Volume 21, Number 1, 2021 ISSN 1562-8353

NONLINEAR DYNAMICS AND SYSTEMS THEORY
An International Journal of Research and Surveys

Volume 21 Number 1 2021

CONTENTS

 ISSN 1562-8353 Printed in Ukraine
To receive contents and abstracts by e-mail, visit our Website at: **http://www.e-ndst.kiev.ua**

NONLINEAR DYNAMICS & SYSTEMS THEORY Volume 21, No. 1, 2021

Nonlinear Dynamics and Systems Theory

An International Journal of Research and Surveys

InforMath Publishing Group
http://www.e-ndst.kiev.ua

Nonlinear Dynamics and Systems Theory

An International Journal of Research and Surveys

INSTRUCTIONS FOR CONTRIBUTORS

(1) General. Nonlinear Dynamics and Systems Theory (ND&ST) is an international journal devoted to publishing peer-refereed, high quality, original papers, brief notes and review articles focusing on nonlinear dynamics and systems theory and their practical applications in engineering, physical and life sciences. Submission of a manuscript is a representation that the submission has been approved by all of the authors and by the institution where the work was carried out. It also represents that the manuscript has not been previously published, has not been copyrighted, is not being submitted for publication elsewhere, and that the authors have agreed that the copyright in the article shall be assigned exclusively to InforMath Publishing Group by signing a transfer of copyright form. Before submission, the authors should visit the website:

http://www.e-ndst.kiev.ua

for information on the preparation of accepted manuscripts. Please download the archive Sample_NDST.zip containing example of article file (you can edit only the file Samplefilename.tex).

(2) Manuscript and Correspondence. Manuscripts should be in English and must meet common standards of usage and grammar. To submit a paper, send by e-mail a file in PDF format directly to

Professor A.A. Martynyuk, Institute of Mechanics,
Nesterov str.3, 03057, MSP 680, Kiev-57, Ukraine
e-mail: journalndst@gmail.com

or to one of the Regional Editors or to a member of the Editorial Board. Final version of the manuscript must typeset using LaTex program which is prepared in accordance with the style file of the Journal. Manuscript texts should contain the title of the article, name(s) of the author(s) and complete affiliations. Each article requires an abstract not exceeding 150 words. Formulas and citations should not be included in the abstract. AMS subject classifications and key words must be included in all accepted papers. Each article requires a running head (abbreviated form of the title) of no more than 30 characters. The sizes for regular papers, survey articles, brief notes, letters to editors and book reviews are: (i) 10-14 pages for regular papers, (ii) up to 24 pages for survey articles, and (iii) 2-3 pages for brief notes, letters to the editor and book reviews.

(3) Tables, Graphs and Illustrations. Each figure must be of a quality suitable for direct reproduction and must include a caption. Drawings should include all relevant details and should be drawn professionally in black ink on plain white drawing paper. In addition to a hard copy of the artwork, it is necessary to attach the electronic file of the artwork (preferably in PCX format).

(4) References. Each entry must be cited in the text by author(s) and number or by number alone. All references should be listed in their alphabetic order. Use please the following style:

Journal: [1] H. Poincare, Title of the article. *Title of the Journal* **volume** (issue) (year) pages. [Language]

Book: [2] A.M. Lyapunov, *Title of the Book.* Name of the Publishers, Town, year.

Proceeding: [3] R. Bellman, Title of the article. In: *Title of the Book.* (Eds.). Name of the Publishers, Town, year, pages. [Language]

(5) Proofs and Sample Copy. Proofs sent to authors should be returned to the Editorial Office with corrections within three days after receipt. The corresponding author will receive a sample copy of the issue of the Journal for which his/her paper is published.

(6) Editorial Policy. Every submission will undergo a stringent peer review process. An editor will be assigned to handle the review process of the paper. He/she will secure at least two reviewers' reports. The decision on acceptance, rejection or acceptance subject to revision will be made based on these reviewers' reports and the editor's own reading of the paper.

NONLINEAR DYNAMICS AND SYSTEMS THEORY
An International Journal of Research and Surveys
Published by InforMath Publishing Group since 2001

Volume 21 Number 1 2021

CONTENTS

Founded by A.A. Martynyuk in 2001.
Registered in Ukraine Number: KB No 5267 / 04.07.2001.

NONLINEAR DYNAMICS AND SYSTEMS THEORY

An International Journal of Research and Surveys

Impact Factor from SCOPUS for 2019: SJR – 0.328, SNIP – 0.591 and CiteScore – 1.4

Nonlinear Dynamics and Systems Theory (ISSN 1562–8353 (Print), ISSN 1813–7385 (Online)) is an international journal published under the auspices of the S.P. Timoshenko Institute of Mechanics of National Academy of Sciences of Ukraine and Curtin University of Technology (Perth, Australia). It aims to publish high quality original scientific papers and surveys in areas of nonlinear dynamics and systems theory and their real world applications.

AIMS AND SCOPE

Nonlinear Dynamics and Systems Theory is a multidisciplinary journal. It publishes papers focusing on proofs of important theorems as well as papers presenting new ideas and new theory, conjectures, numerical algorithms and physical experiments in areas related to nonlinear dynamics and systems theory. Papers that deal with theoretical aspects of nonlinear dynamics and/or systems theory should contain significant mathematical results with an indication of their possible applications. Papers that emphasize applications should contain new mathematical models of real world phenomena and/or description of engineering problems. They should include rigorous analysis of data used and results obtained. Papers that integrate and interrelate ideas and methods of nonlinear dynamics and systems theory will be particularly welcomed. This journal and the individual contributions published therein are protected under the copyright by International InforMath Publishing Group.

PUBLICATION AND SUBSCRIPTION INFORMATION

Nonlinear Dynamics and Systems Theory will have 4 issues in 2021, printed in hard copy (ISSN 1562–8353) and available online (ISSN 1813–7385), by InforMath Publishing Group, Nesterov str., 3, Institute of Mechanics, Kiev, MSP 680, Ukraine, 03057. Subscription prices are available upon request from the Publisher, EBSCO Information Services (`mailto:journals@ebsco.com`), or website of the Journal: `http://e-ndst.kiev.ua`. Subscriptions are accepted on a calendar year basis. Issues are sent by airmail to all countries of the world. Claims for missing issues should be made within six months of the date of dispatch.

ABSTRACTING AND INDEXING SERVICES

Papers published in this journal are indexed or abstracted in: Mathematical Reviews / MathSciNet, Zentralblatt MATH / Mathematics Abstracts, PASCAL database (INIST–CNRS) and SCOPUS.

Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 1–12

Upper and Lower Solutions for Fractional q-Difference Inclusions

S. Abbas [1], M. Benchohra [2] and J. R. Graef [3*]

[1] *Department of Mathematics, Tahar Moulay University of Saïda, P.O. Box 138, EN-Nasr, 20000 Saïda, Algeria.*
[2] *Laboratory of Mathematics, Djillali Liabes University of Sidi Bel-Abbès, P.O. Box 89, Sidi Bel-Abbès 22000, Algeria, and Department of Mathematics, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia.*
[3] *Department of Mathematics, University of Tennessee at Chattanooga, Chattanooga, TN 37403, USA.*

Received: October 29, 2019; Revised: December 24, 2020

Abstract: This paper deals with some existence results for a class of boundary value problems for Caputo fractional q-difference inclusions by using set-valued analysis, fixed point theory, and the method of upper and lower solutions.

Keywords: *fractional q-difference inclusion; upper solution; lower solution; boundary condition; fixed point.*

Mathematics Subject Classification (2010): 26A33, 34A08, 34A60, 34B15, 39A13.

1 Introduction

Fractional differential equations and inclusions have been applied in various areas of engineering, mathematics, physics, and other applied sciences. Recently, considerable attention has been given to the existence of solutions of initial and boundary value problems for fractional differential equations and inclusions with Caputo fractional derivatives. The method of upper and lower solutions has been successfully applied to study the existence of solutions for differential equations and inclusions; see [1–5, 11, 12] and the references therein.

The study of fractional q-difference equations was initiated early in the 20-th century [6, 14] and has received significant attention in recent years [10, 16]. Some interesting details about initial and boundary value problems for q-difference and fractional

* Corresponding author: mailto:John-Graef@utc.edu

q-difference equations can be found in [8, 9, 15–17] and the included references. In this paper, we discuss the existence of solutions to the fractional q-difference inclusion

$$({}^cD_q^\alpha u)(t) \in F(t, u(t)),\ t \in I := [0, T], \tag{1}$$

with the boundary condition

$$L(u(0), u(T)) = 0, \tag{2}$$

where $q \in (0,1),\ \alpha \in (0,1],\ T > 0,\ F : I \times \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is a multivalued map, $\mathcal{P}(\mathbb{R})$ is the family of all nonempty subsets of $\mathbb{R}$, ${}^cD_q^\alpha$ is the Caputo fractional q-difference derivative of order α, and $L : \mathbb{R}^2 \to \mathbb{R}$ is a given continuous function.

This paper initiates the application of the method of upper and lower solutions to Caputo q-fractional difference equations.

2 Preliminaries

Consider the Banach space $C(I) := C(I, \mathbb{R})$ of continuous functions from I into $\mathbb{R}$ equipped with the supremum (uniform) norm

$$\|u\|_\infty := \sup_{t \in I} |u(t)|.$$

As usual, $L^1(I)$ denotes the space of measurable functions $v : I \to \mathbb{R}$ that are Lebesgue integrable with the norm

$$\|v\|_1 = \int_0^T |v(t)| dt.$$

Let us recall some definitions and properties of the fractional q-calculus. For $a \in \mathbb{R}$, we set

$$[a]_q = \frac{1 - q^a}{1 - q}.$$

The q analogue of the power $(a - b)^n$ is

$$(a - b)^{(0)} = 1,\ (a - b)^{(n)} = \Pi_{k=0}^{n-1}(a - bq^k),\ a, b \in \mathbb{R},\ n \in \mathbb{N}.$$

In general,

$$(a - b)^{(\alpha)} = a^\alpha \Pi_{k=0}^{\infty} \left(\frac{a - bq^k}{a - bq^{k+\alpha}} \right),\ a, b, \alpha \in \mathbb{R}.$$

Definition 2.1 ([19]) The q-gamma function is defined by

$$\Gamma_q(\xi) = \frac{(1 - q)^{(\xi - 1)}}{(1 - q)^{\xi - 1}} \text{ for } \xi \in \mathbb{R} - \{0, -1, -2, \dots\}.$$

Notice that the q-gamma function satisfies $\Gamma_q(1 + \xi) = [\xi]_q \Gamma_q(\xi)$.

Next, we give definitions of different types of q-derivatives and q-integrals and indicate some of their properties.

Definition 2.2 ([19]) The q-derivative of order $n \in \mathbb{N}$ of a function $u : I \to \mathbb{R}$ is defined by $(D_q^0 u)(t) = u(t)$,

$$(D_q u)(t) := (D_q^1 u)(t) = \frac{u(t) - u(qt)}{(1 - q)t},\ t \neq 0,\quad (D_q u)(0) = \lim_{t \to 0} (D_q u)(t),$$

and

$$(D_q^n u)(t) = (D_q D_q^{n-1} u)(t),\ t \in I,\ n \in \{1, 2, \dots\}.$$

We set $I_t := \{tq^n : n \in \mathbb{N}\} \cup \{0\}$.

Definition 2.3 ([19]) The q-integral of a function $u : I_t \to \mathbb{R}$ is defined by

$$(I_q u)(t) = \int_0^t u(s) d_q s = \sum_{n=0}^{\infty} t(1-q)q^n f(tq^n),$$

provided that the series converges.

We note that $(D_q I_q u)(t) = u(t)$, while if u is continuous at 0, then

$$(I_q D_q u)(t) = u(t) - u(0).$$

Definition 2.4 ([7]) The Riemann-Liouville fractional q-integral of order $\alpha \in \mathbb{R}_+ := [0,\infty)$ of a function $u : I \to \mathbb{R}$ is defined by $(I_q^0 u)(t) = u(t)$, and

$$(I_q^\alpha u)(t) = \int_0^t \frac{(t-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} u(s) d_q s, \ t \in I.$$

Lemma 2.1 ([20]) *For $\alpha \in \mathbb{R}_+ := [0,\infty)$ and $\lambda \in (-1,\infty)$, we have*

$$(I_q^\alpha (t-a)^{(\lambda)})(t) = \frac{\Gamma_q(1+\lambda)}{\Gamma(1+\lambda+\alpha)} (t-a)^{(\lambda+\alpha)}, \ 0 < a < t < T.$$

In particular,

$$(I_q^\alpha 1)(t) = \frac{1}{\Gamma_q(1+\alpha)} t^{(\alpha)}.$$

Definition 2.5 ([21]) The Riemann-Liouville fractional q-derivative of order $\alpha \in \mathbb{R}_+$ of a function $u : I \to \mathbb{R}$ is defined by $(D_q^0 u)(t) = u(t)$, and

$$(D_q^\alpha u)(t) = (D_q^{[\alpha]} I_q^{[\alpha]-\alpha} u)(t), \ t \in I,$$

where $[\alpha]$ is the integer part of α.

Definition 2.6 ([21]) The Caputo fractional q-derivative of order $\alpha \in \mathbb{R}_+$ of a function $u : I \to \mathbb{R}$ is defined by $({}^C D_q^0 u)(t) = u(t)$ and

$$({}^C D_q^\alpha u)(t) = (I_q^{[\alpha]-\alpha} D_q^{[\alpha]} u)(t), \ t \in I.$$

Lemma 2.2 ([21]) *Let $\alpha \in \mathbb{R}_+$. Then the following equality holds:*

$$(I_q^\alpha \ {}^C D_q^\alpha u)(t) = u(t) - \sum_{k=0}^{[\alpha]-1} \frac{t^k}{\Gamma_q(1+k)} (D_q^k u)(0).$$

In particular, if $\alpha \in (0,1)$, then

$$(I_q^\alpha \ {}^C D_q^\alpha u)(t) = u(t) - u(0).$$

For a given Banach space $(X, \|\cdot\|)$, we define the following subsets of $\mathcal{P}(X)$:

$$P_{cl}(X) = \{Y \in \mathcal{P}(X) : Y \text{ is closed}\}, \quad P_b(X) = \{Y \in \mathcal{P}(X) : Y \text{ is bounded}\},$$

$$P_{cp}(X) = \{Y \in \mathcal{P}(X) : Y \text{ is compact}\}, \quad P_{cv}(X) = \{Y \in \mathcal{P}(X) : Y \text{ is convex}\},$$

$$P_{cp,cv}(X) = P_{cp}(X) \Cap P_{cv}(X).$$

The following properties of multivalued maps will be needed.

Definition 2.7 A multivalued map $G : X \to \mathcal{P}(X)$ is said to be convex (closed) valued if $G(x)$ is convex (closed) for all $x \in X$. A multivalued map G is bounded on bounded sets if $G(B) = \cup_{x\in B} G(x)$ is bounded in X for all $B \in P_b(X)$ (i.e., $\sup_{x\in B}\{\sup\{|y| : y \in G(x)\}$ exists).

Definition 2.8 A multivalued map $G : X \to \mathcal{P}(X)$ is called upper semi-continuous (u.s.c.) on X if for each $x_0 \in X$, the set $G(x_0)$ is a nonempty closed subset of X, and for each open set $N \subset X$ containing $G(x_0)$, there exists an open neighborhood N_0 of x_0 such that $G(N_0) \subset N$. Moreover, G is said to be completely continuous if $G(B)$ is relatively compact for every $B \in P_b(X)$.

Definition 2.9 Let $G : X \to \mathcal{P}(X)$ be completely continuous with nonempty compact values. Then G is u.s.c. if and only if G has a closed graph (i.e., $x_n \to x_*, y_n \to y_*, y_n \in G(x_n)$ imply $y_* \in G(x_*)$). We say that G has a fixed point if there is $x \in X$ such that $x \in G(x)$.

We denote by $Fix\, G$ the set of fixed points of the multivalued operator G.

Definition 2.10 A multivalued map $G : J \to P_{cl}(\mathbb{R})$ is said to be measurable if for every $y \in \mathbb{R}$, the function

$$t \to d(y, G(t)) = \inf\{|y - z| : z \in G(t)\}$$

is measurable.

The following relationship between upper semi-continuous maps and closed graphs is well known.

Lemma 2.3 ([18]) *Let G be a completely continuous multivalued map with nonempty compact values. Then G is u.s.c. if and only if G has a closed graph.*

Definition 2.11 A multivalued map $F : I \times \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is said to be Carathéodory if:

(1) $t \to F(t, u)$ is measurable for each $u \in \mathbb{R}$;

(2) $u \to F(t, u)$ is upper semicontinuous for almost all $t \in I$.

Moreover, F is said to be L^1-Carathéodory if (1), (2), and the following condition hold:

(3) For each $q > 0$, there exists $\varphi_q \in L^1(I, \mathbb{R}_+)$ such that

$$\|F(t, u)\|_{\mathcal{P}} = \sup\{|v| : v \in F(t, u)\} \le \varphi_q \text{ for all } |u| \le q \text{ and for a.e. } t \in I.$$

For each $u \in C(I, \mathbb{R})$, we define the set of selections of F by

$$S_{F\circ u} = \{v \in L^1(I, \mathbb{R}) : v(t) \in F(t, u(t)) \text{ a.e. } t \in I\}.$$

Let (X, d) be a metric space induced from the normed space $(X, |\cdot|)$. The function $H_d : \mathcal{P}(X) \times \mathcal{P}(X) \to \mathbb{R}_+ \cup \{\infty\}$ given by

$$H_d(A, B) = \max\{\sup_{a\in A} d(a, B), \sup_{b\in B} d(A, b)\}$$

is known as the Hausdorff–Pompeiu metric. For more details on multivalued maps see the books of Hu and Papageorgiou [18].

In the sequel, we need the following fixed point theorem.

Theorem 2.1 (Bohnenblust-Karlin [13]) *Let X be a Banach space and $K \in \mathcal{P}_{cl,cv}(X)$, and suppose that the operator $G : K \to \mathcal{P}_{cl,cv}(K)$ is upper semicontinuous and the set $G(K)$ is relatively compact in X. Then G has a fixed point in K.*

3 Main Results

We begin by defining what we mean by a solution, an upper solution, and a lower solution to our problem.

Definition 3.1 A function $u \in C(I)$ is said to be a solution of (1)–(2) if there exists a function $f \in S_{F\circ u}$ such that ${}^C D_q^\alpha u(t) = f(t)$ a.e. $t \in I$ and the boundary condition $L(u(0), u(T)) = 0$ is satisfied.

Definition 3.2 A function $w \in C(I)$ is said to be an upper solution of (1)–(2) if $L(w(0), w(T)) \geq 0$, and there exists a function $v_1 \in S_{F\circ w}$ such that ${}^C D_q^\alpha w(t) \geq v_1(t)$ a.e. $t \in I$. Similarly, a function $v \in C(I)$ is said to be a lower solution of (1)–(2) if $L(v(0), v(T)) \leq 0$, and there exists a function $v_2 \in S_{F\circ v}$ such that ${}^C D^\alpha v(t) \leq v_2(t)$ a.e. $t \in I$.

We now present the main result in this paper.

Theorem 3.1 *Assume that the following conditions hold:*

(H1) *$F : I \times \mathbb{R} \to \mathcal{P}_{cp,cv}(\mathbb{R})$ is Carathéodory;*

(H2) *There exist $v, w \in C(I)$, which are the lower and upper solutions, respectively, for problem (1)–(2) such that $v \leq w$;*

(H3) *The function $L(\cdot,\cdot)$ is continuous on $[u(0), w(0)]\times[u(T), w(T)]$ and is nonincreasing in each of its arguments;*

(H4) *There exists $l \in L^1(I, \mathbb{R}^+)$ such that*

$$H_d(F(t, u), F(t, \bar{u})) \leq l(t)|u - \bar{u}| \text{ for every } u, \bar{u} \in \mathbb{R},$$

and

$$d(0, F(t, 0)) \leq l(t) \text{ a.e. } t \in I.$$

Then the problem (1)–(2) has at least one solution u defined on I such that

$$v \leq u \leq w.$$

Proof. Consider the following modified problem

$$^{C}D_q^{\alpha}u(t) \in F(t, \tau(u(t))), \text{ for a.e. } t \in I, \tag{3}$$

$$u(0) = \tau(u(0)) - L(\overline{u}(0), \overline{u}(T)), \tag{4}$$

where

$$\tau(u(t)) = \max\{v(t), \min\{u(t), w(t)\}\},$$

and

$$\overline{u}(t) = \tau(u(t)).$$

A solution to (3)–(4) is a fixed point of the operator $N : C(I) \to \mathcal{P}(C(I))$ defined by

$$N(u) = \{h \in C(I) \ : \ h(t) = u(0) + (I_q^{\alpha}\nu)(t)\},$$

where

$$\nu \in \{x \in \widetilde{S}^1_{F\circ\tau(u)} : x(t) \geq v_1(t) \text{ on } A_1 \text{ and } x(t) \leq v_2(t) \text{ on } A_2\},$$

$$S^1_{F\circ\tau(y)} = \{x \in L^1(I) \ : x(t) \in F(t, (\tau u)(t)), \ a.e. \ t \in I\},$$

$$A_1 = \{t \in I : u(t) < v(t) \leq w(t)\}, \ A_2 = \{t \in I \ : v(t) \leq w(t) < u(t)\}.$$

Remark 3.1 (1) For each $u \in C(I)$, the set $\widetilde{S}^1_{F\circ\tau(u)}$ is nonempty. In fact, $(H1)$ implies that there exists $v_3 \in S^1_{F\circ\tau(u)}$, so we set

$$v = v_1\chi_{A_1} + v_2\chi_{A_2} + v_3\chi_{A_3},$$

where

$$A_3 = \{t \in I : v(t) \leq u(t) \leq w(t)\}.$$

Then, by decomposability, $x \in \widetilde{S}^1_{F\circ\tau(u)}$.

(2) From the definition of τ, it is clear that $F(\cdot, \tau u(\cdot))$ is an L^1-Carathéodory multi-valued map with compact convex values and there exists $\phi_1 \in C(I, \mathbb{R}^+)$ such that

$$\|F(t, \tau u(t))\|_{\mathcal{P}} \leq \phi_1(t) \text{ for each } u \in \mathbb{R}.$$

(3) Since $\tau(u(t)) = v(t)$ for $t \in A_1$, and $\tau(u(t)) = w(t)$ for $t \in A_2$, in view of (H3), equation (4) implies that

$$|u(0)| \leq |v(0)| + |L(v(0), v(T)| \leq |v(0)| + |L(u(0), u(T))| = |v(0)| \ on \ A_1,$$

and

$$u(1) = w(0) - L(w(0), w(T) \leq w(0) - L(u(0), u(T)) = w(0) \ on \ A_2.$$

Thus,

$$|u(0)| \leq \min\{|v(0)|, |w(0)|\}.$$

Now set

$$L := \sup_{t \in I} \int_0^t \frac{(t-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} d_q s,$$

let

$$R := \min\{|v(0)|, |w(0)|\} + L\|\phi_1\|_\infty,$$

and consider the closed and convex subset of $C(I)$ given by

$$B = \{u \in C(I) : \|u\|_\infty \le R\}.$$

We shall show that the operator $N : B \to \mathcal{P}_{cl,cv}(B)$ satisfies all the assumptions of Theorem 2.1. The proof will be given in steps.

Step 1: *$N(u)$ is convex for each $y \in B$.*
Let h_1, h_2 belong to $N(u)$; then there exist ν_1, $\nu_2 \in \widetilde{S}^1_{F\circ\tau(u)}$ such that, for each $t \in I$ and any $i = 1, 2$, we have

$$h_i(t) = u(0) + (I_q^\alpha \nu_i)(t).$$

Let $0 \le d \le 1$. Then, for each $t \in I$, we have

$$(dh_1 + (1-d)h_2)(t) = u(0) + \int_0^t \frac{(t-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} [d\nu_1(s) + (1-d)\nu_2(s)] d_q s.$$

Since $S_{F\circ\tau(u)}$ is convex (because F has convex values), we have

$$dh_1 + (1-d)h_2 \in N(u).$$

Step 2: *N maps bounded sets into bounded sets in B.*
For each $h \in N(u)$, there exists $\nu \in \widetilde{S}^1_{F\circ\tau(u)}$ such that

$$h(t) = u(0) + \int_0^t \frac{(t-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} \nu(s) d_q s.$$

From conditions (H1)–(H3), for each $t \in I$, we have

$$\begin{aligned} |h(t)| &\le |u(0)| + \left| \int_0^t \frac{(t-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} |\nu(s)| d_q s \right| \\ &\le \min\{|v(0)|, |w(0)|\} + \int_0^t \frac{(t-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} |\nu(s)| d_q s \\ &\le \min\{|v(0)|, |w(0)|\} + L\|\phi_1\|_\infty. \end{aligned}$$

Thus,

$$\|h\|_\infty \le R.$$

Step 3: *N maps bounded sets into equicontinuous sets of B.*

Let $t_1,\, t_2 \in I$ with $t_1 < t_2$, and let $u \in B$ and $h \in N(u)$. Then

$$\begin{aligned}
|h(t_2) - h(t_1)| &= \Bigg| \int_0^{t_1} \frac{|(t_2 - qs)^{(\alpha-1)} - (t_1 - qs)^{(\alpha-1)}|}{\Gamma_q(\alpha)} \nu(s) d_q s \\
&\quad + \int_{t_1}^{t_2} \frac{(t_2 - qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} \nu(s) d_q s \Bigg| \\
&\le \int_0^{t_1} \frac{|(t_2 - qs)^{(\alpha-1)} - (t_1 - qs)^{(\alpha-1)}|}{\Gamma_q(\alpha)} |\nu(s)| d_q s \\
&\quad + \int_{t_1}^{t_2} \frac{|(t_2 - qs)^{(\alpha-1)}|}{\Gamma_q(\alpha)} |\nu(s)| d_q s \\
&\le \|\phi_1\|_\infty \int_0^{t_1} \frac{|(t_2 - qs)^{(\alpha-1)} - (t_1 - qs)^{(\alpha-1)}|}{\Gamma_q(\alpha)} d_q s \\
&\quad + \|\phi_1\|_\infty \int_{t_1}^{t_2} \frac{|(t_2 - qs)^{(\alpha-1)}|}{\Gamma_q(\alpha)} d_q s \\
&\to 0 \ as\ t_1 \to t_2.
\end{aligned}$$

As a consequence of the three steps above, we can conclude from the Arzelà-Ascoli theorem that $N : C(I) \to \mathcal{P}(C(I))$ is continuous and completely continuous.

