-value problems. *BIT*, 20(2):223–232, 1980.

[134] I. A. Tirmizi, F. I. Haq, and S. U. Islam. Non-polynomial spline solution of singularly perturbed boundary-value problems. *Appl. Math. Comput.*, 196(1):6–16, 2008.

[135] I. A. Tirmizi and E. H. Twizell. Higher-order finite difference methods for nonlinear second-order two-point boundary-value problems. *Appl. Math. Lett.*, 15(7):897–902, 2002.

[136] R. A. Usmani. Smooth spline approximations for the solution of a boundary-value problem with engineering applications. *J. Comput. Appl. Math.*, 6(2):93–98, 1980.

[137] R. A. Usmani. The use of quartic splines in the numerical solution of a fourth-order boundary-value problem. *J. Comput. Appl. Math.*, 44(2):187–200, 1992.

[138] R. A. Usmani and M. Sakai. A connection between quartic spline solution and numerov solution of a boundary-value problem. *Int. J. Comput. Math.*, 26(3-4):263–273, 1989.

[139] R. A. Usmani and S. A. Warsi. Smooth spline solutions for boundary-value problems in plate deflection theory. *Comput. Math. Appl.*, 6(2):205–211, 1980.

[140] M. Van Daele, G. V. Berghe, and H. De Meyer. A smooth approximation for the solution of a fourth-order boundary-value problem based on nonpolynomial splines. *J. Comput. Appl. Math.*, 51(3):383–394, 1994.

[141] N. Vijender and A. K. B. Chand. Shape preserving affine fractal interpolation surfaces. *Nonlinear Stud.*, 21(2), 2014.

[142] P. Viswanathan and A. K. B. Chand. α-fractal rational splines for constrained interpolation. *Electron. Trans. Numer. Anal.*, 41:420–442, 2014.

[143] P. Viswanathan and A. K. B. Chand. A fractal procedure for monotonicity preserving interpolation. *Appl. Math. Comput.*, 247:190–204, 2014.

[144] P. Viswanathan and A. K. B. Chand. A $\mathcal{C}^1$-rational cubic fractal interpolation function: Convergence and associated parameter identification problem. *Acta Appl. Math.*, 136(1):19–41, 2015.

[145] P. Viswanathan, A. K. B. Chand, and R. P. Agarwal. Preserving convexity through rational cubic spline fractal interpolation function. *J. Comput. Appl. Math.*, 263:262–276, 2014.

[146] P. Viswanathan, A. K. B. Chand, and M. A. Navascués. A rational iterated function system for resolution of univariate constrained interpolation. *Rev. R. Acad. Cienc. Exactas Fs. Nat. Ser. A Math. RACSAM*, 109(2):483–509, 2015.

[147] R. F. Voss. Random fractal forgeries. In *Fundamental Algorithms for Computer Graphics*, pages 805–835. Springer, 1985.

[148] R. F. Voss. Random fractals: Characterization and measurement. In *Scaling Phenomena in Disordered Systems*, pages 1–11. Springer, 1991.

[149] H. Xie and H. Sun. The study on bivariate fractal interpolation functions and creation of fractal interpolated surfaces. *Fractals*, 5(04):625–634, 1997.

[150] Y. Zhang, Q. Duan, and E. H. Twizell. Convexity control of a bivariate rational interpolating spline surfaces. *Comput. Graphics*, 31(5):679–687, 2007.

[151] N. Zhao. Construction and application of fractal interpolation surfaces. *Vis. Comput.*, 12(3):132–146, 1996.

List of published and communicated papers

Based on the work carried out in this thesis, the following published and communicated papers have resulted:

- N. Balasubramani. Shape preserving rational cubic fractal interpolation function. *Journal of Computational and Applied Mathematics*, 319, 277-295, 2017.
- N. Balasubramani, M. G. P. Prasad and S. Natesan. Rational cubic fractal spline for visualization of shaped data. *Neural, Parallel, and Scientific Computations*, 26(2), 183-210, 2018.
- N. Balasubramani, M. G. P. Prasad and S. Natesan. Constrained and convex interpolation through rational cubic fractal interpolation surface. *Computational and Applied Mathematics*, 37, 6308–6331, 2018.
- N. Balasubramani, M. G. P. Prasad and S. Natesan. Fractal quintic spline method for non-linear boundary-value problems (communicated).
- N. Balasubramani, M. G. P. Prasad and S. Natesan. Fractal quintic spline solutions for fourth-order boundary-value problems (communicated).