Step 4: *N has a closed graph.*
Let $u_n \to u_*$, $h_n \in N(u_n)$, and $h_n \to h_*$. We need to show that $h_* \in N(u_*)$. Now $h_n \in N(u_n)$ implies there exists $\nu_n \in \widetilde{S}^1_{F\circ\tau(u_n)}$ such that, for each $t \in I$,

$$h_n(t) = u(0) + \int_0^t \frac{(t - qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} \nu_n(s) d_q s.$$

We must show that there exists $\nu_* \in \widetilde{S}^1_{F\circ\tau(u_*)}$ such that, for each $t \in I$,

$$h_*(t) = u(0) + \int_0^t \frac{(t - qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} \nu_*(s) d_q s.$$

Since $F(t, \cdot)$ is upper semi-continuous, for every $\epsilon > 0$, there exists a natural number $n_0(\epsilon)$ such that, for every $n \ge n_0(\epsilon)$, we have

$$\nu_n(t) \in F(t, \tau u_n(t)) \subset F(t, u_*(t)) + \epsilon B(0, 1) \quad \text{a.e. } t \in I.$$

Since $F(\cdot, \cdot)$ has compact values, there exists a subsequence $\nu_{n_m}(\cdot)$ such that

$$\nu_{n_m}(\cdot) \to \nu_*(\cdot) \quad \text{as} \quad m \to \infty,$$

and

$$\nu_*(t) \in F(t, \tau u_*(t)) \quad \text{a.e. } t \in I.$$

For every $w \in F(t, \tau u_*(t))$, we have

$$|\nu_{n_m}(t) - \nu_*(t)| \le |\nu_{n_m}(t) - w| + |w - \nu_*(t)|.$$

Hence,

$$|\nu_{n_m}(t) - \nu_*(t)| \le d(\nu_{n_m}(t), F(t, \tau u_*(t))).$$

We obtain an analogous relation by interchanging the roles of v_{n_m} and v_* to obtain

$$|\nu_{n_m}(t) - \nu_*(t)| \leq H_d(F(t, \tau u_n(t)), F(t, \tau u_*(t))) \leq l(t)\|y_n - y_*\|_\infty.$$

Thus,

$$\begin{aligned} |h_{n_m}(t) - h_*(t)| &\leq \int_0^t \frac{|(t-qs)^{(\alpha-1)}|}{\Gamma_q(\alpha)} |\nu_{n_m}(s) - \nu_*(s)| d_q s \\ &\leq \|u_{n_m} - u_*\|_\infty \int_0^t \frac{|(t-qs)^{(\alpha-1)}|}{\Gamma_q(\alpha)} l(s) d_q s. \end{aligned}$$

Therefore,

$$\|h_{n_m} - h_*\|_\infty \leq \|u_{n_m} - u_*\|_\infty \int_0^t \frac{|(t_1-qs)^{(\alpha-1)}|}{\Gamma_q(\alpha)} l(s) d_q s \to 0 \quad \text{as } m \to \infty,$$

so Lemma 2.3 implies that N is upper semicontinuous.

Step 5: *Every solution u of (3)–(4) satisfies $v(t) \leq u(t) \leq w(t)$ for all $t \in I$.*
Let u be a solution of (3)–(4). To prove that $v(t) \leq u(t)$ for all $t \in I$, suppose this is not the case. Then there exist t_1, t_2, with $t_1 < t_2$, such that $v(t_1) = u(t_1)$ and $v(t) > u(t)$ for all $t \in (t_1, t_2)$. In view of the definition of τ,

$${}^C D_q^\alpha u(t) \in F(t, v(t)) \text{ for all } t \in (t_1, t_2).$$

Thus, there exists $y \in S_{F\circ\tau(v)}$ with $y(t) \geq v_1(t)$ a.e. on (t_1, t_2) such that

$${}^C D_q^\alpha u(t) = y(t) \text{ for all } t \in (t_1, t_2).$$

An integration on $(t_1, t]$, with $t \in (t_1, t_2)$, yields

$$u(t) - y(t_1) = \int_{t_1}^t \frac{(t-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} \nu(s) d_q.$$

Since v is a lower solution of (1)–(2),

$$v(t) - v(t_1) \leq \int_{t_1}^t \frac{(t-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} v_1(s) d_q, \ t \in (t_1, t_2).$$

From the facts that $u(t_0) = v(t_0)$ and $\nu(t) \geq v_1(t)$, it follows that

$$v(t) \leq u(t) \text{ for all } t \in (t_1, t_2).$$

This is a contradiction, since $v(t) > u(t)$ for all $t \in (t_1, t_2)$. Consequently,

$$v(t) \leq u(t) \text{ for all } t \in I.$$

Similarly, we can prove that

$$u(t) \leq w(t) \text{ for all } t \in I.$$

This shows that

$$v(t) \leq u(t) \leq w(t) \text{ for all } t \in I.$$

Therefore, the problem (3)–(4) has a solution u satisfying $v \leq u \leq w$.

Step 6: *Every solution of problem (3)–(4) is a solution of (1)–(2).* Suppose that u is a solution of the problem (3)–(4). Then, we have

$$^{C}D_q^{\alpha}u(t) \in F(t, \tau(u(t))) \text{ for a.e. } t \in I,$$

and

$$u(0) = \tau(u(0)) - L(\overline{u}(0), \overline{u}(T)).$$

Since, for all $t \in I$, we have $v(t) \leq u(t) \leq w(t)$, it follows that $\tau(u(t)) = u(t)$. Thus, we have

$$^{C}D_q^{\alpha}u(t) \in F(t, u(t)) \text{ for a.e. } t \in I,$$

and $L(u(0), u(T)) = 0$. We only need to prove that

$$v(0) \leq u(0) - L(u(0), u(T)) \leq w(1),$$

so suppose that

$$u(0) - L(u(0), u(T)) < u(0).$$

Since $L(v(0), v(T)) \leq 0$, we have

$$u(0) \leq u(1) - L(v(0), v(T)),$$

and since $L(\cdot, \cdot)$ is nonincreasing with respect to both of its arguments,

$$u(0) \leq u(0) - L(v(0), v(T)) \leq u(0) - L(u(0), u(T)) < v(0).$$

Hence, $u(0) < v(0)$, which is a contradiction. Similarly, we can prove that

$$u(0) - L(u(0), u(T)) \leq w(1).$$

Thus, u is a solution of (1)–(2).

This shows that the problem (1)–(2) has a solution u satisfying $v \leq u \leq w$, and completes the proof of the theorem.

Remark 3.2 In the case where $L(x, y) = ax - by - c$, Theorem 3.1 yields existence results to the problem

$$^{C}D_q^{\alpha}u(t) \in F(t, u(t)) \text{ for a.e. } \; t \in I, \tag{5}$$

$$ay(1) - by(T) = c, \tag{6}$$

where $-b < a \leq 0 \leq b$, $c \in \mathbb{R}$, which includes the anti-periodic problem $b = -a$, $c = 0$, the initial value problem, and the terminal value problem.

4 An Example

Consider the following problem of a Caputo fractional $\frac{1}{4}-$difference inclusion of order $\alpha = \frac{1}{2}$,

$$\begin{cases} \left(^{c}D_{\frac{1}{4}}^{\frac{1}{2}}u\right)(t) \in \frac{7t^2}{27(1+|u(t)|)}[u(t), 33(1 + u(t))], \; t \in [0, 1], \\ \\ u(0) + u(1) = 1. \end{cases} \tag{7}$$

Set
$$F(t,u(t)) = \frac{7t^2}{27(1+|u(t)|)}[u(t), 33(1+u(t))],\ t \in [0,1],$$
and $L(x,y) = -x-y+1$ for $x,\, y \in \mathbb{R}$. It is easy to see that $F : [0,1]\times\mathbb{R} \to \mathcal{P}_{cp,cv}(\mathbb{R})$ is Carathéodory.

In order to see that $(H2)$ holds, let $v,\ w \in C([0,1],\mathbb{R})$ be given by $v(t) = t^{\frac{5}{2}}$ and $w(t) = t^{\frac{3}{2}}$. Now $L(v(0), v(1)) = 0 \le 0$ and
$$\left({}^cD^{\frac{1}{2}}_{\frac{1}{4}}v\right)(t) = \frac{217}{27}t^2 \le \frac{7t^2}{27(1+v(t))}(31+31v(t)) \in F(t,v(t)).$$
Also, $L(w(0), w(1)) = 0 \ge 0$ and
$$\left({}^cD^{\frac{1}{2}}_{\frac{1}{4}}w\right)(t) = \frac{7}{9}t \ge \frac{7}{9}t^2 = \frac{7t^2}{27(1+w(t))}(3+3w(t)) \in F(t,w(t)).$$
Therefore, v and w are lower and upper solutions, respectively, for problem (7) with $v \le w$. To see that condition $(H3)$ is satisfied, note that L is continuous and
$$\frac{\partial L(x,y)}{\partial x} = \frac{\partial L(x,y)}{\partial y} = -1 < 0.$$
Finally, for each $u,\ \bar{u} \in \mathbb{R}$ and $t \in [0,1]$, we have
$$H_d(F(t,u), F(t,\bar{u})) \le \frac{7}{27}t^2|u-\bar{u}| \quad \text{and} \quad d(0, F(t,0)) = \|F(t,0)\|_{\mathcal{P}} \le \frac{7}{27}t^2,$$
so $(H4)$ is satisfied with $l(t) = \frac{7}{27}t^2$.

Consequently, all conditions of Theorem 3.1 are satisfied, and so problem (7) has at least one solution u defined on $[0,1]$ with $t^2\sqrt{t} \le u(t) \le t\sqrt{t}$.

5 Concluding Remarks

In this paper the authors study the existence of solutions to a boundary value problem for a fractional q-difference inclusion involving the Caputo fractional derivative. This topic fits well in the scope of problems covered by the journal *Nonlinear Dyamics and Systems Theory*.

This paper is the first attempt at using the method of upper and lower solutions to study problems of this type. In order to illustrate the applicability of the results, an example is given detailing how the various hypotheses are satisfied.

References

[1] S. Abbas and M. Benchohra. Upper and lower solutions method for Darboux problem for fractional order implicit impulsive partial hyperbolic differential equations. *Acta Univ. Palacki. Olomuc.* **51** (2012) 5–18.

[2] S. Abbas and M. Benchohra. The method of upper and lower solutions for partial hyperbolic fractional order differential inclusions with impulses. *Discuss. Math. Differ. Incl. Control Optim.* **30** (2010) 141–161.

[3] S. Abbas, M. Benchohra, S. Hamani and J. Henderson. Upper and lower solutions method for Caputo–Hadamard fractional differential inclusions. *Math. Moravica.* **23** (2019) 107–118.

[4] S. Abbas, M. Benchohra and A. Hammoudi. Upper, lower solutions method and extremal solutions for impulsive discontinuous partial fractional differential inclusions. *Panamer. Math. J.* **24** (2014) 31–52.

[5] S. Abbas, M. Benchohra and J. J. Trujillo. Upper and lower solutions method for partial fractional differential inclusions with not instantaneous impulses. *Prog. Frac. Diff. Appl.* **1** (2015) 11–22.

[6] C. R. Adams. On the linear ordinary q-difference equation. *Annals Math.* **30** (1928) 195–205.

[7] R. Agarwal. Certain fractional q-integrals and q-derivatives. *Proc. Camb. Philos. Soc.* **66** (1969) 365–370.

[8] B. Ahmad. Boundary value problem for nonlinear third order q-difference equations. *Electron. J. Differential Equations* **2011** (94) (2011) 1–7.

[9] B. Ahmad, S. K. Ntouyas and L. K. Purnaras. Existence results for nonlocal boundary value problems of nonlinear fractional q-difference equations. *Adv. Difference Equ.* **2012**, 2012:140.

[10] M. H. Annaby and Z. S. Mansour. *q-fractional Calculus and Equations.* Lecture Notes in Mathematics, vol. 2056. Springer, Heidelberg, 2012.

[11] M. Benchohra and S. Hamani. The method of upper and lower solution and impulsive fractional differential inclusions. *Nonlinear Anal. Hybrid Syst.* **3** (2009) 433–440.

[12] M. Benchohra and S. K. Ntouyas. The lower and upper method for first order differential inclusions with nonlinear boundary condition. *J. Inequ. Pure Appl. Math.* **3** (2002), Article 14, 1–20.

[13] H. F. Bohnenblust and S. Karlin. On a theorem of Ville. Contribution on the theory of games. In: *Annals of Mathematics Studies*, Vol. 24. Princeton University Press, Princeton, 1950, 155–160.

[14] R. D. Carmichael. The general theory of linear q-difference equations. *American J. Math.* **34** (1912) 147–168.

[15] M. El-Shahed and H. A. Hassan. Positive solutions of q-difference equation. *Proc. Amer. Math. Soc.* **138** (2010) 1733–1738.

[16] T. Ernst. *A Comprehensive Treatment of q-Calculus.* Birkhäuser, Basel, 2012.

[17] S. Etemad, S. K. Ntouyas and B. Ahmad. Existence theory for a fractional q-integro-difference equation with q-integral boundary conditions of different orders. *Mathematics* **7** 659 (2019) 1–15.

[18] Sh. Hu and N. Papageorgiou. *Handbook of Multivalued Analysis, Volume I: Theory.* Kluwer, Dordrecht, 1997.

[19] V. Kac and P. Cheung. *Quantum Calculus.* Springer, New York, 2002.

[20] P. M. Rajkovic, S. D. Marinkovic and M. S. Stankovic. Fractional integrals and derivatives in q-calculus. *Appl. Anal. Discrete Math.* **1** (2007) 311–323.

[21] P. M. Rajkovic, S. D. Marinkovic and M. S. Stankovic. On q-analogues of Caputo derivative and Mittag-Leffler function. *Fract. Calc. Appl. Anal.* **10** (2007) 359–373.

Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 13–26

Stopping Rules for Selecting the Optimal Subset

Mohammad H. Almomani [1]* and Mahmoud H. Alrefaei [2] and Mohanad H. Almomani

[1] *Department of Mathematics, Faculty of Science, The Hashemite University, P.O. box 330127, Zarqa 13133, JORDAN.*
[2] *Department of Mathematics and Statistics, Jordan University of Science and Technology, Irbid, JORDAN.*
[3] *The Ministry of Education, Ajloun, JORDAN.*

Received: September 25, 2020; Revised: December 26, 2020

Abstract: Selecting the best of a finite set of alternatives is a very important area of research. In this paper, we discuss the stopping rules of the procedure of selecting the optimum subset out of a very large alternative dynamic system. A combined procedure with two stages is studied. The first stage employs the ordinal optimization to select a subset that overlaps with the set of actual best $k\%$ designs with high probability. After that, the optimal computing budget allocation is used in the second stage to select the best m designs from the selected subset. The efficiency of selection procedures with two different stopping rules is studied by implementing them on two test problems to see the efficiency of the procedure in the context of the most effective stopping rule. The first problem is a generic example and the second one is a buffer allocation problem.

Keywords: *large scale problems; simulation optimization; ordinal optimization; stopping rules; optimal computing budget allocation.*

1 Introduction

Statistical selection procedures are designed to answer the question "*which treatment can be considered the best?*", where the *best* refers to the design that has the maximum or minimum expected performance measure. Different sampling assumptions, approximations, parameters and stopping rules were combined to define a procedure. Due to the increasing demands that are being placed upon simulation optimization algorithms together with having many differences between the statistical selection procedures, it is getting important to find out which of these procedures is the most convenient one to

* Corresponding author: mailto:mh_momani@hu.edu.jo

use. However, evaluating selection procedures can be done in several ways, including the theoretical, empirical, and practical perspectives.

Simulation is often used by managers to help make a decision by exploring different options. Creating different physical models and conducting experiments on them in order to choose the best model usually cost a lot. Simulation is a relatively cheap alternative for collecting information, since it provides some estimates about the performance of a system that does not exist physically. The experiments are conducted as statistical experiments; the number of times the simulation runs is determined, the experiments are performed, and the output data are analyzed. A stopping rule that decides when the experiment ends -i. e., how many replicates of the simulation are made - must be chosen by the manager. This paper studies how different stopping rules affect the simulation output analysis.

Selection procedures used simulation to estimate the performance measure. Since the simulation methods are used to indicate the performance measure for each alternative, an incorrect selection is possible. Thus, some measures are needed to determine the selection quality. Two measures of selection quality exist; the first one is the Probability of Correct Selection ($P(CS)$) and the second measure is the Expected Opportunity Cost ($E(OC)$) of a potentially incorrect selection, see He et al. [1]. Traditional selection procedures identify the best system with high probability of correct selection, by maximizing the $P(CS)$. However, the $E(OC)$ is applied in business, engineering and many other applications. This led recently to a new selection procedure that reduces the cost of a potentially incorrect selection.

The measures of selection quality can be used to decide when to stop the sampling process. In particular, Brank et al. [2] proposed the following stopping rules:

1. Sequential (S): Repeat sampling while $\sum_{i=1}^{n} T_i < T$, for some specified total budget T and T_i is the number of samples allocated to design i, where $i = 1, 2, \ldots, n$.

2. Expected opportunity cost (EOC): Repeat sampling while $E(OC) > \varepsilon$, for a specified expected opportunity cost target $\varepsilon > 0$.

3. Probability of good selection ($P(GS)_{\delta^*}$): Repeat sampling while $P(GS)_{\delta^*} < 1 - \varphi^*$, for a specified probability target $1 - \varphi^* \in [1/n, 1)$ and given $\delta^* \geq 0$. In the Indifference Zone (IZ) procedure, see Bechhofer et al. [3], δ^* is the difference between the favorable design and the best design and is called the indifference zone. It represents the smallest difference that one wants to achieve.

In many practical problems, selecting a set of m best solutions out of n solutions is more convenient than selecting only one solution. This is done based on the simulation output from each design. In case of having a small size of the feasible solution set, the best design or a subset of the best designs can be selected using Ranking and Selection procedures, see Bechhofer et al. [3], Law and Kelton [4], and Kim and Nelson [5], [6]. Ranking and Selection procedures for large alternatives require a very big computational time. Thus, such procedures might not be feasible for large scale problems. For comprehensive reviews of the Ranking and Selection procedures, see Gibbons et al. [7] and Gupta and Panchapakesan [8].

The Ordinal Optimization (OO) that was proposed by Ho et al. [9] relaxes the objective to finding good enough designs, rather than estimating the performance of the designs accurately. In fact, the OO procedure seeks to isolate a subset of solutions with high goodness probability. After that, the optimal solution(s) can be located from the isolated set by using any simulation optimization procedure. Previously, many procedures have been proposed to select a good design (designs) to solve this selection problem, see for example Alrefaei and Almomani [10], Almomani and Alrefaei [11], Almomani and Abdul Rahman [12], Al-Salem et al. [13], Almomani et al. [14], Almomani et al. [15], Alrefaei et al. [16].

Consider the problem of distributing an available budget computation in simulating the different solutions. Instead of distributing this budget evenly on different alternatives, the available budget can be distributed in a way that maximizes the probability of correct selecting the good solutions.Therefore, the idea of the Optimal Computing Budget Allocation $OCBA_m$ has been proposed by Chen et al. [17] for selecting the best m designs.

Recently, a sequential selection procedure was considered by Almomani and Alrefaei [18] for selecting a good subset of solutions from a large size problem. The procedure combines the $OCBA_m$ and the OO procedures. In the first stage, an isolation of a subset of good enough designs with high probability is made by the OO procedure. This reduces the feasible solution set size and makes it appropriate to apply the $OCBA_m$ procedure. In the second stage, a maximization problem, that seeks to maximize the probability of selecting all best m designs correctly from the subset found in the first stage, is formulated using the $OCBA_m$. A constraint on the total number of available simulation replications is considered for this maximization problem. The procedure starts by simulating each alternative in the set and considers by initial simulation replications of size t_0. After that, a fixed increment of replications Δ is added and distributed among all solutions in the set. The process is repeated until all available computations are consumed.

In this paper, the effect of the stopping rules is studied; we study two stopping rules and implement them in our proposed algorithm; these are the sequential S and the expected opportunity cost EOC. However, the third stopping rule that uses the probability of good selection $P(GS)_{\delta^*}$ is not applicable in the proposed algorithm. These stopping rules are tested and compared by different examples using different measures to study their effect on the final solution of the selection procedure. These rules give the chance to stop the procedure earlier whenever the evidence for correct selection is high enough, and allow for additional sampling when it is not, which gives a kind of flexibility. Furthermore, we apply a numerical illustration for this approach to display the advantages and the disadvantages for each stopping rule, and to determine the most effective stopping rule that works better with this selection procedure.

The rest of this paper is organized as follows. In Section 2, we provide the background of the problem statement, OO procedure, and $OCBA_m$ procedure. In Section 3, the sequential selection procedure is presented with two different stopping rules. The performance of the selection approach under these stopping rules is illustrated with a series of numerical examples in Section 4. Finally, Section 5 includes concluding remarks.

2 Background

2.1 Problem statement

Consider the following simulation optimization problem

$$\min_{\theta \in \Theta} Y(\theta), \tag{1}$$

where Θ is an arbitrary, large and finite feasible solution. Let $Y(\theta) = E[L(\theta, X)]$ be the expected performance measure in a specific complex stochastic design, where θ represents the system parameters as a vector, X represents the random effects on the system and L is a deterministic function of θ and X.

Simulation is used to infer the set of the best m designs, which are the designs with the m smallest means out of the n designs we have in the feasible solution set. Let Y_{ij} (observation) represent the j^{th} sample of $Y(i)$ for the design i. We assume that Y_{ij} are independent and identically distributed ($i.i.d.$) normally distributed with unknown means $Y_i = E(Y_{ij})$ and variances $\sigma_i^2 = Var(Y_{ij})$, i.e. $Y_{i1}, Y_{i2}, \ldots, Y_{iT_i}$ are $i.i.d.$ $N(Y_i, \sigma_i^2)$. There is no a problem with the normality assumption here since it holds for sure. This can be shown using the Central Limit Theorem, regarding that simulation outputs are acquired through an average performance or from batch means. In practice, we estimate the variance σ_i^2 using the sample variance s_i^2 for Y_{ij} because it is unknown. Our aim is to select a set, S_m, containing the best m designs. The word "best", in a minimization problem, refers to the one with the smallest sample mean. Define $\overline{Y}_{[r]}$ to be the r-th smallest (statistic order) of $\{\overline{Y}_1, \overline{Y}_2, \ldots, \overline{Y}_n\}$, i.e., $\overline{Y}_{[1]} \leq \overline{Y}_{[2]} \leq \ldots \leq \overline{Y}_{[n]}$, where $\overline{Y}_i = \frac{1}{T_i}\sum_{i=1}^{T_i} Y_{ij}$ is the sample mean for the design i. After that, let $S_m = \{[1], [2], \ldots, [m]\}$, which gives the correct selection that contains all of the m designs with smallest means, i.e., $CS_m = \{\max_{i \in S_m} \overline{Y}_i \leq \min_{i \notin S_m} \overline{Y}_i\}$.

2.2 Ordinal optimization

For large scale selection problems, the Ordinal Optimization (OO) procedure was proposed by Ho et al. [9]. Due to the high cost of accurate estimating the design performance values in the optimization process, it would be more practical to select a subset of the feasible set containing some of the best designs with high probability. This means that ordinal optimization is first used to isolate a subset of good enough designs, then cardinal optimization is applied on this isolated set. The main objective is to reduce the required simulation time for the discrete event simulation. A review of the OO procedure can be found in Ho et al. [19], Horng and Lin [20] and Ma et al. [21].

2.3 The selection procedure

Minimizing the total computational time for different designs in the simulation process is important to make the OO procedure more effective. Therefore, it is necessary to allocate the simulation samples cleverly, where a greater number of samples is applied to the designs that are more effective in identifying the best design. In this case, noncritical designs with smaller effect on discovering the best designs are not given much simulation samples. Chen et al. [22] have proposed the Optimal Computing Budget Allocation ($OCBA$) procedure which focuses on selecting the best design. On the other hand, an efficient allocation procedure, ($OCBA_m$), was also proposed by Chen et al. [17] for selecting the top m designs.

The problem is formulated by Chen et al. [17] so as to maximize the probability of selecting the best m designs $P(CS_m)$ correctly, subject to a constraint on the available number of samples. In mathematical notation, the problem can be written as

$$\begin{aligned} &\max_{T_1,\dots,T_n} \; P(CS_m) \\ &\text{s.t.} \; \sum_{i=1}^{n} T_i = T, \end{aligned} \tag{2}$$

where T is the number of available simulation samples, and for the design i we set T_i simulation samples. n is the total number of designs, m is the size of the optimal subset, S_m, that contains the best designs. The goal is to allocate the simulation samples we have in a way that maximizes the probability of the correct selection, $P(CS_m)$, given the total number of samples as $\sum_{i=1}^{n} T_i$. By this formulation, the computational cost of each sample is implicitly assumed to be constant across designs. Chen and Lee [23] suggested approximating $P(CS_m)$ by a lower bound of it, $APCS_m$, which determines an asymptotic approximation for the samples $T_i,\ i = 1, \dots, n$ that maximize $APCS_m$. The following theorem of Chen and Lee [23] gives the estimates of these $T_i's$.

Theorem 2.1 *Given a total number of simulation samples T to be distributed to n competing designs whose performance is represented by random variables with means $Y_1, Y_2, \dots, Y_n$, and finite variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, respectively. The Approximate Probability of Correct Selection for m best ($APCS_m$) as $T \longrightarrow \infty$ can be asymptotically maximized when $\frac{T_i}{T_j} = \left(\frac{\sigma_i/\delta_i}{\sigma_j/\delta_j}\right)^2$; for any $i, j \in \{1, 2, \dots, n\}$ and $i \neq j$, where $\delta_i = \bar{Y}_i - c$, for some constant c.*

3 The Selection Procedure with Stopping Rules

We consider the sequential selection procedure that consists of two stages, see Almomani and Alrefaei [18]. In the first stage, out of the search space, a subset G is selected randomly using the OO procedure, such that G overlaps the set containing the actual best $k\%$ designs with high probability. After that, in the second stage, the best m designs are identified from the subset G using the $OCBA_m$ procedure.

In this section, we present the sequential selection procedure for selecting the optimal subset, then we discuss the different stopping rules used to stop the procedure. These are the sequential, the expected opportunity cost and the probability of good selection.

3.1 A selection procedure for selecting the best m designs

The run length of a simulation experiment is often determined using sequential stopping rules. These rules can be used within confidence interval procedures for simulation output analysis. In the sequential stopping rule if the total number of samples is exceeded (i.e., if $\sum_{i=1}^{g} T_i^l \geq T$), then the algorithm stops. In fact, most traditional selection procedures use this stopping rule. Since the target is selecting the best design with minimum elapsed time, we can control this by increasing or decreasing the total budget T.

We first present the sequential algorithm proceedure and then we discuss the stopping rules used to stop the algorithm.

Setup: Determine the precision level p_0 and let $|G| = g$, where G is defined as the required subset from Θ, that satisfies $P(G$ contains at least m of the best $k\%$ designs$) \geq p_0$. Determine the number of initial simulation samples $t_0 \geq 5$.

Determine the total computing budget T, and the value of m (best top m). Let $l = 0$ and let $T_1^l = T_2^l = \ldots = T_g^l = t_0$, where l is the iteration number.

Select a subset G of g alternatives randomly from Θ. Take random samples of t_0 observations Y_{ij} $(1 \leq j \leq t_0)$ for each design i in G, where $i = 1, 2, \ldots, g$.

Initialization: For each $i \in G$, find an estimate of the sample mean and the sample variance as $\bar{Y}_i = \frac{1}{T_i^l} \sum_{j=1}^{T_i^l} Y_{ij}$ and $s_i^2 = \frac{1}{T_i^l - 1} \sum_{j=1}^{T_i^l} (Y_{ij} - \bar{Y}_i)^2$.

Order the sample means $\bar{Y}_{[1]} \geq \bar{Y}_{[2]} \geq \ldots \geq \bar{Y}_{[g]}$. Then select the set of top m designs, S_m.

Stopping Rule: Test the stopping rule, if it is satisfied, then stop, return S_m as the required subset. Otherwise, select randomly a subset S_z of $g - m$ alternatives from $\Theta - S_m$. Take random samples of t_0 observations Y_{ij} $(1 \leq j \leq t_0)$ for each design i in S_z. Compute $\bar{Y}_i$ and s_i as in the **Initialization** step above. Let $G = S_m \bigcup S_z$.

Simulation Budget Allocation: Increase the computing budget by Δ and compute the new budget allocation, $T_1^{l+1}, T_2^{l+1}, \ldots, T_g^{l+1}$ using $\frac{T_1^{l+1}}{\left(\frac{s_1}{\delta_1}\right)^2} = \frac{T_2^{l+1}}{\left(\frac{s_2}{\delta_2}\right)^2} = \cdots = \frac{T_g^{l+1}}{\left(\frac{s_g}{\delta_g}\right)^2}$, where $\delta_i = \bar{Y}_i - c$ and $c = \frac{\hat{\sigma}_{i_{m+1}} \bar{Y}_{i_m} + \hat{\sigma}_{i_m} \bar{Y}_{i_{m+1}}}{\hat{\sigma}_{i_m} + \hat{\sigma}_{i_{m+1}}}$ with $\hat{\sigma}_i = s_i / \sqrt{T_i^l}$, for all $i = 1, 2, \ldots, g$, see Chen and Lee [23].

Perform additional $\max\{0, T_i^{l+1} - T_i^l\}$ simulations samples for each designs i, where $i = 1, 2, \ldots, g$, let $l \longleftarrow l + 1$. Go to **Initialization**.

We consider two **Stopping Rules** for the proposed algorithm.

3.2 The sequential cost stopping rule

In this rule, a total number of samples T is predetermined. If the number of samples used in the algorithm reaches T, stop the algorithm, otherwise, continue. Therefore, the stopping rule becomes:

Stopping Rule: If $\sum_{i=1}^{g} T_i^l \leq T$, for a specified total number of samples T, then stop. Otherwise, proceed.

3.3 The expected opportunity cost stopping rule

This rule uses the expected opportunity cost EOC to stop the algorithm. In fact, the EOC stopping rule is recommended when the goal is to select the best design with the minimum $E(OC)$, especially, in business applications. Therefore, the stopping rule becomes:

Stopping Rule: If $E(OC) \leq \varepsilon$, for a specified expected opportunity cost significant $\varepsilon > 0$, then stop. Otherwise, proceed.

As we stated before, there is another stopping rule that was used in the literature but it is not applicable in our algorithm. This stopping rule is called the probability of good selection stopping rule. A selected design within δ^* from the best design is called the "good" design. However, since the objective in this paper is to select good enough designs from the actual best designs, but we are not concerned with the difference between the selected design and the actual best design(s), this stopping rule is not applicable. For more details about these three stopping rules, please, refer to Brank et al. [2] who provide an illustration about the difference between these three stopping rules through $KN++$ procedure, see also Goldsman et al. [24].

4 Numerical Examples

In this section, we present two numerical examples: a generic monotone increasing mean example and a queuing model example. In both examples, the algorithm is applied under different experiment settings and different stopping rules.

4.1 Example 1

Consider n different designs, each one is normally distributed $N(\mu_i, \sigma^2)$ with mean μ_i and variance σ^2 for $i = 1, 2, \ldots, n$. Such problem is called the Monotone Increasing Mean (MIM), which aims to find the best m designs with the minimum mean. If we let Θ be the feasible solution set, then $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ and $f(\theta_i)$ represents the index of θ_i in the feasible region, the optimization problem is

$$\min_{i=1,\ldots,n} f(\theta_i). \tag{3}$$

The proposed algorithm is applied on this example where $n = 1,000$, using the proposed selection procedure by implementing the first two stopping rules. It is assumed that for each $i \in \Theta, \mu_i = 10 + \frac{(i-1)}{10}$ and variance $\sigma_i^2 = 1$. Let $\theta_{[1]}, \theta_{[2]}, \ldots, \theta_{[n]}$ be the order of alternatives and we seek to select a subset of $m = 5$ solutions from the best 10% designs that have minimum means. If the selected design is in $\{\theta_{[1]}, \theta_{[2]}, \ldots, \theta_{[100]}\}$, then it is considered as a correct selection. The selection algorithm is applied using $g = 100$ solutions in G in order to study the effect of the simulation parameters, such as t_0 and Δ, on the performance of the algorithm. Furthermore, to achieve the normality assumption we use multiple replications method, where the number of multiple replications for each alternative equals M.

In the first experiment, we implement the proposed algorithm using the first stopping rule- the sequential S stopping rule. Here $n = 1,000$, $g = 100$, $k\% = 10\%$, $\Delta = 40$, $t_0 = 10$ and the total budget $T = 10,000$ (these settings are chosen arbitrary). In the second experiment, we implement the algorithm using the second stopping rule- the expected opportunity cost stopping rule with the same parameters setting as in the first experiment. The total budget condition is removed and replaced with the expected opportunity cost condition such that $E(OC) \leq 0.05$ (i.e., the significance level $\varepsilon = 0.05$). Table 1 contains the average performance of the algorithm over 100 replications for selecting 5 of the best 10% designs, for the first and the second experiment. In Table 1, $\overline{T}$ represents the average total sample size used in the algorithm $\sum_{i=1}^{g} T_i$ over the 100 replications. $\overline{P}$ represents the average probability of correctly selecting the best m designs; $P(CS_m)$ over the 100 replications, $\overline{E}$ represents the average expected opportunity cost for selecting the best m designs, $E(OC_m)$ over the 100 replications.

From Table 1, we note that the performance of the algorithm under the two stopping rules, the sequential S and the expected opportunity cost EOC, are almost the same with a preference of EOC on S when the measure is the expected opportunity cost used.

Table 1: The performance of the proposed procedure under different stopping rules for $n = 1000, g = 100, \Delta = 40, t_0 = 10, k\% = 10\%$ over 100 replications.

	Sequential Stopping Rule	EOC Stopping Rule
$\overline{T}$	6913	6995
$\overline{P}$	93%	93%
$\overline{E}$	0.010549	0.000523

To see the effect of the two stopping rules over different values of simulation budget on the probability of correct selection $P(CS_m)$, the results are depicted in Figure 1. It is clear from Figure 1 that the proposed algorithm with these two stopping rules produces a high $P(CS_m)$ quickly. Moreover, the performance of the two stopping rules is almost the same.

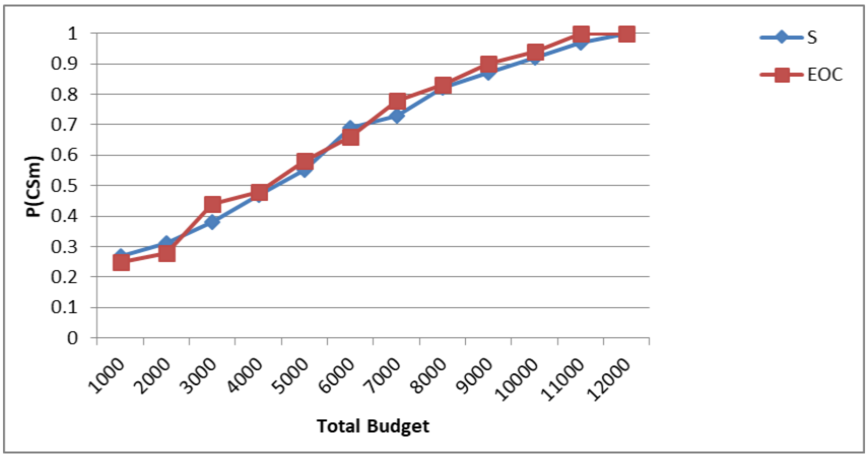


Figure 1: Comparison of the $P(CS_m)$ of S and EOC stopping rules over T budget for (MIM).

To study the effect of the two stopping rules on the proposed algorithm from the prospective of $E(OC_m)$, the results are depicted in Figure 2. Figure 2 shows the $E(OC_m)$ for the proposed algorithm using these two stopping rules over different values of simulation budget. From Figure 2, it is clear that the second stopping rule that uses the EOC gives better performance over the sequential procedure and that $E(OC_m)$ becomes close to 0 using this stopping rule under reasonable number of samples.

To enhance the performance of the algorithm, we increase the number of samples used in the multiple replication simulation method in order to get better estimates of the sample mean. Figure 3 shows the performance of the algorithm on the $E(OC_m)$ performance measure using the two stopping rules. Here M represents the samples used for each alternative. Obviously, it shows that the increase in the number of multiple replications M decreases the $E(OC_m)$ in both of the stopping rules. Moreover, the algorithm gives better performance when the EOC stopping rule is used. This is because when we increase the value of M we get better estimate value of the mean, therefore,

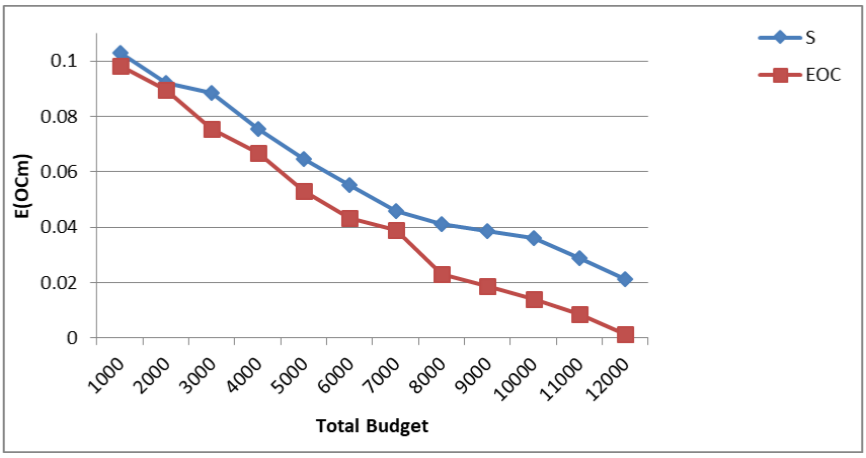


Figure 2: Comparison of the $E(OC_m)$ of S and EOC stopping rules over T budget for (MIM).

the difference between the estimated mean and the actual mean will be very small. In particular, when M is increased, then the $E(OC_m)$ approaches zero.

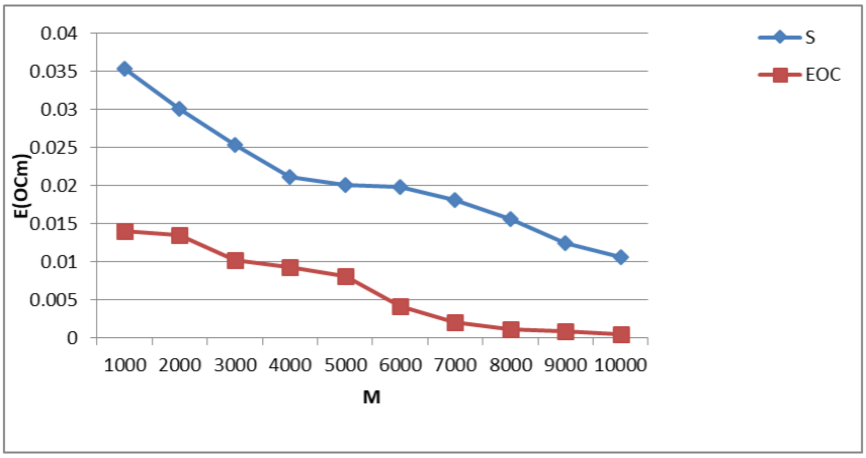


Figure 3: Comparison of the $E(OC_m)$ of S and EOC stopping rules over M replications for (MIM).

4.2 The buffer allocation problem (BAP)

We consider the Buffer Allocation Problem (BAP). The BAP consists of $q+1$ machines and q intermediate buffers in between. The question is how to distribute the available Q buffer slots over the q buffers in a way that meets a specific purpose.

Each station is modeled as a single server queuing model with $q + 1$ machines $M_0, M_1, \ldots, M_q$ and q intermediate buffers $B_1, B_2, \ldots, B_q$ in a production line as shown in Figure 4. Suppose that there are limits neither on the jobs in front of machine M_0, nor on the space for completed jobs after machine M_q. The service time at each machine is assumed to be independent and exponentially distributed random variable with rate $\mu_i,\ i = 1, 2, \ldots, q$. After the service is finished in machine M_i, the job tries to enter the queue of machine $M_{(i+1)}$. This cannot be done if the queue is full, and this prevents machine M_i from receiving new jobs to serve until the current job leaves it. Our goal is to maximize the production rate (throughput) by allocating the available spaces optimally on the intermediate buffers. Let Θ be the solution set, then Θ contains $\binom{Q+q-1}{Q}$ different solutions, see Almomani et al. [25], Papadopoulos et al. [26], Yuzukirmizia and Smith [27] and Alrefaei and Andradóttir [28].

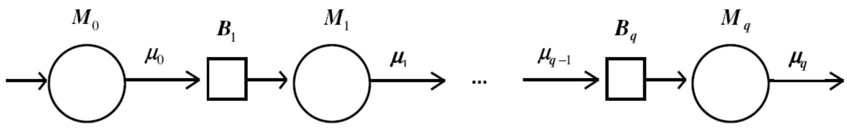


Figure 4: A production line with $q+1$ machines, q buffers, no limits either on the jobs in front of the machine M_0, or on the space for completed jobs after the machine M_q

The proposed algorithm is applied here on the specific type of BAP, with two stopping rules under some assumptions. In fact, there are different classifications for BAP problems. The first one is according to the length of the production line, which was presented by Papadopoulos et al. [26]. A production line is considered as "short" if the number of the machines is up to 6 with no more than 20 buffer spaces. Otherwise, the line is "large". Another point of view defines the BAP according to whether it has a balanced line, with equal mean service time at each machine, or an unbalanced one. Moreover, production lines can be seen as reliable (no machine fails) or unreliable. For more information about these classifications, see Almomani et al. [25].

Suppose that there are $Q = 15$ slots to be allocated over $q = 5$ buffers. Thus, we have 6 workstation and Θ contains 3,876 different designs ($n = 3,876$). In addition, assume that $\mu_0 = \mu_1 = \mu_2 = \mu_3 = 5$ and $\mu_4 = \mu_5 = 10$, which means that we assumed an unbalanced production line in this example. Furthermore, let the size of the set G be $g = 80$, the number of initial simulation samples $t_0 = 20$, the total computing budget $T = 100,000$, and the increment in simulation samples $\Delta = 50$. Suppose that our objective is to select the two design from the best 5% designs of Θ. This means that the correct selection here is to select the 2 designs that belong to the set $\{\theta_{[1]}, \theta_{[2]}, \ldots, \theta_{[193]}\}$, where $\theta_{[i]}, i = 1, 2, \ldots, 193$ represents the set of the actual top 5% designs with the maximum throughput in the set Θ.

We apply the proposed algorithm using the two stopping rules, the sequential S and the expected opportunity cost EOC rules. The experiment is repeated for 100 replications and the results are summarized in Table 2. In Table 2, $\overline{T}$ represents the average total sample size used in the algorithm $\sum_{i=1}^{g} T_i$ over the 100 replications. $\overline{P}$ represents the average probability of correct selecting the best m designs; $P(CS_m)$ over the 100 replications, $\overline{E}$ represents the average expected opportunity cost for selecting the best m designs, $E(OC_m)$ over the 100 replications. Clearly, the proposed algorithm selected the best buffer profile with high $P(CS_m)$ and small $E(OC_m)$. At the same time, there is a relatively small number of simulation samples needed.

Table 2: The performance of the proposed procedure under different stopping rules for $n = 3876, g = 80, \Delta = 50, t_0 = 20, k\% = 5\%$ over 100 replications.

	The S Stopping Rule	The EOC Stopping Rule
$\overline{T}$	241334	245921
$\overline{P}$	85%	89%
$\overline{E}$	0.002032	0.0000876

Figure 5 shows the average $P(CS_m)$ for selecting the 2 designs of the best 5% designs, with the two stopping rules, the sequential S and the expected opportunity cost EOC over different values of simulation budget. It is clear that the proposed algorithm selects the best designs with hight $P(CS_m)$ for the two stopping rules S and EOC and the two stopping rules give almost the same results.

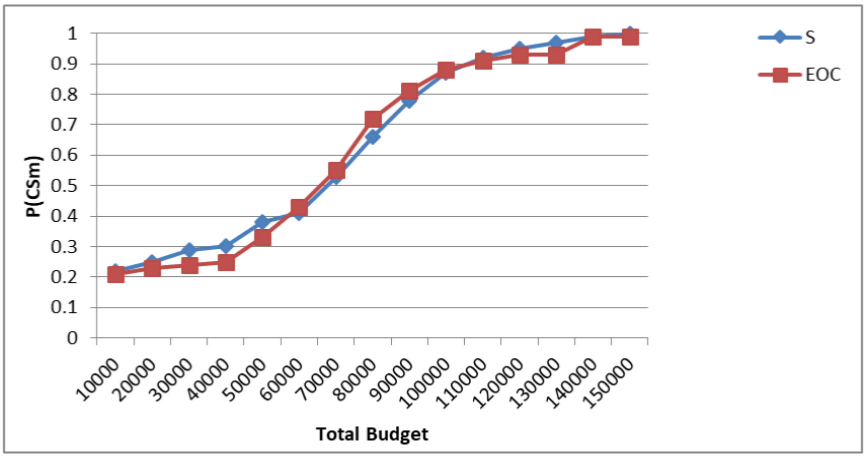


Figure 5: Comparison of the $P(CS_m)$ of S and EOC stopping rules over T budget for (BAP).

Figure 6 gives the $E(OC_m)$ performance measure for the proposed algorithm with two stopping rules, S and EOC, over different values of simulation budget. Clearly, the algorithm produces a very small $E(OC_m)$ under the two stopping rules with a little preferance of the EOC stopping rule over the sequential stopping rule. Also, with high value of total budget, the algorithm gives a very small value of $E(OC_m)$ which is close to 0, especially when the expected opportunity cost EOC stopping rule is used.

To enhance the performance of the algorithm, we increase the number of samples used in the multiple replication simulation method in order to get better estimates of the sample mean. Figure 7 gives the value of the $E(OC_m)$ against the number of samples M in the multiple replications method, for the two stopping rules S and EOC. Obviously, it shows that the increase in M gives a smaller value of $E(OC_m)$. It is clear again that the second stopping rule that uses EOC gives a slight better performance over the sequential stopping rule.

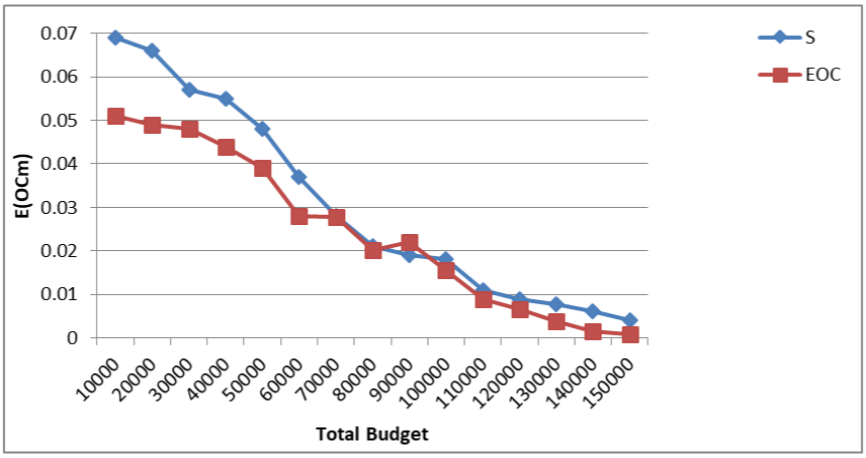


Figure 6: Comparison of the $E(OC_m)$ of S and EOC stopping rules over T budget for (BAP).

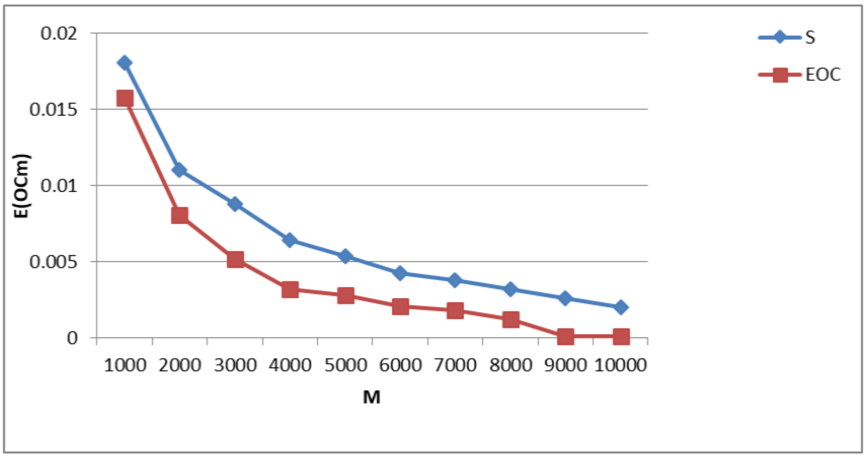


Figure 7: Comparison of the $E(OC_m)$ of S and EOC stopping rules over M replications for (BAP).

From MIM and BAP examples, it is clear that if objective is to select the best designs with high $P(CS_m)$ and minimum elapsed time, then the algorithm with the two stopping rules gives almost the same results. On the other hand, if the objective is to select the best design with minimum $E(OC_m)$, then the algorithm behaves better when the expected opportunity cost EOC is used as a stopping rule. Moreover, increasing the samples in the multiple replications M increases the performance of the algorithm under the second stopping rule. However, it is clear that when the value of the significant level ε is decreased, then the $E(OC_m)$ will decrease, but we get the optimal value for the $E(OC_m)$ which is 0 when the $\varepsilon = 0$. In this case the mean of the selected design will be equal to the mean of the actual best design. Nevertheless, we cannot take too small value of ε since this will require that the number of multiple replications M to be increased, and, of course, this leads to a huge computational time.

5 Conclusion

In this paper, we have discussed the effect of two stopping rules on the performance of a sequential selection procedure that is used to select a set of good enough simulated designs when the number of alternatives is very large. These two rules include the sequential S stopping rule and the expected opportunity cost EOC stopping rule. We have applied these rules on two different examples.

The results obtained from the numerical applications of the procedure using the two stopping rules indicate that to improve the efficiency of the approach using the EOC stopping rule, we need to increase the number of multiple replications M. We conclude that if the objective of the experiment is to select the best designs with high $P(CS_m)$, then both stopping rules give almost the same performance. However, if the objective is to select the best designs with minimum $E(OC_m)$, then the second stopping rule that uses EOC gives better performance.

References

[1] D. He, S.E. Chick, and C.H. Chen, Opportunity Cost and OCBA Selection Procedures in Ordinal Optimization for a Fixed Number of Alternative Systems. *IEEE Transactions on Systems* **37** (2007) 951–961.

[2] J. Branke, S.E. Chick and C. Schmidt. Selecting a Selection Procedure. *Management Science* **53** (12) (2007) 1916–1932.

[3] R.E. Bechhofer, T.J. Santner and D.M. Goldsman. *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons.* New York, Wiley, 1995.

[4] A.M. Law and W.D. Kelton. *Simulation Modeling and Analysis.* New York, McGraw-Hill, 2000.

[5] S.H. Kim and B.L. Nelson. Selecting the best system. In: *Handbooks in Operations Research and Management Science*, Chapter 17. Elsevier, 2006, 501–534.

[6] S.H. Kim and B.L. Nelson. On the asymptotic validity of fully sequential selection procedures for steady-state simulation. *Operations Research* **54** (3) (2006) 475–488.

[7] J.D. Gibbons, I. Olkin and M. Sobel. *Selecting and Ordering Populations: a New Statistical Methodology.* Society for Industrial and Applied Mathematics, 1999.

[8] S.S. Gupta and S. Panchapakesan. *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations*, Wiley, 1979.

[9] Y.C. Ho, R.S. Sreenivas and P. Vakili. Ordinal optimization of DEDS. *Discrete Events Dynamic Systems* **2** (1992) 61–88.

[10] M.H. Alrefaei and M.H. Almomani. Subset selection of best simulated systems. *Journal of the Franklin Institute* **344** (2007) 495–506.

[11] M.H. Almomani and M.H. Alrefaei. A three-stage procedure for selecting a good enough simulated system. *Journal of Applied Probability and Statistics* **6** (1&2) (2012) 13–27.

[12] M.H. Almomani and R. Abdul Rahman. Selecting a good stochastic system for the large number of alternatives. *Communications in Statistics-Simulation and Computation* **41** (2012) 222–237.

[13] M. Al-Salem, M.H. Almomani, M.H. Alrefaei and Ali Diabat. On the optimal computing budget allocation problem for large scale simulation optimization. *Simulation Modelling Practice and Theory* **71** (2017) 149–159.

[14] M.H. Almomani, M.H. Alrefaei and S. Al Mansour. A combined statistical selection procedure measured by expected opportunity cost. *Arabian Journal for Science and Engineering* **43** (6) (2018) 3163–3171, DOI: 10.1007/s13369-017-2865-8.

[15] M.H. Almomani, M.H. Alrefaei and S. Al Mansour. A method for selecting the best performance systems. *International Journal of Pure and Applied Mathematics* **120** (2) (2018) 191–202, DOI: 10.12732/ijpam.v120i2.3.

[16] M.H. Alrefaei, M.H. Almomani and S.N. Alabed Alhadi. Asymptotic approximation of the probability of correctly selecting the best systems. In: *AIP Conference Proceedings* **1991** (1) (2018) 020022, AIP Publishing LLC, DOI: https://doi.org/10.1063/1.5047895.

[17] C.H. Chen, D. He, M. Fu and L.H. Lee. Efficient simulation budget allocation for selecting an optimal subset. *Informs Journal on Computing* **20** (4) (2008) 579–595.

[18] M.H. Almomani and M.H. Alrefaei. Ordinal optimization with computing budget allocation for selecting an optimal subset. *Asia-Pacific Journal of Operational Research* **33** (2) (2016), DOI: 10.1142/S0217595916500093.

[19] Y.C. Ho, Q.C. Zhao and Q.S. Jia. *Ordinal Optimization: Soft Optimization for Hard Problems*. Springer, 2007.

[20] S.C. Horng and S.S. Lin. Bat algorithm assisted by ordinal optimization for solving discrete probabilistic bicriteria optimization problems. *Mathematics and Computers in Simulation* **166** (2019) 346–364.

[21] T. Ma, F. Tian and B. Dong. Ordinal optimization-based performance model estimation method for HDFS. *IEEE Access* **8** (2019) 889–899.

[22] C.H. Chen, E. Yucesan and S.E. Chick. Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization. *Discrete Event Dynamic Systems* **10** (3) (2000) 251–270.

[23] C.H. Chen and L.H. Lee. Stochastic Simulation Optimization: An Optimal Computing Budget Allocation, Chapter 5. World Scientific Publishing Company, 2010.

[24] D. Goldsman, S.H. Kim, W.S. Marshall and B.L. Nelson. Ranking and selection for steady-state simulation: Procedures and perspectives. *INFORMS Journal on Computing* **14** (1) (2002) 2–19.

[25] M.H. Almomani, R. Abdul Rahman, A. Baharum and M.H. Alrefaei. A Selection Approach for Solving the Buffer Allocation Problem. *International Journal of the Physical Sciences* **7** (3) (2012) 413–422.

[26] C.T. Papadopoulos, M.E.J. O'kelly, M. J. Vidalis and D. Spinellis. *Analysis and Design of Discrete Part Production Lines*. Springer, 2009.

[27] M. Yuzukirmizi and J.M. Smith. Optimal buffer allocation in finite closed networks with multiple servers. *Computers & Operations Research* **35** (2008) 2579–2598.

[28] M.H. Alrefaei and S. Andradóttir. Discrete stochastic optimization using variants of the stochastic ruler method. *Naval Research Logistics* **52** (4) (2005) 344–360.

Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 27–55

Existence and Asymptotic Behavior of Unbounded Positive Solutions of a Nonlinear Degenerate Elliptic Equation

A. Bouzelmate* and A. Gmira

Department of Mathematics, Faculty of Sciences, University Abdelmalek Essaâdi, B.P. 2121, Tetouan, Morocco

Received: July 8, 2020; Revised: December 29, 2020

Abstract: This paper is a contribution to the study of the elliptic equation

$$\Delta_p u + \alpha u + \beta x \cdot \nabla u + |x|^l \, u^q = 0 \quad \text{in } \mathbb{R}^N,$$

where $p > 2$, $q > 1$, $N \geq 1$, $\alpha < 0$, $\beta < 0$ and $l < 0$.
If $q \leq p-1$ or $q > p-1$ and $\dfrac{\alpha}{\beta} \neq \dfrac{l+p}{q+1-p}$ or $\dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p} \geq \dfrac{N-p}{p}$, we prove the existence of unbounded radial solutions u and we obtain their asymptotic behavior. In particular, if $\dfrac{\alpha}{\beta} < \dfrac{-l}{q-1}$, $\lim\limits_{r\to+\infty} r^{l/(q-1)} u(r) = \left(\dfrac{\beta l}{q-1} - \alpha\right)^{1/(q-1)}$.

Keywords: *nonlinear parabolic problem; nonlinear degenerate elliptic equation; self-similar solutions; nonlinear dynamical system; unbounded solutions; energy function; asymptotic behavior.*

Mathematics Subject Classification (2010): 34L30, 35J60, 35K92, 35K55.

* Corresponding author: mailto:abouzelmatearij@uae.ac.ma

1 Introduction

This paper is devoted to the study of the elliptic equation

$$\Delta_p u + \alpha u + \beta x \cdot \nabla u + |x|^l \; u^q = 0 \qquad \text{in } \mathbb{R}^N, \tag{1}$$

where $\Delta_p u = div\left(|\nabla u|^{p-2}\nabla u\right)$, $p > 2$, $\nabla u = \left(\dfrac{\partial u}{\partial x_1}, \cdots, \dfrac{\partial u}{\partial x_N}\right)$, $N \geq 1$, $q > 1$, $\alpha < 0$, $\beta < 0$ and l is a real number such that $-p < l < 0$ and $-N < l < 0$.

Equations of the above form occur in the study of self-similar solutions of the nonlinear parabolic problem

$$u_t = \Delta_p u + |x|^l \, |u|^{q-1} u \quad \text{in } \mathbb{R}^N \times (0, +\infty). \tag{2}$$

A lot of work has been done concerning equation (1) when $l = 0$; discussions and bibliographies are found in [3], [4], [10], [11], [13], [15], [16] and [18]. When $p = 2$ and $-2 < l < 0$, equation (1) was studied in [8]. Note also that when $p > 2$ and $l < 0$, the equation was investigated for $\alpha > 0$ and $\beta > 0$ in [9] and was initiated in [7] by the authors for $\alpha < 0$ and $\beta < 0$.

In our paper [7], we studied radial solutions near 0 of the equation

$$\left(|u'|^{p-2}u'\right)' + \frac{N-1}{r}|u'|^{p-2}u' + \alpha u + \beta r u' + r^l |u|^{q-1}u = 0, \quad r > 0.$$

Among the results obtained, we showed that for any radial solution u with $u(0) > 0$, $\lim\limits_{r\to 0} r^{(N-1)/(p-1)}u'(r)$ exists and is finite. Moreover, for any $a > 0$ and $b \in \mathbb{R}$, there exists a unique function $u \in C^0([0, +\infty[) \cap C^1(]0, +\infty[)$ such that $|u'|^{p-2}u' \in C^1(]0, +\infty[)$, and satisfying the problem

$$(\mathbf{P}) \begin{cases} \left(|u'|^{p-2}u'\right)' + \dfrac{N-1}{r}|u'|^{p-2}u' + \alpha u + \beta r u' + r^l |u|^{q-1}u = 0, \quad r > 0, \\ u(0) = a, \quad \lim\limits_{r\to 0} r^{(N-1)/(p-1)}u'(r) = b, \end{cases}$$

where $p > 2$, $q > 1$, $N \geq 1$, $-p < l < 0$, $-N < l < 0$, $\alpha < 0$ and $\beta < 0$.

It is also proved that if a is small and $b = 0$, u is strictly positive.

Our aim in this paper, is to continue our study on the problem $(\mathbf{P})$. For this we must start with the analysis of solutions of the equation

$$\left(|u'|^{p-2}u'\right)' + \frac{N-1}{r}|u'|^{p-2}u' + \alpha u + \beta r u' + r^l u^q = 0, \quad r > 0. \tag{3}$$

First of all, it should be noted that when $0 < \dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p} < \dfrac{N-p}{p-1}$, we have an explicit solution $L\, r^{-\alpha/\beta}$, where

$$L = \left(N - p - \frac{\alpha}{\beta}(p-1)\right)^{1/(q+1-p)} \left(\frac{\alpha}{\beta}\right)^{(p-1)/(q+1-p)}.$$

This solution is bounded near infinity but singular at the origin. However, using the theory of ODE, the equation (3), for bounded solutions whether they are singular or not, can be considered in $+\infty$ as a perturbation of the equation

$$\left(|u'|^{p-2}u'\right)' + \frac{N-1}{r}|u'|^{p-2}u' + \alpha u + \beta r u' = 0,$$

whose solutions behave like the function $r^{-\alpha/\beta}$ near infinity.

On the other hand, the term $r^l u^q$ plays a key role in (3) for unbounded solutions and therefore, the perturbation theory can not be applicable. Hence the interest of focusing our study on unbounded solutions.

For this purpose, let us represent equation (3) in an equivalent but useful form. For any real c, we set

$$\upsilon_c(t) = r^c u(r) \quad \text{where } r > 0 \text{ and } t = \ln r. \tag{4}$$

Then υ_c satisfies

$$\omega_c'(t) + A_c\,\omega_c(t) + \alpha e^{K_c t}\upsilon_c(t) + \beta e^{K_c t}h_c(t) + e^{M_c t}\upsilon_c^q(t) = 0, \tag{5}$$

where

$$\omega_c(t) = |h_c|^{p-2}h_c(t), \qquad h_c(t) = \upsilon_c'(t) - c\upsilon_c(t), \tag{6}$$

$$A_c = N - p - c(p-1),\ \ K_c = c(p-2) + p \ \text{ and } \ M_c = l + p - c(q+1-p). \tag{7}$$

Five critical values of the parameter c will be involved: $\dfrac{\alpha}{\beta}$, $\dfrac{l}{q-1}$ and those which cancel A_c, K_c or M_c, that is, $c = \dfrac{-p}{p-2}$, $c = \dfrac{N-p}{p-1}$ or $c = \dfrac{l+p}{q+1-p}$.

The study of monotonicity of $r^c u(r)$ for these last five values, combined with the behavior of bounded solutions, allows us to show the existence of unbounded solutions of problem (**P**).

A fine analysis of the equation in logarithmic form, using some energy function, gives the asymptotic behavior of solutions. Our main results are given by the following theorems.

Theorem 1.1 *Assume that* $q \le p-1$ *or* $q > p-1$ *and* $\dfrac{\alpha}{\beta} \neq \dfrac{l+p}{q+1-p}$ *or* $\dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p} \ge \dfrac{N-p}{p}$. *Then any positive solution of problem* (**P**) *is unbounded.*

Theorem 1.2 *Assume* $q \ge p\left(2 + 2^{p-1}\right) - 1$ *and* $\dfrac{l}{q-1} < \min\left(\dfrac{-\alpha}{\beta}, \dfrac{N-p}{p-1}\right)$. *Let* u *be an unbounded positive solution of problem* (**P**). *Then*

$$\lim_{r\to+\infty} r^{l/(q-1)}u(r) = \Gamma$$

and

$$\lim_{r\to+\infty} r^{l/(q-1)+1}u'(r) = \frac{-l}{q-1}\Gamma,$$

where

$$\Gamma = \left(\frac{\beta l}{q-1} - \alpha\right)^{1/(q-1)}.$$

The rest of the paper is organized as follows. In the second section, we present fundamental properties of solutions of equation (3). The third section concerns the existence and the asymptotic behavior near infinity of unbounded positive solutions of problem (**P**).

2 Fundamental Properties

In this section, we give some fundamental properties that are the key stone of the main results. For this purpose, we introduce, for any real $c \neq 0$, the function

$$E_c(r) = cu(r) + ru'(r), \quad r > 0. \tag{8}$$

It is clear that

$$(r^c u(r))' = r^{c-1} E_c(r), \quad r > 0. \tag{9}$$

With the logarithmic change (4), we have

$$v_c'(t) = r^c\, E_c(r) \text{ and } h_c(t) = r^{c+1} u'(r). \tag{10}$$

The monotonicity of the function $r^c u(r)$ can be obtained by the sign of the function $E_c(r)$. Observe that for any $r > 0$ such that $u'(r) \neq 0$, we have

$$\begin{aligned}(p-1)\,|u'|^{p-2}(r) E_c'(r) &= \Big(p - N + c(p-1)\Big)\,|u'|^{p-2}\,u'(r) - \beta r^2 u'(r) - \\ &\quad \alpha r u(r) - r^{l+1} u^q(r) \\ &= \Big(p - N + c(p-1)\Big)\,|u'|^{p-2}\,u'(r) + |\beta| r\, E_{\alpha/\beta}(r) - \\ &\quad r^{l+1} u^q(r). \end{aligned} \tag{11}$$

Consequently, if $E_c(r_0) = 0$ for some $r_0 > 0$, equation (3) gives

$$\begin{aligned}(p-1)\,|u'|^{p-2}(r_0) E_c'(r_0) &= -r_0 u(r_0) \Big[\alpha - c\beta + r_0^l u^{q-1}(r_0) + \\ &\quad \Big(p - N + c(p-1)\Big)|c|^{p-2} c\, r_0^{-p} u^{p-2}(r_0)\Big] \\ &= -r_0^{l+1} u^q(r_0) \Big[1 + (\alpha - c\beta) r_0^{-l} u^{1-q}(r_0) + \\ &\quad \Big(p - N + c(p-1)\Big)|c|^{p-2} c\, r_0^{-l-p} u^{p-1-q}(r_0)\Big], \end{aligned} \tag{12}$$

from which we can study the sign of $E_c(r)$ and we use the following remarks.

Remark 2.1 If there exists r_0 such that $E_c(r_0) = 0$ and $E_c'(r_0) \neq 0$, then $E_c(r) \neq 0$ for any $r > r_0$.

Remark 2.2 If u is a bounded solution of equation (3), then, by expression (12) and Remark 2.1, we have, for any $c > 0$ such that $c\beta - \alpha \neq 0$, $E_c(r) \neq 0$ for large r.

We first give the sign of $E_{l/(q-1)}$ and $E_{-p/(p-2)}$.

Proposition 2.1 *Let u be a solution of equation (3). We put*

$$\Gamma = \left(\frac{\beta l}{q-1} - \alpha\right)^{1/(q-1)}. \tag{13}$$

The following holds:

(i) If $\dfrac{N-p}{p-1} > \dfrac{l}{q-1}$ and $u(r) > \Gamma r^{-l/(q-1)}$ for large r, then $E_{l/(q-1)}(r) \neq 0$ for large

r.
(ii) If $q \geq p-1$, $\lim_{r\to+\infty} u(r) = +\infty$ *and* $\lim_{r\to+\infty} r^{l/(q-1)}u(r) = 0$, *then* $E_{l/(q-1)}(r) < 0$ *for large* r.
(iii) If $q \geq p-1$ *and* $\lim_{r\to+\infty} r^{l/(q-1)}u(r) = +\infty$, *then* $E_{-p/(p-2)}(r) < 0$ *for large* r.

The proof requires the following result.

Lemma 2.1 *Let* u *be a solution of equation (3) such that for large* r,

$$u(r) > \Gamma r^{-l/(q-1)},$$

where Γ *is given by (13). Then* $u'(r) > 0$ *for large* r.

Proof. Suppose that there exists a large r_0 such that $u'(r_0) = 0$, then

$$\Big(|u'|^{p-2}u'\Big)'(r_0) = -\Big[\alpha + r_0^l u^{q-1}(r_0)\Big]u(r_0) < -\frac{\beta l}{q-1}u(r_0) < 0.$$

Hence, $u'(r) \neq 0$ for any $r > r_0$. Moreover, since $\lim_{r\to+\infty} u(r) = +\infty$, we have $u'(r) > 0$ for large r. □

Now, we turn to the proof of Proposition 2.1.
Proof. (of Proposition 2.1). The cases (i) and (ii) follow easily from Remark 2.1 and relation (9). Assume now that we are in the case (iii), then again Remark 2.1 gives $E_{-p/p-2}(r) \neq 0$ for large r. Suppose by contradiction that $E_{-p/(p-2)}(r) > 0$ for large r. Hence, by (9), $\lim_{r\to+\infty} r^{-p/(p-2)}u(r) \in]0, +\infty]$. Set

$$\varphi(r) = r^{N-1}|u'|^{p-2}u'(r) + \beta r^N u(r), \qquad r > 0, \tag{14}$$

then, by equation (3), we get

$$\varphi'(r) = r^{N-1}u(r)\Big[N\beta - \alpha - r^l u^{q-1}(r)\Big], \qquad \text{for any } r > 0, \tag{15}$$

so $\lim_{r\to+\infty} \varphi'(r) = -\infty$, in particular, $\varphi(r) < 0$ for large r, that is,

$$r^{N-1}|u'|^{p-2}u'(r) < |\beta| r^N u(r) \tag{16}$$

for large r. Note that by Lemma 2.1, we have $u'(r) > 0$ for large r and then a simple integration of this last inequality on (r_0, r) gives $\lim_{r\to+\infty} r^{-p/(p-2)}u(r) = d_1 > 0$.
Now we are going to the logarithmic change. First, from (5) we obtain

$$\begin{aligned} \omega'_{-p/(p-2)}(t) \quad &= -A_{-p/(p-2)}\,\omega_{-p/(p-2)}(t) - \alpha\upsilon_{-p/(p-2)}(t) - \\ &\beta h_{-p/(p-2)}(t) - e^{M_{-p/(p-2)}t}\upsilon^q_{-p/(p-2)}(t). \end{aligned} \tag{17}$$

Note that, as $E_{-p/(p-2)}(r) > 0$ for large r, then by (10) $\upsilon'_{-p/(p-2)}(t) > 0$ for large t. Then, since $\lim_{r\to+\infty} r^{-p/(p-2)}u(r) = d_1 > 0$, necessarily $\lim_{t\to+\infty} \upsilon'_{-p/(p-2)}(t) = 0$, which implies that $\lim_{t\to+\infty} h_{-p/(p-2)}(t) = \dfrac{p}{p-2}d_1$ and therefore $\lim_{t\to+\infty} \omega_{-p/(p-2)}(t) = \Big(\dfrac{p}{p-2}d_1\Big)^{p-1}$. As $M_{-p/(p-2)} > 0$, it follows by letting $t \to +\infty$ in (17) that $\lim_{t\to+\infty} \omega'_{-p/(p-2)}(t) = -\infty$. This is impossible as $\lim_{t\to+\infty} \omega_{-p/(p-2)}(t)$ is finite.

Consequently, $E_{-p/(p-2)}(r) < 0$ for large r. The proof is complete. □

Proposition 2.2 *Let u be a solution of equation (3). Then, for any $c \geq \max\left(\frac{N-p}{p-1}, \frac{\alpha}{\beta}\right)$, $E_c(r) > 0$ for large r.*

We need the following lemma.

Lemma 2.2 *Assume that u is a bounded solution of equation (3). Then $u'(r) < 0$ for large r.*

Proof. First, we claim that $u'(r)$ cannot vanish for large enough r. Assume by contradiction that there exists a large extremum r_0 of u. Then, according to equation (3), we get $\left(|u'|^{p-2}u'\right)'(r_0) = -\left[\alpha + r_0^l u^{q-1}(r_0)\right]u(r_0) > 0$, then $u'(r) > 0$ for any $r > r_0$, from which it turns out that $u'(r) \neq 0$ for large r. If $u'(r) > 0$ for large r, it follows by the boundedness of u that $\lim_{r\to+\infty} u(r) = L > 0$ and from (8), $E_N(r) > 0$ for large r. On the other hand, by (11), we have that for large r

$$\begin{aligned}(p-1)\left|u'\right|^{p-2}(r)E_N'(r) &= (p+N(p-2))\left|u'\right|^{p-1} + |\beta|r^2u'(r)\\ &+ \left[|\alpha| - r^l u^{q-1}\right]ru(r)\\ &> \left[|\alpha| - r^l u^{q-1}\right]ru(r) > 0. \qquad (18)\end{aligned}$$

So, $\lim_{r\to+\infty} E_N(r) \in]0, +\infty]$.

Note that if $\lim_{r\to+\infty} E_N(r) = +\infty$, then by (8), $\lim_{r\to+\infty} ru'(r) = +\infty$, which contradicts the boundedness of u. On the other hand, if $\lim_{r\to+\infty} E_N(r)$ is finite and strictly positive, necessarily $\lim_{r\to+\infty} ru'(r) = 0$ and by letting r to $+\infty$ in equation (3), we obtain $\lim_{r\to+\infty}(|u'|^{p-2}u')'(r) = -\alpha L > 0$, which implies that $\lim_{r\to+\infty} u'(r) = +\infty$ and we have again a contradiction. Consequently, $u'(r) < 0$ for large r. □

Now we prove Proposition 2.2.

Proof. (of Proposition 2.2). We distinguish two cases.

Case 1: u is bounded.

Assume $\frac{\alpha}{\beta} \geq \frac{N-p}{p-1}$. We have easily from Remark 2.1, $E_{\alpha/\beta}(r) \neq 0$ for large r. Suppose that $E_{\alpha/\beta}(r) < 0$ for large r, it turns out by (9) that the function $r^{\alpha/\beta}u(r)$ is decreasing for large r, hence $\lim_{r\to+\infty} r^{\alpha/\beta}u(r)$ exists and is finite. But $\frac{\alpha}{\beta} > 0$, then necessarily $\lim_{r\to+\infty} u(r) = 0$. On the other hand, using the equation (11), Lemma 2.2, the fact that $E_{\alpha/\beta}(r) < 0$ and $\frac{\alpha}{\beta} \geq \frac{N-p}{p-1}$, we get $E'_{\alpha/\beta}(r) < 0$ for large r, hence $\lim_{r\to+\infty} E_{\alpha/\beta}(r) \in [-\infty, 0[$. Combining this with $\lim_{r\to+\infty} u(r) = 0$, we obtain from (8), $\lim_{r\to+\infty} ru'(r) \in [-\infty, 0[$. But this contradicts the fact that u is positive. Consequently, $E_{\alpha/\beta}(r) > 0$ for large r and $E_c(r) > 0$ for $c \geq \frac{\alpha}{\beta}$.

Assume now $\frac{\alpha}{\beta} < \frac{N-p}{p-1}$. Then, from Remark 2.2, $E_{(N-p)/(p-1)}(r) \neq 0$ for large r. Suppose on the contrary that $E_{(N-p)/(p-1)}(r) < 0$ for large r. Since $\frac{\alpha}{\beta} < \frac{N-p}{p-1}$, we have $E_{\alpha/\beta}(r) < 0$ for large r. On the other hand, by (11), we have $E'_{(N-p)/(p-1)}(r) < 0$ and thereby $\lim_{r\to+\infty} E_{(N-p)/(p-1)}(r) \in [-\infty, 0[$. Using similar arguments as before, we get a contradiction. In fact, as the function $r^{(N-p)/(p-1)}u(r)$ is decreasing, then it admits a finite limit, this implies that $\lim_{r\to+\infty} u(r) = 0$. So, by (8), $\lim_{r\to+\infty} ru'(r) \in [-\infty, 0[$, which is impossible. Consequently, $E_{(N-p)/(p-1)}(r) > 0$ for large r and $E_c(r) > 0$ for $c \geq \dfrac{N-p}{p-1}$.

Case 2: u is unbounded.

It is easy to see by Remark 2.1 that for $c \geq \max\left(\dfrac{N-p}{p-1}, \dfrac{\alpha}{\beta}\right)$, $E_c(r) \neq 0$ for large r, that is, by (9), $r^c u(r)$ is strictly monotone. Since u is unbounded, necessarily $\lim_{r\to+\infty} r^c u(r) = +\infty$. Consequently, by (9), $E_c(r) > 0$ for large r. □

3 Unbounded Solutions

In this section we study the unbounded positive solutions of problem (**P**) and we give their asymptotic behavior near infinity.

Theorem 3.1 *Assume that* $q \leq p-1$ *or* $q > p-1$ *and* $\dfrac{\alpha}{\beta} \neq \dfrac{l+p}{q+1-p}$ *or* $\dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p} \geq \dfrac{N-p}{p}$. *Then any positive solution of problem* (**P**) *is unbounded.*

Before giving the proof, we need the behavior of bounded solutions near infinity. For this purpose we start with the following result.

Proposition 3.1 *Assume that* u *is a bounded solution of equation (3). Then*

$$\lim_{r\to+\infty} u(r) = \lim_{r\to+\infty} ru'(r) = 0. \tag{19}$$

Proof. Recall Lemma 2.2 and Proposition 2.2, we deduce that $\lim_{r\to+\infty} u(r) = L$ exists and for any $c \geq \max\left(\dfrac{N-p}{p-1}, \dfrac{\alpha}{\beta}\right)$,

$$-cu(r) < r\,u'(r) < 0 \quad \text{for large } r. \tag{20}$$

Thus $ru'(r)$ is bounded for large r. Assume by contradiction that $L > 0$.

First, suppose that $ru'(r)$ is monotone for large r, then necessarily $\lim_{r\to+\infty} ru'(r) = 0$ and we get from equation (3)

$$\lim_{r\to+\infty} \left(|u'|^{p-2}u'\right)'(r) = -\alpha L > 0.$$

This is a contradiction with $u'(r) < 0$ for large r.

Next, suppose that $ru'(r)$ is oscillating for large r. Since u is positive and strictly decreasing, one has $\limsup_{r\to+\infty} ru'(r) = 0$. Otherwise, there exists a constant $C > 0$ such that $ru'(r) < -C$ for large r. This contradicts $u(r) > 0$. Consequently, there exists a sequence $\{\xi_i\}$ going to $+\infty$ as $i \to +\infty$ such that the function $ru'(r)$ has a local maximum in ξ_i satisfying $\lim_{i\to+\infty} \xi_i u'(\xi_i) = \limsup_{r\to+\infty} ru'(r) = 0$ and $u'(\xi_i) + \xi_i u''(\xi_i) = 0$ (u'' exists because $u' < 0$). Therefore

$$\lim_{i\to+\infty} u'(\xi_i) = \lim_{i\to+\infty} u''(\xi_i) = 0. \tag{21}$$

On the other hand, take $r = \xi_i$ in equation (3), we obtain

$$\lim_{i\to+\infty} \Big(|u'|^{p-2}u'\Big)'(\xi_i) = -\alpha L > 0,$$

that is,

$$\lim_{i\to+\infty} |u'|^{p-2}u''(\xi_i) = -\frac{\alpha L}{p-1} > 0.$$

This contradicts (21).

It follows from both cases that $\lim_{r\to+\infty} u(r) = 0$. Hence, by inequality (20), we have $\lim_{r\to+\infty} ru'(r) = 0$ and the proof is complete. $\square$

Proposition 3.2 *Assume that u is a bounded solution of equation (3). Then*

(i) If $0 < c < \dfrac{\alpha}{\beta} < N$, $\lim_{r\to+\infty} r^c u(r) = 0$.

(ii) If $\dfrac{\alpha}{\beta} < c < N$, $\lim_{r\to+\infty} r^c u(r) = +\infty$.

Proof. According to Remark 2.2 and expression (9), $r^c u(r)$ is monotone for large r for any $c \neq \dfrac{\alpha}{\beta}$. Hence $\lim_{r\to+\infty} r^c u(r) \in [0, +\infty]$.

(i) Assume by contradiction that $\lim_{r\to+\infty} r^c u(r) \in]0, +\infty]$. Then, for $0 < c < \dfrac{\alpha}{\beta} < N$, $\lim_{r\to+\infty} r^N u(r) = +\infty$ and by Remark 2.2, $E_N(r) > 0$ for large r. Hence, using the fact that u is bounded, $u'(r) < 0$ for large r and expression of E_N, we find

$$\lim_{r\to+\infty} \frac{|u'(r)|^{p-1}}{ru} = 0. \tag{22}$$

Then by (22), (14) and (15), we have

$$\varphi(r) \underset{+\infty}{\sim} \beta r^N u(r) < 0 \tag{23}$$

and

$$\varphi'(r) \underset{+\infty}{\sim} (N\beta - \alpha) r^{N-1} u(r). \tag{24}$$

Combining these two estimates, we get

$$r(r^{c-N}\varphi)' \underset{+\infty}{\sim} (c\beta - \alpha) r^c u(r).$$

Since $\lim_{r\to+\infty} r^c u(r) \in]0, +\infty]$ and $c\beta - \alpha > 0$, there exists some $C_1 > 0$ such that

$$r(r^{c-N}\varphi)'(r) > C_1 \quad \text{for large } r.$$

Whence $\lim_{r\to+\infty}\varphi(r) = +\infty$. This is a contradiction with (23). We deduce that $\lim_{r\to+\infty} r^c u(r) = 0$.
(ii) Assume by contradiction that $\lim_{r\to+\infty} r^c u(r) = K \in [0,+\infty[$. We distinguish two cases:
- $K = 0$. Then necessarily $E_c(r) < 0$ for large r, this means that

$$\frac{u(r)}{r|u'(r)|} < \frac{1}{c} \quad \text{for large}\, r. \tag{25}$$

Using equation (3) and the fact that $u > 0, u' < 0$ and $\frac{\alpha}{\beta} < c$, we obtain

$$\Big(|u'|^{p-2}u'\Big)'(r) < r|u'(r)|\Big[\beta + \frac{|\alpha|}{c} + \frac{N-1}{r^2}|u'(r)|^{p-2}\Big] < 0 \quad \text{ for large } r.$$

Thus $u'(r) > 0$ for large r, which is a contradiction with Lemma 2.2.
- $K > 0$. Since $c < N$, $\lim_{r\to+\infty} r^N u(r) = +\infty$ and $E_N(r) > 0$ for large r. Therefore, (22) is satisfied and thereby from (23) and (24), we get

$$\lim_{r\to+\infty}\varphi(r) = -\infty \text{ and } \lim_{r\to+\infty} r^{c+1-N}\varphi'(r) = K(N\beta - \alpha). \tag{26}$$

Then Hopital's rule implies

$$\lim_{r\to+\infty} r^{c-N}\varphi(r) = \frac{K(N\beta-\alpha)}{N-c}. \tag{27}$$

But from (23), this limit is exactly $K\beta$. This contradicts the fact that $c > \frac{\alpha}{\beta}$. Consequently, $\lim_{r\to+\infty} r^c u(r) = +\infty$ and the proof of the proposition is complete. □

Proposition 3.3 *Assume that u is a bounded solution of equation (3). Then the function $r^{\alpha/\beta}u(r)$ is not strictly monotone for large r.*

Proof. We argue by contradiction and assume that $r^{\alpha/\beta}u(r)$ is strictly monotone for large r. Therefore, according to (9), $E_{\alpha/\beta}(r) \neq 0$ for large r. We distinguish two cases.
Case 1: $E_{\alpha/\beta}(r) > 0$ for large r.
We set

$$J_1(r) = u(r) - r^{p-1}|u'|^{p-1}. \tag{28}$$

Then for large r,

$$J_1'(r) = r^{p-1}u\Big[-\alpha - r^l u^{q-1}\Big] - r^p|u'|\Big[-\beta + r^{-p} + (p-N)r^{-2}|u'|^{p-2}\Big]. \tag{29}$$

Using now Proposition 3.1 and $E_{\alpha/\beta}(r) > 0$ for large r, we get

$$\lim_{r\to+\infty} J_1(r) = 0,$$

$$J_1(r) > u(r)\left[1 - \left(\frac{\alpha}{\beta}\right)^{p-1} u^{p-2}(r)\right] > 0 \quad \text{ for large } r$$

and

$$J_1'(r) \underset{+\infty}{\sim} -\alpha r^{p-1}u(r) + \beta r^p|u'(r)| = -\beta r^{p-1}E_{\alpha/\beta}(r) > 0 \quad \text{for large } r.$$

This is a contradiction.
Case 2: $E_{\alpha/\beta}(r) < 0$ for large r.
Then necessarily $\dfrac{\alpha}{\beta} < \dfrac{N-p}{p-1}$, by Proposition 2.2.
Now, we set

$$J_2(r) = r^k u^p(r) - r^{p-1}|u'|^{p-1}, \tag{30}$$

with $0 < k < \min\left(\dfrac{\alpha}{\beta}, p\right)$. Then, for large r,

$$\begin{aligned} J_2'(r) \quad &= r^{p-1}u\Big[-\alpha - r^l u^{q-1} + kr^{k-p}u^{p-1}\Big] - r^p|u'|\Big[-\beta + \\ &+ pr^{k-p}u^{p-1} + (p-N)r^{-2}|u'|^{p-2}\Big]. \end{aligned} \tag{31}$$

As $k < p$, $k < \dfrac{\alpha}{\beta} < \dfrac{N-p}{p-1} < N$, $E_{\alpha/\beta}(r) < 0$ for large r, by Proposition 3.1 and Proposition 3.2, we obtain

$$\lim_{r\to+\infty} J_2(r) = 0,$$

$$J_2(r) < u^{p-1}(r)\left[-\left(\frac{\alpha}{\beta}\right)^{p-1} + r^k u(r)\right] < 0 \quad \text{for large } r$$

and

$$J_2'(r) \underset{+\infty}{\sim} -\alpha r^{p-1}u(r) + \beta r^p|u'(r)| = -\beta r^{p-1}E_{\alpha/\beta}(r) < 0 \quad \text{for large } r.$$

Again, we have a contradiction. Consequently, the function $r^{\alpha/\beta}u(r)$ is not strictly monotone for large r. The proof of the proposition is complete. □

Proposition 3.4 *Assume that u is a bounded solution of equation (3). If* $\dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p} < \dfrac{N-p}{p-1}$, *then*

$$\lim_{r\to+\infty} r^{\alpha/\beta}u(r) = L \tag{32}$$

and

$$\lim_{r\to+\infty} r^{\alpha/\beta+1}u'(r) = \frac{-\alpha}{\beta}L, \tag{33}$$

where

$$L = \left(N - p - \frac{\alpha}{\beta}(p-1)\right)^{1/(q+1-p)}\left(\frac{\alpha}{\beta}\right)^{(p-1)/(q+1-p)}. \tag{34}$$

Proof. Define the following function:

$$I(r) = r^{\alpha/\beta}u(r)\left[\frac{|u'|^{p-2}u'(r)}{ru(r)} + \beta\right]. \tag{35}$$

We have $I(r) < 0$ for large r. Its derivative is given by

$$I'(r) = \left(\frac{\alpha}{\beta} - N\right) r^{\alpha/\beta-2}|u'|^{p-2}u'(r) - r^{\alpha/\beta+l-1}u^q(r). \tag{36}$$

The proof will be done in five steps.
Step 1: $I(r) \underset{+\infty}{\sim} \beta r^{\alpha/\beta} u(r)$.
Since $\dfrac{\alpha}{\beta} < \dfrac{N-p}{p-1}$, Proposition 2.2 implies that $E_{(N-p)/(p-1)}(r) > 0$ for large r and thus, from the boundedness of u, $\lim\limits_{r\to+\infty} \dfrac{|u'|^{p-2}u'(r)}{ru} = 0$ and one deduces that $I(r) \underset{+\infty}{\sim} \beta r^{\alpha/\beta} u(r)$.
Step 2: $\lim\limits_{r\to+\infty} r^{\alpha/\beta} u(r)$ exists and is finite.
According to Step 1, it suffices to prove that $\lim\limits_{r\to+\infty} I(r)$ exists and is finite. For this purpose, consider a real σ such that

$$0 < \sigma < \min\left(\frac{\alpha}{\beta}, \frac{1}{q}\left(\frac{\alpha}{\beta}(q-1) - l\right), \frac{1}{p-1}\left(\frac{\alpha}{\beta}(p-2) + p\right)\right).$$

So, by Proposition 3.2, $\lim\limits_{r\to+\infty} r^{\alpha/\beta-\sigma} u(r) = 0$. In particular, there exists a constant $C > 0$ such that

$$u(r) \le C r^{\sigma-\alpha/\beta} \quad \text{for large } r. \tag{37}$$

Recall the positivity of $E_{(N-p)/(p-1)}(r)$ for large r, we get that there exists $C_1 > 0$ such that

$$r^{\alpha/\beta-2}|u'|^{p-1} < C_1^{p-1} r^{\gamma} \quad \text{for large } r, \tag{38}$$

where $\gamma = \dfrac{\alpha}{\beta}(2-p) + \sigma(p-1) - p - 1$.
By the choice of σ, the functions $r \to r^{\alpha/\beta+l-1}u^q(r)$ and $r \to r^{\alpha/\beta-2}|u'|^{p-1}$ are integrable near $+\infty$, therefore, $I'(r)$ is also integrable near $+\infty$. To conclude, we observe that for any $r_0 > 0$,

$$\lim_{r\to+\infty} I(r) = \int_{r_0}^{+\infty} I'(s)ds + I(r_0)$$

exists and is finite. Therefore, $\lim\limits_{r\to+\infty} r^{\alpha/\beta} u(r)$ exists and is finite.
Set $\lim\limits_{r\to+\infty} r^{\alpha/\beta} u(r) = L \ge 0$.
Step 3: $\lim\limits_{r\to+\infty} r^{\alpha/\beta} u(r) = L > 0$.
Assume that $\lim\limits_{r\to+\infty} r^{\alpha/\beta} u(r) = 0$. Then $\lim\limits_{r\to+\infty} I(r) = 0$. Therefore, applying Hopital's rule and using the first step, we obtain

$$\lim_{r\to+\infty} \frac{I'(r)}{\left(r^{\alpha/\beta}u(r)\right)'} = \lim_{r\to+\infty} \frac{I(r)}{r^{\alpha/\beta}u} = \beta. \tag{39}$$

On the other hand, using (36), we have

$$I'(r) = r^{\alpha/\beta-2}|u'|^{p-1}\left[N - \frac{\alpha}{\beta} - \frac{r^{l+1}u^q(r)}{|u'|^{p-1}}\right]. \tag{40}$$

Let $0 < c < \dfrac{\alpha}{\beta} < N$, then $\lim\limits_{r\to+\infty} r^c u(r) = 0$ and according to Remark 2.2, we have $E_c(r) \neq 0$ for large r. Therefore, necessarily by (9), $E_c(r) < 0$ for large r. Hence,

$$0 < \frac{r^{l+1}u^q(r)}{|u'|^{p-1}} < c^{1-p} r^{l+p} u^{q+1-p}(r). \tag{41}$$

Since $\dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p}$, then $\lim\limits_{r\to+\infty} r^{l+p}u^{q+1-p}(r) = 0$ and therefore $\lim\limits_{r\to+\infty} \dfrac{r^{l+1}u^q(r)}{|u'|^{p-1}} = 0$. Hence, as $\dfrac{\alpha}{\beta} < N$ and $|u'(r)| > 0$ for large r, we have by (40), $I'(r) > 0$ for large r. Consequently, by (39) and the fact that $\beta < 0$, $\left(r^{\alpha/\beta}u(r)\right)' < 0$ for large r. But this contradicts Proposition 3.3.

Consequently, $\lim\limits_{r\to+\infty} r^{\alpha/\beta}u(r) = L > 0$.

Step 4: $\lim\limits_{r\to+\infty} r^{\alpha/\beta+1}u'(r) = \dfrac{-\alpha}{\beta}L$.
Since $\lim\limits_{r\to+\infty} u(r) = 0$, then applying Hopital's rule, we obtain

$$\lim_{r\to+\infty} r^{\alpha/\beta+1}u'(r) = \frac{-\alpha}{\beta}\lim_{r\to+\infty} r^{\alpha/\beta}u(r) = \frac{-\alpha}{\beta}L.$$

Step 5: $L = \left(N - p - \dfrac{\alpha}{\beta}(p-1)\right)^{1/(q+1-p)}\left(\dfrac{\alpha}{\beta}\right)^{(p-1)/(q+1-p)}$.
According to (11), we have

$$\begin{aligned} -\beta r E_{\alpha/\beta}(r) &= |u'|^{p-2}u'(r)\Bigg[\left(N - p - \frac{\alpha}{\beta}(p-1)\right) \\ &\quad + (p-1)\frac{E'_{\alpha/\beta}(r)}{u'(r)} + \frac{r^{l+1}u^q(r)}{|u'|^{p-2}u'(r)}\Bigg]. \end{aligned} \tag{42}$$

Using Step 3 and Step 4 and applying Hopital's rule, we get

$$\lim_{r\to+\infty}\frac{E'_{\alpha/\beta}(r)}{u'(r)} = \lim_{r\to+\infty}\frac{E_{\alpha/\beta}(r)}{u(r)} = \lim_{r\to+\infty}\left(\frac{\alpha}{\beta} + \frac{ru'(r)}{u}\right) = 0 \tag{43}$$

and

$$\lim_{r\to+\infty}\frac{r^{l+1}u^q(r)}{|u'|^{p-2}u'(r)} = \frac{-L^{q+1-p}}{\left(\dfrac{\alpha}{\beta}\right)^{p-1}}, \tag{44}$$

when $\dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p}$. Suppose by contradiction

$$N - p - \frac{\alpha}{\beta}(p-1) - \frac{L^{q+1-p}}{\left(\dfrac{\alpha}{\beta}\right)^{p-1}} \neq 0.$$

After combining these estimates, equation (42) gives

$$-\beta r E_{\alpha/\beta}(r) \underset{+\infty}{\sim} \left[N - p - \frac{\alpha}{\beta}(p-1) - \frac{L^{q+1-p}}{\left(\dfrac{\alpha}{\beta}\right)^{p-1}}\right]|u'|^{p-2}u'(r).$$

So, $E_{\alpha/\beta}(r) \neq 0$ for large r, that is, $r^{\alpha/\beta}u(r)$ is strictly monotone for large r. Again, this contradicts Proposition 3.3.

Consequently,

$$L = \left(N - p - \frac{\alpha}{\beta}(p-1)\right)^{1/(q+1-p)} \left(\frac{\alpha}{\beta}\right)^{(p-1)/(q+1-p)}.$$

The proof is complete. □

The following figure illustrates the behavior of the bounded solution.

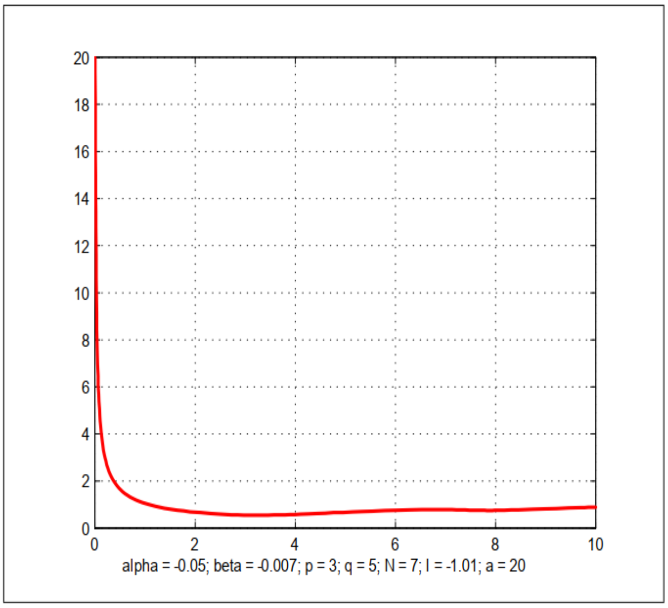


Figure 1: Bounded solution.

Now, we turn to the proof of Theorem 3.1.
Proof. (of Theorem 3.1). We argue by contradiction and assume that u is bounded. We claim that if one of the first two cases holds, then $E_{\alpha/\beta}(r) \neq 0$ for large r and this contradicts Proposition 3.3. In fact, if $\dfrac{\alpha}{\beta} \geq \dfrac{N-p}{p-1}$, we have $E_{\alpha/\beta}(r) > 0$ for large r, from Proposition 2.2.

Now we consider the case where $\dfrac{\alpha}{\beta} < \dfrac{N-p}{p-1}$.
Assume that there exists a large r_0 such that $E_{\alpha/\beta}(r_0) = 0$. Recall formula (12) with $c = \dfrac{\alpha}{\beta}$, we have in the case $q \leq p-1$, $E'_{\alpha/\beta}(r_0) \neq 0$. In the case $q > p-1$ and $\dfrac{\alpha}{\beta} \neq \dfrac{l+p}{q+1-p}$, we have by Proposition 3.2, $\lim\limits_{r\to+\infty} r^{l+p}u^{q+1-p}(r) = 0$ or $\lim\limits_{r\to+\infty} r^{l+p}u^{q+1-p}(r) = +\infty$ (because $\dfrac{\alpha}{\beta} < \dfrac{N-p}{p-1} < N$). Then we have also $E'_{\alpha/\beta}(r_0) \neq 0$. Consequently, Remark 2.1 gives $E_{\alpha/\beta}(r) \neq 0$ for any $r > r_0$.

Suppose now that we are in the third case $\dfrac{N-p}{p} \le \dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p} < \dfrac{N-p}{p-1}$. Recall the logarithmic change with $c = \dfrac{\alpha}{\beta}$, then

$$\omega'_{\alpha/\beta}(t) + A_{\alpha/\beta}\,\omega_{\alpha/\beta}(t) + \alpha e^{K_{\alpha/\beta}t}\upsilon_{\alpha/\beta}(t) + \beta e^{K_{\alpha/\beta}t} h_{\alpha/\beta}(t) + \upsilon^q_{\alpha/\beta}(t) = 0 \tag{45}$$

and Proposition 3.4 gives

$$\lim_{t\to+\infty} \upsilon_{\alpha/\beta}(t) = L \quad \text{and} \quad \lim_{t\to+\infty} h_{\alpha/\beta}(t) = \frac{-\alpha}{\beta} L. \tag{46}$$

Define the following energy function:

$$Z(t) = \frac{p-1}{p}\left|h_{\alpha/\beta}(t)\right|^p + \frac{\alpha}{\beta}\omega_{\alpha/\beta}(t)\upsilon_{\alpha/\beta}(t) + \frac{\upsilon^{q+1}_{\alpha/\beta}(t)}{q+1} + \frac{\varrho}{p}\left(\frac{\alpha}{\beta}\right)^{p-1}\upsilon^p_{\alpha/\beta}(t), \tag{47}$$

where

$$\varrho = \frac{\alpha}{\beta} - A_{\alpha/\beta} = \frac{\alpha}{\beta}p - (N-p) \ge 0. \tag{48}$$

According to [7], we have $\lim\limits_{r\to0} ru'(r) = 0$. Therefore

$$\lim_{r\to0} r^{\alpha/\beta}u(r) = \lim_{r\to0} r^{1+\alpha/\beta}u'(r) = 0.$$

It gives

$$\lim_{t\to-\infty} \upsilon_{\alpha/\beta}(t) = \lim_{t\to-\infty} h_{\alpha/\beta}(t) = 0.$$

This implies by (47) that $\lim\limits_{t\to-\infty} Z(t) = 0$.

On the other hand, by a straightforward calculation, the function Z satisfies

$$Z'(t) = \varrho Y(t) - \beta e^{K_{\alpha/\beta}t}\left(h_{\alpha/\beta}(t) + \frac{\alpha}{\beta}\upsilon_{\alpha/\beta}(t)\right)^2, \tag{49}$$

where

$$Y(t) = \left(\left|h_{\alpha/\beta}(t)\right|^{p-2} h_{\alpha/\beta}(t) + \left(\frac{\alpha}{\beta}\right)^{p-1}\upsilon^{p-1}_{\alpha/\beta}(t)\right)\left(h_{\alpha/\beta}(t) + \frac{\alpha}{\beta}\upsilon_{\alpha/\beta}(t)\right). \tag{50}$$

As $\varrho \ge 0$, $\beta < 0$ and the function $s \to |s|^{p-2}s$ is increasing, then $Z'(t) \ge 0$ for any $t \in (-\infty, +\infty)$, that is, Z is increasing on $(-\infty, +\infty)$. Therefore, $Z(t) \ge 0$ for any $t \in (-\infty, +\infty)$. But letting $t \to +\infty$ in (47), we get

$$\lim_{t\to+\infty} Z(t) = A_{\alpha/\beta}\, L^p \left(\frac{\alpha}{\beta}\right)^{p-1}\left(\frac{p-q-1}{p(q+1)}\right) < 0.$$

This is a contradiction. Consequently, u is unbounded. The proof of Theorem 3.1 is complete. □

In the following, we are concerned with the asymptotic behavior of unbounded solutions near infinity.

Theorem 3.2 *Assume $q \geq p-1$ and $\dfrac{N-p}{p-1} > \dfrac{l}{q-1}$. Let u be an unbounded solution of equation (3). Then*

$$\lim_{r\to+\infty} r^{-p/(p-2)}u(r) = \lim_{r\to+\infty} r^{-2/(p-2)}u'(r) = 0. \tag{51}$$

The proof requires some preliminary results.

Proposition 3.5 *Assume $q \geq p-1$ and $\dfrac{N-p}{p-1} > \dfrac{l}{q-1}$. Then there is no solution of equation (3) such that*

$$u(r) > \Gamma r^{-l/(q-1)} \qquad \textit{for large } r, \tag{52}$$

where Γ is given by (13).

Proof. We argue by contradiction and assume that u satisfies (52). Then, according to Lemma 2.1 and Proposition 2.1, we have $u'(r) > 0$ and $E_{l/(q-1)}(r) \neq 0$ for large r. This gives, with logarithmic change (4), that $\omega_{l/(q-1)}(t) > 0$ and $\upsilon'_{l/(q-1)}(t) \neq 0$ for large t. We distinguish two cases.
Case 1: $\upsilon'_{l/(q-1)}(t) < 0$ for large t. Then $\lim\limits_{t\to+\infty} \upsilon_{l/(q-1)}(t) = d \in [\Gamma, +\infty[$.
From equation (5), we have

$$\begin{aligned} \omega'_{l/(q-1)}(t) + A_{l/(q-1)}\,\omega_{l/(q-1)}(t) &= -e^{K_{l/(q-1)}t}\upsilon_{l/(q-1)}(t)\Big[\alpha + \\ &\qquad \beta\frac{h_{l/(q-1)}(t)}{\upsilon_{l/(q-1)}(t)} + \upsilon^{q-1}_{l/(q-1)}(t)\Big]. \end{aligned} \tag{53}$$

Note that

$$\frac{h_{l/(q-1)}(t)}{\upsilon_{l/(q-1)}(t)} = \frac{\upsilon'_{l/(q-1)}(t)}{\upsilon_{l/(q-1)}(t)} - \frac{l}{q-1} < -\frac{l}{q-1}. \tag{54}$$

Then we deduce from (52) and (53) that

$$\omega'_{l/(q-1)}(t) + A_{l/(q-1)}\,\omega_{l/(q-1)}(t) < 0 \quad \text{for large } t.$$

This means that the function $e^{A_{l/(q-1)}t}\omega_{l/(q-1)}(t)$ is decreasing for large t. As $\omega_{l/(q-1)}(t) > 0$ for large t, then $e^{A_{l/(q-1)}t}\omega_{l/(q-1)}(t)$ has a finite limit. Since $A_{l/(q-1)} > 0$, necessarily $\lim\limits_{t\to+\infty} \omega_{l/(q-1)}(t) = 0$. Now, recalling (6), we obtain $\lim\limits_{t\to+\infty} h_{l/(q-1)}(t) = 0$ and $\lim\limits_{t\to+\infty} \upsilon'_{l/(q-1)}(t) = \dfrac{l}{q-1}d < 0$. But this contradicts the fact that $\upsilon_{l/(q-1)}$ is positive.
Case 2: $\upsilon'_{l/(q-1)}(t) > 0$ for large t. Then $\lim\limits_{t\to+\infty} \upsilon_{l/(q-1)}(t) \in]\Gamma, +\infty]$.
(a) Assume that $\lim\limits_{t\to+\infty} \upsilon_{l/(q-1)}(t) = d < +\infty$, then necessarily $\lim\limits_{t\to+\infty} \upsilon'_{l/(q-1)}(t) = 0$. This implies by (6) that $\lim\limits_{t\to+\infty} \omega_{l/(q-1)}(t) = \Big(-\dfrac{l}{q-1}d\Big)^{p-1}$. Therefore,

$$\lim_{t\to+\infty} \Big[\alpha + \beta\frac{h_{l/(q-1)}(t)}{\upsilon_{l/(q-1)}(t)} + \upsilon^{q-1}_{l/(q-1)}(t)\Big] = \alpha - \frac{\beta l}{q-1} + d^{q-1} > 0,$$

by letting $t \to +\infty$ in (53), we obtain $\lim_{t\to+\infty} \omega'_{l/(q-1)}(t) = -\infty$. This is a contradiction with $\omega_{l/(q-1)}$ being positive.

(b) Assume that $\lim_{t\to+\infty} v_{l/(q-1)}(t) = +\infty$, that is, $\lim_{r\to+\infty} r^{l/(q-1)}u(r) = +\infty$. Then, according to Proposition 2.1, $E_{-p/(p-2)}(r) < 0$ for large r. Hence, according to (8) and the fact that $u'(r) > 0$ for large r, we have

$$0 < \frac{ru'(r)}{u} < \frac{p}{p-2} \quad \text{for large } r.$$

On the other hand, we have by equation (3),

$$\Big(|u'|^{p-2}u'\Big)'(r) < -u\Big[\alpha + \beta\frac{ru'}{u} + r^l u^{q-1}\Big].$$

So, $\lim_{r\to+\infty} \left(|u'|^{p-2}u'\right)'(r) = -\infty$, which means that $\lim_{r\to+\infty} u'(r) = -\infty$. This is impossible.

In conclusion, the two cases can not hold, so there is no solution satisfying (52). The proof is complete. □

As a consequence of the previous proposition, we have the following result.

Corollary 3.1 *Assume $q \geq p-1$ and $\dfrac{N-p}{p-1} > \dfrac{l}{q-1}$. Let u be a solution of equation (3). Then $\liminf_{r\to+\infty} r^c u(r) = 0$ for any $c < \dfrac{l}{q-1}$ and $\liminf_{r\to+\infty} r^{l/(q-1)}u(r) \leq \Gamma$.*

Before giving the proof of Theorem 3.2, we need a comparison between the solutions of equation (3) and their derivatives.

Proposition 3.6 *Assume $q \geq p-1$ and $\dfrac{N-p}{p-1} > \dfrac{l}{q-1}$. Let u be an unbounded solution of equation (3). Then*

$$|u'|^{p-2}u'(r) < \max\left(\frac{|\alpha|}{N}, |\beta|\right) ru(r) \quad \textit{for large } r. \tag{55}$$

Proof. Let $\lambda = \min\left(\dfrac{\alpha}{N}, \beta\right) < 0$ and set

$$G(r) = r^{N-1}\Big[|u'|^{p-2}u'(r) + \lambda r\, u(r)\Big]. \tag{56}$$

From equation (3), we have

$$G'(r) = r^{N-1}u(r)\Big[\lambda N - \alpha - r^l|u|^{q-1}\Big] + (\lambda-\beta)r^N u'(r). \tag{57}$$

We will show that $G(r) < 0$ for large r.

Suppose that there exists a large r_0 such that $G(r_0) = 0$, then $G'(r_0)$ can be written in the following form:

$$G'(r_0) = r_0^{N-1}u(r_0)\Big[\lambda N - \alpha - r_0^l u^{q-1}(r_0)\Big] + |\lambda|^{1/(p-1)}(\lambda-\beta)r_0^{N+1/(p-1)}u^{1/(p-1)}(r_0). \tag{58}$$

Therefore, according to the choice of λ, we have $G'(r_0) < 0$. Hence, $G(r) \neq 0$ for any $r > r_0$.
If $G(r) > 0$ for large r, then $u'(r) > 0$ for large r and by a simple integration, we deduce that there exists a constant $c > 0$ such that $r^{-p/(p-2)}u(r) \geq c$ for large r. But this is a contradiction with Proposition 3.5.

In conclusion, $G(r) < 0$ for large r and the proof is complete. □

Proposition 3.7 *Assume $q \geq p-1$ and $\dfrac{N-p}{p-1} > \dfrac{l}{q-1}$. Let u be an unbounded solution of equation (3). Then the functions $r^{-p/(p-2)}u(r)$ and $r^{-2/(p-2)}u'(r)$ are bounded for large r.*

Proof. The proof will be done in two steps.
Step 1: The function $r^{-p/(p-2)}u(r)$ is bounded for large r.
We argue by contradiction and assume that $\upsilon_{-p/(p-2)}(t) = r^{-p/(p-2)}u(r)$ is unbounded for large t, where we use the notation (4). Since $\dfrac{l}{q-1} > \dfrac{-p}{p-2}$, Proposition 3.5 ensures that $\upsilon_{-p/(p-2)}(t)$ can not converge to $+\infty$. Hence, it must necessarily oscillate. Then there exists a sequence $\{\xi_i\}$ going to $+\infty$ as $i \to +\infty$ such that $\upsilon_{-p/(p-2)}$ has a local maximum in ξ_i satisfying $\lim\limits_{i\to+\infty} \upsilon_{-p/(p-2)}(\xi_i) = +\infty$. Since $\upsilon'_{-p/(p-2)}(\xi_i) = 0$, we have

$$h_{-p/(p-2)}(\xi_i) = \frac{p}{p-2}\upsilon_{-p/(p-2)}(\xi_i) > 0,$$

therefore

$$\omega_{-p/(p-2)}(\xi_i) = \left(\frac{p}{p-2}\right)^{p-1} \upsilon^{p-1}_{-p/(p-2)}(\xi_i).$$

On the other hand, estimate (55) can be written in the following form:

$$\omega_{-p/(p-2)}(t) < |\lambda|\, \upsilon_{-p/(p-2)}(t) \quad \text{for large } t, \tag{59}$$

where $|\lambda| = \max\left(\dfrac{|\alpha|}{N}, |\beta|\right)$. In particular, for $t = \xi_i$, we obtain

$$\upsilon^{p-2}_{-p/(p-2)}(\xi_i) < |\lambda| \left(\frac{p-2}{p}\right)^{p-1} \quad \text{for large } i.$$

But this contradicts the fact that $\lim\limits_{i\to+\infty} \upsilon_{-p/(p-2)}(\xi_i) = +\infty$. Consequently, $\upsilon_{-p/(p-2)}(t)$ is bounded for large t.
Step 2: The function $r^{-2/(p-2)}u'(r)$ is bounded for large r, that means that $\omega_{-p/(p-2)}(t)$ is bounded for large t.

Observe preliminary that if $u(r)$ is monotone for large r, necessarily $u'(r) \geq 0$ for large r (because u is unbounded) and then $\omega_{-p/(p-2)}(t) \geq 0$ for large t. Therefore, by (59) and Step 1, $\omega_{-p/(p-2)}(t)$ is bounded for large t.

So, we only have to deal with the case where $u(r)$ is not monotone for large r and then the idea of the proof is the same as for Step 1.

Suppose by contradiction that $\omega_{-p/(p-2)}(t)$ is not bounded and let a sequence $\{k_i\}$ go to $+\infty$ as $i \to +\infty$ such that $\omega'_{-p/(p-2)}(k_i) = 0$ and $\lim\limits_{i\to+\infty} \omega_{-p/(p-2)}(k_i) = +\infty$ or $-\infty$.

First, note that as $\upsilon_{-p/(p-2)}(t)$ is bounded for large t, then (59) implies that we can not have $\lim_{i\to+\infty} \omega_{-p/(p-2)}(k_i) = +\infty$. Secondly, take some constant $\delta > \max\left(\dfrac{N-p}{p-1}, \dfrac{\alpha}{\beta}\right) > 0$, then by Proposition 2.2, $E_\delta(r) > 0$ for large r. So,

$$h_{-p/(p-2)}(t) > -\delta\upsilon_{-p/(p-2)}(t) \quad \text{for large } t. \tag{60}$$

In particular,

$$h_{-p/(p-2)}(k_i) > -\delta\upsilon_{-p/(p-2)}(k_i) \quad \text{for large } i. \tag{61}$$

Again, as $\upsilon_{-p/(p-2)}(t)$ is bounded for large t, we deduce that $\omega_{-p/(p-2)}(k_i)$ is bounded for large i. This is a contradiction. It follows that $\omega_{-p/(p-2)}(t)$ is bounded for large t. Hence $r^{-2/(p-2)}u'(r)$ is bounded for large r. □

Now, we are ready to give the proof of Theorem 3.2.

Proof. (of Theorem 3.2). Using the logarithmic change (4) and Proposition 3.7, we deduce that the functions $\upsilon_{-p/(p-2)}(t)$ and $\omega_{-p/(p-2)}(t)$ are bounded for large t. We proceed in two steps.

Step 1: $\lim_{t\to+\infty} \omega_{-p/(p-2)}(t) = 0$.

First, we claim that $\omega_{-p/(p-2)}(t)$ converges when $t \to +\infty$. Assume by contradiction that $\omega_{-p/(p-2)}(t)$ oscillates, that is, there exist two sequences $\{s_i\}$ and $\{k_i\}$ going to $+\infty$ as $i \to +\infty$ such that $\omega_{-p/(p-2)}(t)$ has a local minimum in s_i and a local maximum in k_i satisfying $s_i < k_i < s_{i+1}$ and

$$\liminf_{t\to+\infty} \omega_{-p/(p-2)}(t) = \lim_{i\to+\infty} \omega_{-p/(p-2)}(s_i) < \limsup_{t\to+\infty} \omega_{-p/(p-2)}(t) = \lim_{i\to+\infty} \omega_{-p/(p-2)}(k_i). \tag{62}$$

Applying equation (17) at the point $t = k_i$, we get $e^{M_{-p/(p-2)}k_i}\upsilon^q_{-p/(p-2)}(k_i)$ is bounded. Therefore, since $M_{-p/(p-2)} > 0$, $\lim_{i\to+\infty} \upsilon_{-p/(p-2)}(k_i) = 0$ and thanks to (59), we deduce

$$\lim_{i\to+\infty} \omega_{-p/(p-2)}(k_i) = \limsup_{t\to+\infty} \omega_{-p/(p-2)}(t) \leq 0.$$

As u is unbounded, then u cannot be decreasing and by (6) and (10) necessarily $\limsup_{t\to+\infty} \omega_{-p/(p-2)}(t) = 0$, so, by (62)

$$\lim_{i\to+\infty} \omega_{-p/(p-2)}(s_i) = \liminf_{t\to+\infty} \omega_{-p/(p-2)}(t) < 0.$$

This implies that $\lim_{i\to+\infty} h_{-p/(p-2)}(s_i) < 0$ and therefore, by (60), we find $\lim_{i\to+\infty} \upsilon_{-p/(p-2)}(s_i) > 0$. On the other hand, $\omega'_{-p/(p-2)}(s_i) = 0$, then equation (17) implies that $\lim_{i\to+\infty} \upsilon_{-p/(p-2)}(s_i) = 0$. This is a contradiction. So, $\omega_{-p/(p-2)}(t)$ is monotone and then it has a finite limit when $t \to +\infty$.

As u is positive and unbounded, then necessarily $\lim_{t\to+\infty} \omega_{-p/(p-2)}(t) \geq 0$. If $\lim_{t\to+\infty} \omega_{-p/(p-2)}(t) > 0$, then, by (59), there exists a constant $C > 0$ such that $\upsilon_{-p/(p-2)}(t) > C$ for large t. Therefore, we obtain by equation (17) that $\lim_{t\to+\infty} \omega'_{-p/(p-2)}(t) = -\infty$, which is a contradiction with the boundedness of $\omega_{-p/(p-2)}$. Consequently, $\lim_{t\to+\infty} \omega_{-p/(p-2)}(t) = 0$. Therefore, $\lim_{r\to+\infty} r^{-2/(p-2)}u'(r) = 0$.

Step 2: $\lim_{t\to+\infty} \upsilon_{-p/(p-2)}(t) = 0$.
Knowing that $\upsilon_{-p/(p-2)}(t)$ is bounded, assume by contradiction that it oscillates, that is, there exist two sequences $\{\eta_i\}$ and $\{\xi_i\}$ going to $+\infty$ as $i \to +\infty$ such that $\upsilon_{-p/(p-2)}$ has a local minimum in η_i and a local maximum in ξ_i satisfying $\eta_i < \xi_i < \eta_{i+1}$ and

$$\liminf_{t\to+\infty} \upsilon_{-p/(p-2)}(t) = \lim_{i\to+\infty} \upsilon_{-p/(p-2)}(\eta_i) < \limsup_{t\to+\infty} \upsilon_{-p/(p-2)}(t) = \lim_{i\to+\infty} \upsilon_{-p/(p-2)}(\xi_i). \tag{63}$$

Since $\upsilon'_{-p/(p-2)}(\eta_i) = \upsilon'_{-p/(p-2)}(\xi_i) = 0$, then we have by (6)

$$h_{-p/(p-2)}(\eta_i) = \frac{p}{p-2}\upsilon_{-p/(p-2)}(\eta_i) \quad \text{and} \quad h_{-p/(p-2)}(\xi_i) = \frac{p}{p-2}\upsilon_{-p/(p-2)}(\xi_i).$$

Since $\lim_{t\to+\infty} h_{-p/(p-2)}(t) = 0$,

$$\lim_{i\to+\infty} h_{-p/(p-2)}(\eta_i) = \lim_{i\to+\infty} h_{-p/(p-2)}(\xi_i) = 0.$$

This implies that

$$\lim_{i\to+\infty} \upsilon_{-p/(p-2)}(\eta_i) = \lim_{i\to+\infty} \upsilon_{-p/(p-2)}(\xi_i) = 0.$$

But this contradicts (63). Therefore, $\upsilon_{-p/(p-2)}$ converges. Hence, by (6), $\upsilon'_{-p/(p-2)}$ converges necessarily to 0. Consequently, $\lim_{t\to+\infty} \upsilon_{-p/(p-2)}(t) = 0$, i.e., $\lim_{r\to+\infty} r^{-p/(p-2)}u(r) = 0$. The proof is complete. □

Theorem 3.3 *Assume* $q \geq p\Big(2 + 2^{p-1}\Big) - 1$ *and* $\dfrac{l}{q-1} < \min\left(\dfrac{-\alpha}{\beta}, \dfrac{N-p}{p-1}\right)$. *Let* u *be an unbounded positive solution of problem* $(\mathbf{P})$. *Then*

$$\lim_{r\to+\infty} r^{l/(q-1)}u(r) = \Gamma \tag{64}$$

and

$$\lim_{r\to+\infty} r^{l/(q-1)+1}u'(r) = \frac{-l}{q-1}\Gamma, \tag{65}$$

where Γ *is given by (13).*

To prove this theorem we will need the following results.

Proposition 3.8 *Assume* $q \geq p-1$ *and* $\dfrac{N-p}{p-1} > \dfrac{l}{q-1}$. *Let* u *be an unbounded positive solution of problem* $(\mathbf{P})$. *Then, for any* $c > \left(\dfrac{\alpha}{\beta}\right)^{p-1}$,

$$|u'|^{p-2}u'(r) > -cr^{1-p}u^{p-1}(r) \quad \textit{for large } r. \tag{66}$$

Proof. Let $c > \left(\dfrac{\alpha}{\beta}\right)^{p-1}$ and

$$F(r) = r^{N-1}|u'|^{p-2}u'(r) + cr^{N-p}u^{p-1}(r). \tag{67}$$

Then, according to equation (3),

$$F'(r) = r^{N-1}u\Big[-\alpha - r^l u^{q-1} + c(N-p)r^{-p}u^{p-2}\Big] + r^N u'\Big[-\beta + c(p-1)r^{-p}u^{p-2}\Big]. \quad (68)$$

We will show that $F(r) > 0$ for large r.

First, we prove that $F(r) \neq 0$ for large r. Suppose that there exists a large r_0 such that $F(r_0) = 0$. Then, according to (67) and (68), we have

$$F'(r_0) = r_0^{N-1}u\Big[-\alpha + \beta c^{1/(p-1)} - r_0^l u^{q-1} + \Big(c(N-p) - c^{p/(p-1)}(p-1)\Big)r_0^{-p}u^{p-2}\Big]. \quad (69)$$

Since $-\alpha + \beta c^{1/(p-1)} < 0$, $r^l u^{q-1}(r) > 0$ and $\lim\limits_{r\to+\infty} r^{-p}u^{p-2}(r) = 0$ (by Theorem 3.2), we get $F'(r_0) < 0$. Hence, $F(r) \neq 0$ for any $r > r_0$ and then, as u is unbounded, $F(r)$ cannot be negative for large r. The proof is complete. □

Proposition 3.9 *Assume $q \geq p\Big(2+2^{p-1}\Big)-1$ and $\dfrac{l}{q-1} < \min\left(\dfrac{-\alpha}{\beta}, \dfrac{N-p}{p-1}\right)$. Let u be an unbounded positive solution of problem* $(\mathbf{P})$*. Then the functions $r^{l/(q-1)}u(r)$ and $r^{l/(q-1)+1}u'(r)$ are bounded for large r.*

Proof. (i) We first show that $r^{l/(q-1)}u(r)$ is bounded for large r.
We make the change (4) for $c = \dfrac{l}{q-1}$ and we set

$$D = \frac{K_{l/(q-1)}(p-1)}{p} - \frac{l}{q-1} + A_{l/(q-1)} > 0, \quad (70)$$

We define, for any real $\theta > 0$, the following energy function:

$$\begin{aligned} F_\theta(t) \quad &= \frac{p-1}{p}e^{-K_{l/(q-1)}t}\left|h_{l/(q-1)}(t)\right|^p - \frac{\Gamma^{q-1}}{2}v^2_{l/(q-1)}(t) + \frac{v^{q+1}_{l/(q-1)}(t)}{q+1} + \\ &\frac{l}{q-1}e^{-K_{l/(q-1)}t}\omega_{l/(q-1)}(t)v_{l/(q-1)}(t) - \frac{\theta}{p}\left(\frac{-l}{q-1}\right)^{p-1}e^{-K_{l/(q-1)}t}v^p_{l/(q-1)}(t). \end{aligned} \quad (71)$$

Using equation (5), a straightforward calculation gives

$$\begin{aligned} F'_\theta(t) \quad = \quad &\theta e^{-K_{l/(q-1)}t}X(t) - \beta\left(h_{l/(q-1)}(t) + \frac{l}{q-1}v_{l/(q-1)}(t)\right)^2 + \\ &D_1 e^{-K_{l/(q-1)}t}\left|h_{l/(q-1)}(t)\right|^p + D_2 e^{-K_{l/(q-1)}t}\omega_{l/(q-1)}(t)v_{l/(q-1)}(t) + \\ &\frac{K_{l/(q-1)}\theta}{p}\left(\frac{-l}{q-1}\right)^{p-1}e^{-K_{l/(q-1)}t}v^p_{l/(q-1)}(t), \end{aligned} \quad (72)$$

where

$$\begin{aligned} X(t) \quad = \quad &\left[\left|h_{l/(q-1)}(t)\right|^{p-2}h_{l/(q-1)}(t) - \left(\frac{-l}{q-1}\right)^{p-1}v^{p-1}_{l/(q-1)}(t)\right]\Big[h_{l/(q-1)}(t) + \\ &\frac{l}{q-1}v_{l/(q-1)}(t)\Big], \end{aligned} \quad (73)$$

$$D_1 = -(D+\theta) < 0 \tag{74}$$

and

$$D_2 = \frac{-l}{q-1}\left(D+\theta+\frac{K_{l/(q-1)}}{p}\right) > 0. \tag{75}$$

We show in two steps that $\upsilon_{l/(q-1)}(t)$ is bounded for large t.
Step 1: $F'_\theta(t) > 0$ for large t by choosing a suitable θ.
We know that for any $\rho \geq 2$ and for any $a, b \in \mathbb{R}$, we have

$$|a-b|^\rho \leq 2^{\rho-1}\left(|a|^\rho + |b|^\rho\right).$$

Therefore, particulary for $\rho = p$, $a = \upsilon'_{l/(q-1)}(t)$ and $b = \dfrac{l}{q-1}\upsilon_{l/(q-1)}(t)$, we obtain

$$\begin{aligned}\left|h_{l/(q-1)}(t)\right|^p &= \left|\upsilon'_{l/(q-1)}(t) - \frac{l}{q-1}\upsilon_{l/(q-1)}(t)\right|^p \\ &\leq 2^{p-1}\left(\left|\upsilon'_{l/(q-1)}(t)\right|^p + \left(\frac{|l|}{q-1}\right)^p \upsilon^p_{l/(q-1)}(t)\right).\end{aligned}$$

Since $D_1 < 0$,

$$\begin{aligned} F'_\theta(t) \geq\ & \theta e^{-K_{l/(q-1)}t}X(t) + \upsilon'^2_{l/(q-1)}(t)\left[-\beta + 2^{p-1}D_1 e^{-K_{l/(q-1)}t}\left|\upsilon'_{l/(q-1)}\right|^{p-2}\right] + \\ & e^{-K_{l/(q-1)}t}\upsilon^p_{l/(q-1)}(t)\Bigg[2^{p-1}D_1\left(\frac{|l|}{q-1}\right)^p + \frac{K_{l/(q-1)}\theta}{p}\left(\frac{|l|}{q-1}\right)^{p-1} + \\ & D_2\,\omega_{l/(q-1)}\upsilon^{1-p}_{l/(q-1)}\Bigg]. \end{aligned} \tag{76}$$

As $\dfrac{|l|}{q-1} > \dfrac{\alpha}{\beta}$, we have by Proposition 3.8,

$$\omega_{l/(q-1)}(t)\upsilon^{1-p}_{l/(q-1)}(t) > -\left(\frac{|l|}{q-1}\right)^{p-1} \quad \text{for large } t. \tag{77}$$

Therefore, according to (76) we get

$$\begin{aligned} F'_\theta(t) \geq\ & \theta e^{-K_{l/(q-1)}t}X(t) + \upsilon'^2_{l/(q-1)}(t)\left[-\beta + 2^{p-1}D_1 e^{-K_{l/(q-1)}t}\left|\upsilon'_{l/(q-1)}\right|^{p-2}\right] + \\ & e^{-K_{l/(q-1)}t}\upsilon^p_{l/(q-1)}\Bigg[2^{p-1}D_1\left(\frac{|l|}{q-1}\right)^p + \frac{K_{l/(q-1)}\theta}{p}\left(\frac{|l|}{q-1}\right)^{p-1} - \\ & D_2\left(\frac{|l|}{q-1}\right)^{p-1}\Bigg]. \end{aligned} \tag{78}$$

Now, we choose in (71)

$$\theta = \frac{2^p|l|pD + 2|l|pD + 2|l|K_{l/(q-1)}}{K_{l/(q-1)}(q-1) + lp(2^{p-1}+1)} > 0. \tag{79}$$

Note that, since $q \geq p\Big(2+2^{p-1}\Big)-1$, we have $K_{l/(q-1)}(q-1)+lp(2^{p-1}+1)>0$, and the expression (79) gives

$$2^{p-1}D_1\left(\frac{|l|}{q-1}\right)^p+\frac{K_{l/(q-1)}\theta}{p}\left(\frac{|l|}{q-1}\right)^{p-1}-D_2\left(\frac{|l|}{q-1}\right)^{p-1}>0. \tag{80}$$

Moreover,

$$e^{-K_{l/(q-1)}t/(p-2)}\upsilon'_{l/(q-1)}(t)=e^{-K_{l/(q-1)}t/(p-2)}\left(h_{l/(q-1)}(t)+\frac{l}{q-1}\upsilon_{l/(q-1)}(t)\right), \tag{81}$$

and from Theorem 3.2,

$$\lim_{t\to+\infty}e^{-K_{l/(q-1)}t/(p-2)}\upsilon_{l/(q-1)}(t)=\lim_{r\to+\infty}r^{-p/(p-2)}u(r)=0 \tag{82}$$

and

$$\lim_{t\to+\infty}e^{-K_{l/(q-1)}t/(p-2)}h_{l/(q-1)}(t)=\lim_{r\to+\infty}r^{-2/(p-2)}u'(r)=0, \tag{83}$$

then

$$\lim_{t\to+\infty}e^{-K_{l/(q-1)}t/(p-2)}\upsilon'_{l/(q-1)}(t)=0. \tag{84}$$

Using the last equality, the estimate (80), the fact that $X(t)\geq 0$ (because the function $s\to|s|^{p-2}s$ is increasing), $\upsilon_{l/(q-1)}(t)>0$, $\theta>0$ and $\beta<0$, we deduce from (78) that $F'_\theta(t)>0$ for large t.

Step 2: The function $\upsilon_{l/(q-1)}(t)$ cannot oscillate about a constant B such that $B>(q+1)^{1/(q-1)}\Gamma>\Gamma$, where Γ is given by (13).

We argue by contradiction and assume that there exist two sequences $\{\eta_i\}$ and $\{\xi_i\}$ going to $+\infty$ as $i\to+\infty$ such that $\upsilon_{l/(q-1)}$ has a local minimum in η_i and a local maximum in ξ_i satisfying $\eta_i<\xi_i<\eta_{i+1}$ and $\upsilon_{l/(q-1)}(\xi_i)>B$.

Since $h_{l/(q-1)}(\eta_i)=-\dfrac{l}{q-1}\upsilon_{l/(q-1)}(\eta_i)>0$, we see that $h'_{l/(q-1)}(\eta_i)$ exists and, more exactly, $h'_{l/(q-1)}(\eta_i)=\upsilon''_{l/(q-1)}(\eta_i)\geq 0$ (because $\upsilon'_{l/(q-1)}(\eta_i)=0$), which implies that $\omega'_{l/(q-1)}(\eta_i)\geq 0$. Taking $c=\dfrac{l}{q-1}$ in (4), we obtain

$$\begin{aligned} e^{-K_{l/(q-1)}t}\omega'_{l/(q-1)}(t) \;&=\; -A_{l/(q-1)}e^{-K_{l/(q-1)}t}\omega_{l/(q-1)}(t)-\alpha\upsilon_{l/(q-1)}(t)\\ &\quad -\beta h_{l/(q-1)}(t)-\upsilon^q_{l/(q-1)}(t). \end{aligned} \tag{85}$$

So, for $t=\eta_i$

$$e^{-K_{l/(q-1)}\eta_i}\omega'_{l/(q-1)}(\eta_i)<-\psi(\upsilon_{l/(q-1)}(\eta_i)),$$

where

$$\psi(s)=\left[s^{q-1}-\Gamma^{q-1}\right]s,\quad s\geq 0. \tag{86}$$

Since $\omega'_{l/(q-1)}(\eta_i)\geq 0$, one has $\psi(\upsilon_{l/(q-1)}(\eta_i))<0$ and therefore, necessarily $\upsilon_{l/(q-1)}(\eta_i)<\Gamma$. On the other hand, according to (71), we have

$$F_\theta(\eta_i)=\frac{\upsilon^{q+1}_{l/(q-1)}(\eta_i)}{q+1}+\upsilon^2_{l/(q-1)}(\eta_i)\left[-\frac{\Gamma^{q-1}}{2}+C_2e^{-K_{l/(q-1)}\eta_i}\upsilon^{p-2}_{l/(q-1)}(\eta_i)\right], \tag{87}$$

where $C_2 = \dfrac{1}{p}\left(\dfrac{-l}{q-1}\right)^{p-1}\left(\dfrac{l}{q-1}-\theta\right) < 0$. In particular, we obtain $F_\theta(\eta_i) < \phi_1(\upsilon_{l/(q-1)}(\eta_i))$, where

$$\phi_1(s) = \frac{s^{q+1}}{q+1} - \frac{\Gamma^{q-1}}{2}s^2, \quad s \geq 0. \tag{88}$$

Therefore, since $0 < \upsilon_{l/(q-1)}(\eta_i) < \Gamma$, a simple study of the function ϕ_1 gives $\phi_1(\upsilon_{l/(q-1)}(\eta_i)) < 0$. Consequently, $F_\theta(\eta_i) < 0$ for large i.

In the same way, since $\upsilon'_{l/(q-1)}(\xi_i) = 0$,

$$F_\theta(\xi_i) = \frac{\upsilon^{q+1}_{l/(q-1)}(\xi_i)}{q+1} + \upsilon^2_{l/(q-1)}(\xi_i)\left[-\frac{\Gamma^{q-1}}{2} + C_2 e^{-K_{l/(q-1)}\xi_i}\upsilon^{p-2}_{l/(q-1)}(\xi_i)\right]. \tag{89}$$

Since $\lim\limits_{i\to+\infty} e^{-K_{l/(q-1)}\xi_i}\upsilon^{p-2}_{l/(q-1)}(\xi_i) = 0$ by (82),

$$F_\theta(\xi_i) > \frac{\upsilon^{q+1}_{l/(q-1)}(\xi_i)}{q+1} - \Gamma^{q-1}\upsilon^2_{l/(q-1)}(\xi_i) \quad \text{for large } i. \tag{90}$$

Set

$$\phi_2(s) = \frac{s^{q+1}}{q+1} - \Gamma^{q-1}s^2, \quad s \geq 0. \tag{91}$$

Therefore, by (90), $F_\theta(\xi_i) > \phi_2(\upsilon_{l/(q-1)}(\xi_i))$ for large i. Since $\upsilon_{l/(q-1)}(\xi_i) > B > (q+1)^{1/(q-1)}\Gamma$, we have $\phi_2(\upsilon_{l/(q-1)}(\xi_i)) > 0$, which implies that $F_\theta(\xi_i) > 0$ for large i. Hence, $F_\theta(\eta_i) < 0$ and $F_\theta(\xi_i) > 0$, for large i, which clearly contradicts the monotonicity of $F_\theta(t)$ for large t. It follows that $\upsilon_{l/(q-1)}(t)$ cannot oscillate about the constant B. Moreover, since $\upsilon_{l/(q-1)}(t)$ cannot stay above B (from Proposition 3.5), one has $\upsilon_{l/(q-1)}(t) \leq B$ for large t. Consequently, $\upsilon_{l/(q-1)}(t)$ is bounded for large t. That is, $r^{l/(q-1)}u(r)$ is bounded for large r.

(ii) Now, we show that $r^{l/(q-1)+1}u'(r)$ is bounded for large r, i.e., by (6) and (10), $\omega_{l/(q-1)}(t)$ is bounded for large t.

We argue by contradiction. As u is positive and unbounded, then we have two possibilities.

- $\lim\limits_{t\to+\infty}\omega_{l/(q-1)}(t) = +\infty$, then $\lim\limits_{t\to+\infty} h_{l/(q-1)}(t) = +\infty$. Using (6) and the fact that $\upsilon_{l/(q-1)}(t)$ is bounded for large t, we obtain $\lim\limits_{t\to+\infty}\upsilon'_{l/(q-1)}(t) = +\infty$ and therefore, $\lim\limits_{t\to+\infty}\upsilon_{l/(q-1)}(t) = +\infty$, which is impossible.
- There exists a sequence $\{k_i\}$ going to $+\infty$ as $i \to +\infty$ such that $\omega_{l/(q-1)}$ has a local extremum in k_i satisfying $\lim\limits_{i\to+\infty}\omega_{l/(q-1)}(k_i) = +\infty$ or $-\infty$.

We use expression (6), equation (85) and the fact that $\omega'_{l/(q-1)}(k_i) = 0$, then

$$\alpha\upsilon_{l/(q-1)}(k_i) + \upsilon^q_{l/(q-1)}(k_i) = h_{l/(q-1)}(k_i)\left[-\beta - A_{l/(q-1)}e^{-K_{l/(q-1)}k_i}|h_{l/(q-1)}(k_i)|^{p-2}\right]. \tag{92}$$

Since $\lim\limits_{t\to+\infty} e^{-K_{l/(q-1)}t}|h_{l/(q-1)}(t)|^{p-2} = 0$ by (83), $\lim\limits_{i\to+\infty} h_{l/(q-1)}(k_i) = +\infty$ or $-\infty$ and $\beta < 0$, we have

$$\lim_{i\to+\infty}\alpha\upsilon_{l/(q-1)}(k_i) + \upsilon^q_{l/(q-1)}(k_i) = +\infty$$

or

$$\lim_{i\to+\infty} \alpha\upsilon_{l/(q-1)}(k_i) + \upsilon^q_{l/(q-1)}(k_i) = -\infty.$$

But this contradicts the fact that $\upsilon_{l/(q-1)}(t)$ is bounded for large t.
Hence $\omega_{l/(q-1)}(t)$ is bounded for large t and therefore $r^{l/(q-1)+1}u'(r)$ is bounded for large r. □

We can now give the proof of Theorem 3.3.

Proof. (of Theorem 3.3). We use the logarithmic change, we have by Proposition 3.9, $\upsilon_{l/(q-1)}(t)$ and $h_{l/(q-1)}(t)$ are bounded for large t. The proof will be done in three steps.

Step 1: The function $\upsilon_{l/(q-1)}(t)$ converges.
We argue by contradiction and assume that $\upsilon_{l/(q-1)}$ oscillates, that is, there exist two sequences $\{\eta_i\}$ and $\{\xi_i\}$ going to $+\infty$ as $i\to+\infty$ such that $\upsilon_{l/(q-1)}$ has a local minimum in η_i and a local maximum in ξ_i satisfying $\eta_i < \xi_i < \eta_{i+1}$ and

$$\begin{aligned}\liminf_{t\to+\infty} \upsilon_{l/(q-1)}(t) = \lim_{i\to+\infty} \upsilon_{l/(q-1)}(\eta_i) = m_1 < \limsup_{t\to+\infty} \upsilon_{l/(q-1)}(t) = \lim_{i\to+\infty} \upsilon_{l/(q-1)}(\xi_i) \\ = M_1. \qquad (93)\end{aligned}$$

On the other hand, we know by the proof of Proposition 3.9 that $F'_\theta(t) > 0$ for large t where θ and F_θ are given, respectively, by (79) and (71). Then $F_\theta(t) \neq 0$ for large t. We show that $F_\theta(t) < 0$ for large t.
If $F_\theta(t) > 0$ for large t, then, since $F'_\theta(t) > 0$ for large t, $\lim_{t\to+\infty} F_\theta(t) \in]0,+\infty]$.
Using the fact that

$$\lim_{i\to+\infty} F_\theta(\eta_i) = \phi_1(m_1) < +\infty \tag{94}$$

and

$$\lim_{i\to+\infty} F_\theta(\xi_i) = \phi_1(M_1) < +\infty, \tag{95}$$

where ϕ_1 is given by (88), we see that $\lim_{t\to+\infty} F_\theta(t)$ is finite and strictly positive. More exactly, we have

$$\lim_{t\to+\infty} F_\theta(t) = \phi_1(m_1) = \phi_1(M_1) > 0.$$

But this contradicts the fact that $\phi_1(m_1) \le 0$ because by Corollary 3.1, we have $0 \le \liminf_{t\to+\infty} \upsilon_{l/(q-1)}(t) = m_1 \le \Gamma$. We deduce that $F_\theta(t) < 0$ for large t.

Since $F'_\theta(t) > 0$ for large t, one has $\lim_{t\to+\infty} F_\theta(t)$ is finite and negative. Therefore, according to (94) and (95), we have

$$\lim_{t\to+\infty} F_\theta(t) = \phi_1(m_1) = \phi_1(M_1) \le 0. \tag{96}$$

Set $L_1 = \lim_{t\to+\infty} F_\theta(t)$. Then

$$L_1 = \phi_1(m_1) = \phi_1(M_1).$$

Therefore, there exist $\gamma \in (m_1, M_1)$ and $t_i \in (\eta_i, \xi_i)$ such that $\upsilon_{l/(q-1)}(t_i) = \gamma$, $\phi'_1(\gamma) = 0$ and $\phi_1(\gamma) \neq L_1$.

On the other hand, $\upsilon_{l/(q-1)}(t)$, $h_{l/(q-1)}(t)$, $\omega_{l/(q-1)}(t)$ are bounded for large t and $\upsilon_{l/(q-1)}(t_i) = \gamma$, we get by (71), $\lim_{i\to+\infty} F_\theta(t_i) = \phi_1(\gamma)$, hence $\phi_1(\gamma) = L_1$. But this

contradicts the fact that $\phi_1(\gamma) \neq L_1$.
Consequently, $\upsilon_{l/(q-1)}$ converges. Set $\lim\limits_{t\to+\infty} \upsilon_{l/(q-1)}(t) = d \geq 0$.

Step 2: The function $h_{l/(q-1)}(t)$ converges.
According to (6), it suffices to show that $\omega_{l/(q-1)}$ converges. Since $h_{l/(q-1)}(t)$ is bounded for large t, we see that $\omega_{l/(q-1)}(t)$ is also bounded for large t. Assume by contradiction that it oscillates, that is, there exist two sequences $\{s_i\}$ and $\{k_i\}$ going to $+\infty$ as $i \to +\infty$ such that $\omega_{l/(q-1)}$ has a local minimum in s_i and a local maximum in k_i satisfying $s_i < k_i < s_{i+1}$ and

$$\liminf_{t\to+\infty} \omega_{l/(q-1)}(t) = \lim_{i\to+\infty} \omega_{l/(q-1)}(s_i) < \limsup_{t\to+\infty} \omega_{l/(q-1)}(t) = \lim_{i\to+\infty} \omega_{l/(q-1)}(k_i). \quad (97)$$

Using equation (85), the fact that $\omega_{l/(q-1)}(t)$ is bounded for large t, $\lim\limits_{t\to+\infty} \upsilon_{l/(q-1)}(t) = d$ and $\omega'_{l/(q-1)}(s_i) = \omega'_{l/(q-1)}(k_i) = 0$, we obtain

$$\lim_{i\to+\infty} -\beta h_{l/(q-1)}(s_i) = \lim_{i\to+\infty} -\beta h_{l/(q-1)}(k_i) = \alpha d + d^q.$$

Since $\beta < 0$,

$$\lim_{i\to+\infty} h_{l/(q-1)}(s_i) = \lim_{i\to+\infty} h_{l/(q-1)}(k_i),$$

which gives

$$\lim_{i\to+\infty} \omega_{l/(q-1)}(s_i) = \lim_{i\to+\infty} \omega_{l/(q-1)}(k_i).$$

But this contradicts (97). Hence, $\omega_{l/(q-1)}$ converges and therefore $h_{l/(q-1)}$ converges.
According to (6), $\lim\limits_{t\to+\infty} \upsilon'_{l/(q-1)}(t)$ exists and must be 0. Hence,

$$\lim_{t\to+\infty} h_{l/(q-1)}(t) = \frac{-l}{q-1} d. \quad (98)$$

Step 3: $\lim\limits_{t\to+\infty} \upsilon_{l/(q-1)}(t) = \Gamma$.
Combining equation (85), expression of Γ given by (13), Step 1 and Step 2, we get

$$\lim_{t\to+\infty} e^{-K_{l/(q-1)}t} \omega'_{l/(q-1)}(t) = d\left(\Gamma^{q-1} - d^{q-1}\right).$$

Since $\omega_{l/(q-1)}$ converges and $\lim\limits_{t\to+\infty} e^{-K_{l/(q-1)}t} \omega'_{l/(q-1)}(t)$ exists, necessarily $\lim\limits_{t\to+\infty} e^{-K_{l/(q-1)}t} \omega'_{l/(q-1)}(t) = 0$. Therefore, $d\left(\Gamma^{q-1} - d^{q-1}\right) = 0$. We claim that $d = \Gamma$.

Assume by contradiction that $d = \lim\limits_{t\to+\infty} \upsilon_{l/(q-1)}(t) = 0$. First, we prove that $\omega_{l/(q-1)}(t) \neq 0$ for large t.

For any large T such that $\omega_{l/(q-1)}(T) = 0$, we have $h_{l/(q-1)}(T) = 0$ and by equation (85)

$$e^{-K_{l/(q-1)}T} \omega'_{l/(q-1)}(T) = \upsilon_{l/(q-1)}(T) \left[-\alpha - \upsilon^{q-1}_{l/(q-1)}(T)\right].$$

Since $\alpha < 0$, $\upsilon_{l/(q-1)}(t) > 0$ and $\lim\limits_{t\to+\infty} \upsilon^{q-1}_{l/(q-1)}(t) = 0$, one has $\omega'_{l/(q-1)}(T) > 0$. Hence, $\omega_{l/(q-1)}(t) \neq 0$ for any $t > T$.

As u is positive and unbounded, then $\omega_{l/(q-1)}(t) > 0$ for large t.
On the other hand, we have by (85)

$$e^{-K_{l/(q-1)}t}\omega'_{l/(q-1)}(t) = h_{l/(q-1)}(t)\left[-\beta - A_{l/(q-1)}e^{-K_{l/(q-1)}t}\left|h_{l/(q-1)}(t)\right|^{p-2}\right] + \upsilon_{l/(q-1)}(t)\left[-\alpha - \upsilon^{q-1}_{l/(q-1)}(t)\right]. \tag{99}$$

Using the fact that $\beta < 0$, $\alpha < 0$, $\lim_{t\to+\infty} e^{-K_{l/(q-1)}t}\left|h_{l/(q-1)}(t)\right|^{p-2} = 0$ (by (83)), $\lim_{t\to+\infty} \upsilon^{q-1}_{l/(q-1)}(t) = 0$, $\upsilon_{l/(q-1)}(t) > 0$ and $h_{l/(q-1)}(t) > 0$ for large t, we have $\omega'_{l/(q-1)}(t) > 0$ for large t. This implies, since $\omega_{l/(q-1)}(t) > 0$ for large t, that $\lim_{t\to+\infty} \omega_{l/(q-1)}(t) \in]0,+\infty]$. But this contradicts the fact that $\lim_{t\to+\infty} \omega_{l/(q-1)}(t) = 0$ by (6) and (98).
Consequently, $d = \Gamma$. It follows that

$$\lim_{t\to+\infty} \upsilon_{l/(q-1)}(t) = \Gamma \quad \text{and} \quad \lim_{t\to+\infty} h_{l/(q-1)}(t) = \frac{-l}{q-1}\Gamma.$$

The proof is complete. □

The behavior of unbounded solution is illustrated by the following figure.

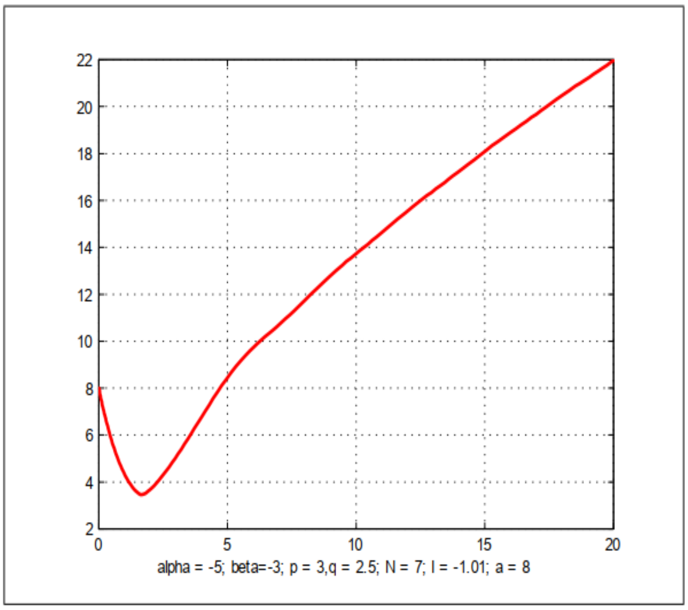


Figure 2: Unbounded solution.

To finish this work, we note that if we have the monotonicity of u, then the asymptotic behavior (64) and (65) is validated by reducing the assumptions of Theorem 3.3. More precisely, we have the following result.

Proposition 3.10 *Assume* $q \geq p\Big(1+2^{p-1}\Big)-1$ *and* $\dfrac{N-p}{p-1} > \dfrac{l}{q-1}$. *Let* u *be an unbounded positive solution of problem* $(\mathbf{P})$. *If* u *is an increasing function for large* r, *then it satisfies (64) and (65).*

Proof. First of all we note that the idea of the proof is similar to that of Theorem 3.3. So, we follow the same steps and we only change the step where we use the monotonicity of u.

According to the proof of Propostition 3.9, we choose a suitable θ to show that $F_\theta(t)$ is strictly increasing for large t, where F_θ is given by (71). In fact, we choose in (71)

$$\theta = \frac{2^p|l|pD}{K_{l/(q-1)}(q-1)+lp2^{p-1}} > 0. \tag{100}$$

Note that, since $q \geq p\Big(1+2^{p-1}\Big)-1$, one has $K_{l/(q-1)}(q-1)+lp2^{p-1} > 0$ and by expression (100),

$$2^{p-1}D_1\left(\frac{|l|}{q-1}\right)^p + \frac{K_{l/(q-1)}\theta}{p}\left(\frac{|l|}{q-1}\right)^{p-1} = \left(\frac{|l|}{q-1}\right)^{p-1}\frac{2^{p-1}|l|D}{q-1} > 0.$$

According to inequality (76) and using expressions (100) and (75) and the fact that $\omega_{l/(q-1)}(t) \geq 0$ for large t (because $u'(r) \geq 0$ for large r), we deduce that

$$\begin{aligned} F'_\theta(t) \geq\ & \theta e^{-K_{l/(q-1)}t}X(t) + \upsilon'^2_{l/(q-1)}(t)\left[-\beta + 2^{p-1}D_1 e^{-K_{l/(q-1)}t}\left|\upsilon'_{l/(q-1)}\right|^{p-2}\right] + \\ & \left(\frac{|l|}{q-1}\right)^{p-1}\frac{2^{p-1}|l|D}{q-1}e^{-K_{l/(q-1)}t}\upsilon^p_{l/(q-1)}. \end{aligned} \tag{101}$$

Using (84), the fact that $X(t) \geq 0$, $\upsilon_{l/(q-1)}(t) > 0$, $\theta > 0$ and $\beta < 0$, we deduce from estimate (101) that $F'_\theta(t) > 0$ for large t. We complete the proof in the same way as that of Theorem 3.3. □

4 Conclusion

We consider equation (1) as a natural generalization of the pure Laplacian case ($p = 2$) already studied by Filippas and Tertikas in [8]. Its appears in studying the self-similar solution of the parabolic equation (2). This equation admits a family of radial self-similar solutions defined in the form

$$v(x,t) = t^{-\alpha}u(t^{-\beta}\,|x|),$$

where u is the solution of equation (3) with

$$\alpha = \frac{l+p}{p(q-1)+l(p-2)}, \quad \beta = \frac{q+1-p}{p(q-1)+l(p-2)}.$$

There is an extensive literature on equation (2) in the case $l = 0$. The study of equation (1) in this case is a reasonable first step towards the understanding of the behavior of blowing up solutions of (2).

Taking the case $l<0$, we have proven in [7] the existence of an entire solution u of (3). In this paper we have proven that if $q \le p-1$ or $q > p-1$ and $\dfrac{\alpha}{\beta} \neq \dfrac{l+p}{q+1-p}$ or $\dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p} \geq \dfrac{N-p}{p}$, this solution is unbounded. The study of its asymptotic behavior depends strongly on the study of the following nonlinear dynamical system by using the logarithmic change (4), $\upsilon_{l/(q-1)}(t) = r^{l/(q-1)}u(r)$,

$$\begin{cases} \upsilon'_{l/(q-1)}(t) = & |\omega_{l/(q-1)}|^{(2-p)/(p-1)}\omega_{l/(q-1)}(t) + \dfrac{l}{q-1}\upsilon_{l/(q-1)}(t), \\ \omega'_{l/(q-1)}(t) = & -A_{l/(q-1)}\omega_{l/(q-1)}(t) - \alpha e^{K_{l/(q-1)}t}\upsilon_{l/(q-1)}(t) - \beta e^{K_{l/(q-1)}t}h_{l/(q-1)}(t) - \\ & e^{K_{l/(q-1)}t}\upsilon^q_{l/(q-1)}(t). \end{cases}$$

It is shown that under some assumptions, the solution $(\upsilon_{l/(q-1)}, \omega_{l/(q-1)})$ of the above system tends to the equilibrium point $\left(\Gamma, \left(\dfrac{-l}{q-1}\Gamma\right)^{p-1}\right)$, where Γ is given by (13). This result can be translated in terms of u and u' by

$$u(r) \underset{+\infty}{\sim} \Gamma\, r^{-l/(q-1)}$$

and

$$|u'|^{p-2}u'(r) \underset{+\infty}{\sim} \left(\frac{-l}{q-1}\Gamma\right)^{p-1} r^{-(l/(q-1)+1)(p-1)}.$$

The more complicated case $\dfrac{\alpha}{\beta} = \dfrac{l+p}{q+1-p} < \dfrac{N-p}{p}$ has not been completely investigated.

Acknowledgment

The authors thank the reviewer for several remarks on a preliminary version of the paper.

References

[1] H. Amann. *Ordinary Differential Equations.* Walter de Gruyter. Belin. NewYork, 1996.

[2] S. Bae. Asymptotic behavior of positive solutions of inhomogeneous semilinear elliptic equations. *Nonlinear Analysis.* **51** (2002) 1373–1403.

[3] B. Bettioui and A. Gmira. On the radial solutions of a degenerate quasilinear elliptic equation in $\mathbb{R}^{\mathbb{N}}$. *Ann. Fac. Sci. Toulouse* **8** (3) (1999) 411–438.

[4] M.F. Bidaut-Véron. Local and global behavior of solutions of quasilinear equations of Emden-Fowler type. *Archive for Rational Mechanics and Analysis* **107** (4) (1989) 293–324.

[5] M. F. Bidaut-Véron. The p-Laplace heat equation with a source term: self-similar solutions revisited. *Advanced Nonlinear Studies* **6** (2006) 69–108.

[6] M. F. Bidaut-Véron. Self-similar solutions of the p-Laplace heat equation: the fast diffusion case. *Pacific J. Math.* **2** (227) (2006) 201–269.

[7] A. Bouzelmate and A. Gmira. On the radial solutions of a nonlinear singular elliptic equation. *International Journal of Mathematical Analysis.* **9** (26) (2015) 1279–1297.

[8] S. Filippas and A. Tertikas. On similarity solutions of heat equation with a nonhomogeneous nonlinearity. *J. Diff. Equ.* **165** (2000) 468–492.

[9] B. Hamri, A. Gmira and Y. Jabri. Existence and uniqueness of selfsimilar solutions of a nonhomogeneous equation. *Intenational Journal of Pure and Applied Mathematics.* **86** (5) (2013) 849–857.

[10] A. Haraux and F. B. Weissler. Non-uniqueness for a semilinear initial value problem. *Indiana Univ. Math. J.* **31** (1982) 167–189.

[11] M. Hirose and E. Yanagida. Global structure of self-similar solutions in a semilinear parabolic equation. *J. Mathematical Analysis and Applications.* **244** (2000) 348–368.

[12] G. Leoni. A classification of positive solutions on the problem $div\left(|\nabla u|^{p-2}\nabla u\right)+x\nabla\left(u^{q}\right)+\alpha u^{q}=0$ in $\mathbb{R}^{N}$. *Differential Equations* **34** (9) (1998) 1172–1180.

[13] Y. Naito and T. Suzuki. Radial symmetry self similar solutions for semilinear heat equations. *J. Diff. Equ.* **163** (2) (2000) 407–428.

[14] W. M. Ni and J. Serrin. Nonexistence theorems for singular solutions of quasilinear partial differential equations. *Pure and Applied Mathematics* **39** (1986) 379–399.

[15] L. A. Peletier, D. Terman and F. B. Weissler. On the equation $\triangle u+\frac{1}{2}x.\nabla u+f(u)=0$. *Arch. Rational Mech. Anal.* **94** (1986) 83–99.

[16] Y. W. Qi. The global existence and non uniqueness of a nonlinear degenerate equation. *Nonlinear Anal.* **31** (1-2) (1998) 117–136.

[17] J. Serrin and H. Zou. Existence and non existence results for ground states of quasilinear elliptic equations. *Arch. Rational Mech. Anal.* **121** (1992) 101–130.

[18] E. Yanagida. Uniqueness of rapidly decaying solutions to the Haraux Weissler equation. *J. Diff. Equ.* **127** (1996) 561–570.

Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 56–67

Analysis of an SIRS Epidemic Model for a Disease Geographic Spread

Mahmoud H. DarAssi [1*] and Mohammad A. Safi [2]

[1] *Department of Basic Sciences, Princess Sumaya University for Technology, Amman, Jordan.*
[2] *Department of Mathematics, The Hashemite University, Zarqa, Jordan.*

Received: August 24, 2020; Revised: December 30, 2020

Abstract: An SIRS epidemic model for the geographic spread is considered. The linear stability analysis is conducted to obtain the threshold condition and a supercritical instability region is found whenever the reproduction number $\mathcal{R} > 1$. An evolution equation for the leading order of infectives is derived by the long wavelength expansion method and full pattern formation analysis is carried out. The Poincaré-Lindstedt method is applied to obtain a uniformly periodic valid solution. Numerical simulations are used to present the results.

Keywords: *evolution equation; SIRS; stability; pattern formation; Poincaré-Lindstedt method.*

Mathematics Subject Classification (2010): 92B05, 35B35, 35B36, 47J35.

1 Introduction

In epidemiology the use of mathematical models starts from the pioneering works of Kermack and McKendrick [1–4]. To describe the Great Plague of London of 1665-1666, Kermack and McKendrick use a simple basic deterministic differential equation model called the SIR model [3], [4]. Many mathematical models in the literature are built based on the modeling framework of Kermack and McKendrick.

Most of existing studies rely on different types of differential equations. For instance, first-order partial differential equations are used for modeling of age structures [5–8]; delay-differential equations or integral equations are suitable when time delay or delay factors appear [9–13]; second-order partial differential equations are more realistic when a diffusion term exists.

* Corresponding author: mailto:m.assi@psut.edu.jo

Recently, the study of geographic spread of epidemics becomes of much interest. Diffusion epidemic models have been studied by many authors [14–19]. Liu and Jin [18] considered an SI model with either constant or nonlinear incidence function. The authors studied numerically the pattern formation of the model. In [19], Hadji applied the long wavelength expansion method to analyze a simple model for the geographic spread of a rabies epidemic in a population of foxes. The author studied the pattern formation of the model. In [14], the authors studied the effect of different types of animal movement on threshold conditions for disease spread by considering a simple SI diffusion model. Jawaz et. al. [10] considered a time delay HIV/AIDS reaction diffusion SIR model. The authors designed a numerical scheme to solve the model. Moreover, the proposed technique was compared with the results obtained by Euler's technique, also the results are presented by numerical simulations.

The main objective of this study is to develop and analyze a mathematical model of an epidemic incorporating with diffusion of the various epidemic sub-population within a geographical region. We conducted the linear and weakly nonlinear stability analysis of an SIRS with diffusion model. The long wavelength expansion is applied to obtain the evolution equation. Furthermore, a periodic uniformly valid solution is obtained by the Poincaré-Lindstedt method.

This paper is organized as follows, The mathematical model formulation is presented in Section 2. In Section 3, the linear stability analysis is conducted to obtain the threshold conditions. The weakly nonlinear stability is investigated in Section 4. In Section 5, a full pattern formation analysis is carried out. A uniformly periodic valid solution is obtained in Section 6. The results are concluded in Section 7.

2 Mathematical Model

We consider a population which consists of three subgroups. The susceptible, S, which can get the disease. The infected, I, those who have the disease and can transmit it. The removed, R, those who recovered, are immune, isolated or dead. The population is considered to have a constant size, N, where $N = S + I + R$. All the classes, S, I and R depend on space and time. The SIR reaction diffusion epidemic model is presented by J. D. Murray [20]:

$$\begin{aligned}
\frac{\partial \widehat{S}}{\partial \widehat{t}} &= \mathcal{D}\,\nabla^2 \widehat{S} - \beta\,\widehat{S}\widehat{I} + \gamma \widehat{R},\\
\frac{\partial \widehat{I}}{\partial \widehat{t}} &= \mathcal{D}\,\nabla^2 \widehat{I} + \beta\,\widehat{S}\widehat{I} - r\widehat{I},\\
\frac{\partial \widehat{R}}{\partial \widehat{t}} &= \mathcal{D}\,\nabla^2 \widehat{R} + r\widehat{I} - \gamma \widehat{R},
\end{aligned} \tag{1}$$

where $\mathcal{D}$ is the diffusion coefficient, β is the disease transmission coefficient, r is the recovery rate and γ is the loss of natural immunity. Upon using the following scaling $I = \widehat{I}/S_0, S = \widehat{S}/S_0, \widehat{R}/S_0, x = \widehat{x}/H$ and $t = \beta\, S_0\, \widehat{t}$, where S_0 is a reference value of the susceptible species, we obtain the dimensionless system which is described by

$$\begin{aligned}
\frac{\partial S}{\partial t} &= \nabla^2 S - SI + \delta\, R,\\
\frac{\partial I}{\partial t} &= \nabla^2 I + SI - \lambda\, I,
\end{aligned}$$

$$\frac{\partial R}{\partial t} = \nabla^2 R + \lambda\, I - \delta\, R, \tag{2}$$

where $\delta = \gamma/\beta\, S_0$ and $\lambda = r/\beta\, S_0$ is the reciprocal of the reproduction rate, $\mathcal{R}$. The corresponding boundary conditions are

$$S = 1, \qquad \frac{\partial\, I}{\partial\, z} = 0 \quad \text{and} \qquad R = 0$$

at $z = 0, 1$. The basic states of S, I and R are

$$S_B = 1, \qquad I_B = 0 \quad \text{and} \quad R_B = 0 \tag{3}$$

for any values of λ and δ. We introduce the perturbations ϕ, θ and ψ to the base state so that $S = 1 + \phi, I = \theta + 0$ and $R = \psi + 0$. Hence the system of equations (2) becomes

$$\begin{aligned}\frac{\partial\phi}{\partial t} &= \nabla^2\phi - \theta - \phi\,\theta + \delta\,\psi,\\ \frac{\partial\theta}{\partial t} &= \nabla^2\theta + \phi\,\theta + (1-\lambda)\,\theta,\\ \frac{\partial\psi}{\partial t} &= \nabla^2\psi + \lambda\,\theta - \delta\,\psi,\end{aligned} \tag{4}$$

subject to

$$\phi = 0, \qquad \frac{\partial\,\theta}{\partial\, z} = 0 \quad \text{and} \qquad \psi = 0$$

at $z = 0, 1$.

3 Stability Threshold Condition

Following a standard procedure (see [21] and [22]), the linearized system of equations governing the convective perturbations is given by

$$\begin{aligned}\frac{\partial\phi}{\partial t} &= \nabla^2\phi - \theta + \delta\,\psi,\\ \frac{\partial\theta}{\partial t} &= \nabla^2\theta + (1-\lambda)\,\theta,\\ \frac{\partial\psi}{\partial t} &= \nabla^2\psi + \lambda\,\theta - \delta\,\psi.\end{aligned} \tag{5}$$

We investigate the linear stability by considering the normal modes $[\Phi, \Theta, \Psi] = [\Phi(z), \Theta(z), \Psi(z)]\, \exp(i\,\mathbf{K}\cdot\,\mathbf{X} + \sigma\, t)$, where $\mathbf{X} = \langle x, y\rangle$, σ is the growth rate and $|\mathbf{K}| = k$ is the wavenumber in the system of equations (5) to obtain the following system of second order ordinary differential equations:

$$\sigma\,\Phi = (D^2 - k^2)\Phi - \Theta + \delta\,\Psi, \tag{6}$$

$$\sigma\,\Theta = (D^2 - k^2)\Theta + (1-\lambda)\,\Theta, \tag{7}$$

$$\sigma\,\Psi = (D^2 - k^2)\Psi + \lambda\,\Theta - \delta\,\Psi. \tag{8}$$

where $D = d/dz$.

Multiply both sides of equation (7) by the complex conjugate of Θ and integrate with respect to z from 0 to 1 to get

$$\sigma = \frac{-\langle |D\,\Theta|^2\rangle - (k^2+\lambda-1)\langle |\Theta|^2\rangle}{\langle |\Theta|^2\rangle}, \tag{9}$$

where $\langle\cdot\rangle = \int_0^1 \cdot\, dz$. If $\sigma = 0$, the solution of equation (7) is $\Theta = A_1 \cosh\alpha\, z + A_2 \sinh\alpha\, z$, where $\alpha = \sqrt{k^2+\lambda-1}$. Apply the boundary conditions $\frac{d\Theta}{dz} = 0$ at $z = 0, 1$. We get $\alpha = 0$ and hence, $\lambda = 1 - k^2$. Thus, the critical λ value, $\lambda_c = 1$ when the wavenumber $k = 0$. Therefore, the reproduction number $\mathcal{R} = \frac{1}{\lambda} = \frac{1}{1-k^2}$ and hence $\mathcal{R}_c = 1$.

Theorem 3.1 *The model (2) has a supercritical instability region whenever the reproduction number $\mathcal{R} > 1$.*

The numerical simulation of the relation between the reproduction number $\mathcal{R}$ and the wavenumber k is depicted in Figure 1. Figure 2 shows the plot of the recovered compartment R as functions of z.

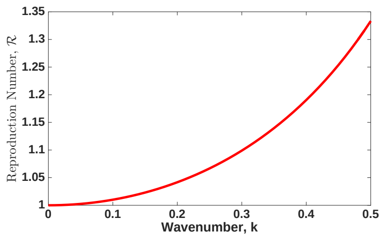


Figure 1: The plot of the reproduction number $\mathcal{R}$ as a function of the wavenumber k.

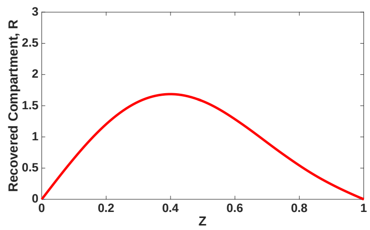


Figure 2: The plot of the recovered compartment, R, as functions of z with $r = 0.1, \beta = 0.9, \delta = 0.01$ and $S_0 = 1000$.

4 Weakly Nonlinear Stability Analysis

A nonlinear evolution equation will be derived in this section. Since the population wavenumber is zero, the long wavelength expansion can be applied to the equations (5). A small perturbation parameter ϵ, $0 < \epsilon \ll 1$, will be introduced. We scale the dimensions

$$\frac{\partial}{\partial x} = \epsilon^{1/2}\frac{\partial}{\partial X}, \quad \frac{\partial}{\partial y} = \epsilon^{1/2}\frac{\partial}{\partial Y}, \quad \frac{\partial}{\partial z} = \frac{\partial}{\partial Z}, \quad \tau = \epsilon^2 t$$

and we expand $\lambda = 1 - \epsilon^2\mu^2$

$$\phi = \epsilon\,\phi_1 + \epsilon^2\,\phi_2 + \cdots\,,$$

$$\theta = \epsilon\,\theta_1 + \epsilon^2\,\theta_2 + \cdots\,,$$

$$\psi = \epsilon\,\psi_1 + \epsilon^2\,\psi_2 + \cdots\ ,$$
$$\delta = \epsilon\,\delta_1 + \epsilon^2\,\delta_2 + \cdots\ ,$$

where μ, δ_1 and δ_2 are of $O(1)$ quantities. For simplicity, we consider the one-dimensional problem. The $O(\epsilon)$ problem is described by

$$\begin{aligned} D^2\phi_1 - \theta_1 &= 0,\\ D^2\theta_1 &= 0,\\ D^2\psi_1 + \theta_1 &= 0, \end{aligned} \tag{10}$$

where $D = \partial/\partial Z$. It is subject to the boundary conditions $\psi_1 = \phi_1 = 0$ and $\partial\theta_1/\partial Z = 0$ at $Z = 0, 1$. Its solution is given by

$$\begin{aligned} \theta_1 &= f(X,\tau),\\ \phi_1 &= \frac{f}{2}(Z^2 - Z),\\ \psi_1 &= -\frac{f}{2}(Z^2 - Z). \end{aligned}$$

Proceed to the next order, the $O(\epsilon^2)$ problem is described by

$$\begin{aligned} D^2\,\phi_2 + (\phi_1)_{XX} - \theta_2 - \theta_1\,\phi_1 + \delta_1\,\psi_1 &= 0,\\ D^2\,\theta_2 + (\theta_1)_{XX} + \theta_1\,\phi_1 &= 0,\\ D^2\,\psi_2 + (\psi_1)_{XX} + \theta_2 - \delta_1\,\psi_1 &= 0, \end{aligned}$$

subject to the boundary conditions $\psi_2 = \phi_2 = 0$ and $\partial\theta_2/\partial Z = 0$ at $Z = 0, 1$. Its solution is given by

$$\begin{aligned} \phi_2 &= -\frac{f_{XX}}{12}\left(Z^4 - Z^3\right) - \frac{f^2}{720}\left(Z^6 - 3\,Z^5 - 30\,Z^4 + 60\,Z^3 - 28\,Z\right),\\ &\quad + \frac{\delta_1\,f}{24}\left(Z^4 - 2\,Z^3 + Z\right) + \frac{B}{2}\left(Z^2 - Z\right),\\ \theta_2 &= -\frac{f_{XX}}{2}\,Z^2 - \frac{f^2}{24}\left(Z^4 - 2\,Z^3\right) + B, \end{aligned}$$

$$\psi_2 = \frac{f_{XX}}{12}\left(Z^4 - Z^3\right) + \frac{f^2}{720}\left(Z^6 - 3\,Z^5 + 2\,Z\right) - \frac{\delta_1\,f}{24}\left(Z^4 - 2\,Z^3 + Z\right) - \frac{B}{2}\left(Z^2 - Z\right),$$

where

$$B = \frac{137}{1512}\,f^2 + \frac{1}{21}\,f_{XX} + \frac{17\,\delta_1}{168}\,f.$$

Proceeding to the order $O(\epsilon^3)$, we get

$$(\theta_1)_\tau = D^2\theta_3 + (\theta_2)_{XX} - \mu^2\,\theta_1 + \theta_2\,\phi_1 + \theta_1\,\phi_2. \tag{11}$$

Application of the orthogonality conditions on equation (11) yields the sought evolution equation:

$$f_\tau = -\frac{5}{42}\,f_{XXXX} - \frac{17\,\delta_1}{168}\,f_{XX} - \mu^2\,f - \frac{199}{45360}\,f^3 - \frac{43}{5040}\,f^2 + \frac{11}{1260}f\,f_{XX} + \frac{1559}{15120}\,\left(f^2\right)_{XX}. \tag{12}$$

5 Pattern Formation

By generalizing the procedure which has been used above, the three-dimensional nonlinear evolution equation will be obtained:

$$f_\tau = -\frac{5}{42}\,\nabla_H^4 f - \frac{17\,\delta_1}{168}\,\nabla_H^2 f - \lambda_1 f - \frac{199}{45360}\,f^3 - \frac{43}{5040}\,f^2 + \frac{11}{1260} f\,\nabla_H^2 f + \frac{1559}{15120}\,\left(\nabla_H^2 f^2\right), \tag{13}$$

where $\nabla_H = (\partial/\partial X, \partial/\partial Y)$. Upon using the following transformation in equation (13):

$$\xi = \sqrt{\frac{17\delta_1}{40}}\,X, \quad \zeta = \sqrt{\frac{17\delta_1}{40}}\,Y, \quad \eta = \frac{289\,\delta_1^2}{13440\,\tau}, \quad f = \mu^2\,F,$$

equation (13) is transferred to

$$F_\eta = -\nabla_H^4 F - 2\,\nabla_H^2 F - \omega\,F - A\,F^3 - B\,F^2 + C\,F\,\nabla_H^2 + E\,\nabla_H^2 F^2, \tag{14}$$

where $\nabla_H = (\partial/\partial\,\xi, \partial/\partial\,\zeta)$, $A = \dfrac{1592\mu^2}{7803\,\delta_1^2}$, $B = \dfrac{344\mu}{867\,\delta_1}$, $C = \dfrac{44\mu}{255\,\delta_1}$, $E = \dfrac{1559\mu}{765\,\delta_1}$ and $\omega = \dfrac{13440}{289\,\delta_1^2}\mu^2$.

The linear stability analysis will be applied, where the growth rate of normal modes with small amplitude can be determined by considering the linearized version of equation (14)

$$\frac{\partial\,F}{\partial\,\eta} = -\nabla_H^4 F - 2\,\nabla_H^2 F - \omega\,F. \tag{15}$$

Application of the normal modes $f = \exp(\sigma\,\eta + i\,\mathbf{K}\cdot\mathbf{r})$, where σ is the growth rate, $\mathbf{K}$ is the wave vector such that $|\mathbf{K}| = k$ and $\mathbf{r} = (\xi, \zeta)$, in equation (15) yields

$$\sigma = -(k^2 - 1)^2 + 1 - \omega.$$

Hence, the range of instability is $0 < \omega < 1$. That is, $0 < mu^2 < \dfrac{289\,\delta_1^2}{13440}$. The effect of nonlinear terms in the evolution equation (14) can be depicted by considering the following expansions:

$$F = \epsilon\,F_1 + \epsilon^2\,F_2 + \cdots,$$

$$\omega = 1 - \epsilon\,\omega_1 - \epsilon^2\,\omega_2 - \cdots$$

and $\eta = \epsilon^2\,\widehat{\eta}$. The solution to the $O(\epsilon)$ which is described by

$$\nabla_H^4 F_1 + 2\,\nabla_H^2 F_1 + F_1 = 0 \tag{16}$$

is given by

$$F_1(\xi, \zeta, \widehat{\eta}) = U(\widehat{\eta})\,\cos(\zeta) + V(\widehat{\eta})\,\cos\left(\frac{\sqrt{3}\,\xi}{3}\right)\,\cos\left(\frac{\zeta}{2}\right). \tag{17}$$

Equation (17) yields a roll structure when $V(\widehat{\eta}) = 0$ and a square structure when $U(\widehat{\eta}) = 0$. Proceed to the next order $O(\epsilon^2)$, the problem is described by

$$\nabla_H^4 F_2 + 2\,\nabla_H^2 F_2 + F_2 = \omega_1\,F_1 - B\,F_1^2 + C\,F_1\,\nabla_H^2 F_1 + E\,\nabla_H^2 F_1^2. \tag{18}$$

Upon averaging equation (18) by multiplying both sides by F_1 and integrating over the domain of ξ and ζ, we get $\omega_1 = 0$. Hence the solution of equation (18) is given by

$$F_2 = -\left(\frac{B+C}{2} + \left(\frac{B+C+4\,E}{2}\right)\cos(\zeta)\right) U^2$$

$$-2\left(E\,\cos\left(\frac{3\,\zeta}{2}\right) + (E+B+C)\,\cos(\zeta)\,\cos\left(\frac{\zeta}{2}\right)\right) \times \cos\left(\sqrt{3}\,\xi\right)\,U\,V +$$

$$\left[\left(\frac{B+C-E}{4}\right)\cos(\zeta) - \left(\frac{C+B}{4}\right) - \left(\frac{3E+C+B}{4}\right)\right.$$

$$\left. + \left(\frac{4E-B-C}{4}\right)\cos(\zeta)\cos(\sqrt{3}\,\xi)\right] V^2.$$

Proceed to the next order, $O(\epsilon^3)$, the problem is described by

$$\frac{\partial\,F_1}{\partial\widehat{\eta}} = -\nabla_H^4\,F_3 - 2\,\nabla_H^2\,F_3 - F_3 + \omega_1\,F_2 + \omega_2\,F_1$$

$$-\,A\,F_1^3 - B\,F_1\,F_2 + C\left(F_1\,\nabla_H^2 F_2 + F_2\,\nabla_H^2 F_1\right) + E\nabla_H^2(F_1\,F_2). \tag{19}$$

In order to obtain the amplitude equations, we will average equation (19) with F_1 to get

$$\langle\frac{\partial\,F_1}{\partial\widehat{\eta}}, F_1\rangle = -\langle\nabla_H^4\,F_3 + 2\,\nabla_H^2\,F_3 + F_3, F_1\rangle$$

$$+\,\langle\omega_2\,F_1 - A\,F_1^3 - B\,F_1\,F_2, F_1\rangle + \langle C\left(F_1\,\nabla_H^2 F_2 + F_2\,\nabla_H^2 F_1\right) + E\nabla_H^2(F_1\,F_2), F_1\rangle. \tag{20}$$

The following are the amplitude equations of U and V:

$$\frac{\partial\,U}{\partial\widehat{\eta}} = -\Gamma_1\,U^3 + \omega_2\,U + \Gamma_2\,U\,V^2, \tag{21}$$

$$\frac{\partial\,V}{\partial\widehat{\eta}} = -\Gamma_3\,V^3 + \omega_2\,V + \Gamma_2\,V\,U^2, \tag{22}$$

where

$$\Gamma_1 = \frac{398}{2601\,\delta_1} - \frac{1341428296}{169130025\,\delta_1^2}, \quad \Gamma_2 = \frac{2504438396}{169130025\,\delta_1^2} - \frac{796}{2601\,\delta_1}$$

and

$$\Gamma_3 = \frac{199}{1734\,\delta_1} - \frac{320488891}{56376675\,\delta_1^2}.$$

We set $V = 0$ in equation (21) to determine the stability of roll structure. Hence the equilibrium points are $(U, V) = (0, 0)$ and $(U, V) = \left(\pm\sqrt{\frac{\omega_2}{\Gamma_1}}, 0\right)$.

Theorem 5.1 *If $\Gamma_1 > 0$ and $\omega_2 > 0$, then the roll solutions exist and the bifurcation is sub-critical. If both Γ_1 and ω_2 are negative, then the roll solutions exist and the bifurcation is supercritical.*

Following the same analysis, the square structure can be determined. If we set $U = 0$ in equation (22), the equilibrium points are $(U, V) = (0, 0)$ and $(U, V) = \left(0, \pm\sqrt{\frac{\omega_2}{\Gamma_3}}\right)$.

Theorem 5.2 *The square solutions exist with sub-critical bifurcation when both Γ_3 and ω_2 are positive. But if both are negative, the square solutions exist with supercritical bifurcation.*

To determine the hexagons structure, we set $V = 2\,U$ in equation (17) and assume that $\eta = \epsilon\,\widehat{\eta}$. Hence, the $O(\epsilon^2)$ problem is described by

$$\frac{\partial F_1}{\partial \widehat{\eta}} = -\nabla_H^4 F_2 - 2\,\nabla_H^2 F_2 - F_2 + \omega_1 F_1 - B\,F_1^2 + C\,F_1\,\nabla_H^2 F_1 + E\nabla_H^2 F_1^2. \tag{23}$$

To obtain the amplitude equation for hexagons, we average equation (23) with F_1 so that

$$\langle \frac{\partial F_1}{\partial \widehat{\eta}}, F_1\rangle = \langle -\nabla_H^4 F_2 - 2\,\nabla_H^2 F_2 - F_2 + \omega_1 F_1 - B\,F_1^2 + C\,F_1\,\nabla_H^2 F_1 + E\nabla_H^2 F_1^2, F_1\rangle. \tag{24}$$

Thus, the amplitude equation for hexagons is given by

$$\frac{\partial U}{\partial \widehat{\eta}} = \left(\omega_1 + \frac{1559}{765\,\delta_1}\right) U + \Gamma\,U^2, \tag{25}$$

where $\Gamma = -\dfrac{2468}{4335\,\delta_1}$.

Theorem 5.3 *Since $\Gamma < 0$, the solution of equation (14) has sub-critical down hexagons formation.*

6 Uniformly Periodic Valid Solution

Upon applying the following scales on equation (12) $f = ah$, $\xi = bX$, $\widehat{\tau} = e\tau$, $\gamma = a\lambda_1$, $e = 1/a$, we obtain

$$\frac{\partial h}{\partial \widehat{\tau}} = -h_{\xi\xi\xi\xi} - 2\mu^2\,h_{\xi\xi} - \gamma\,h - \alpha_1\,h^2 - \alpha_2\,h^3 + \alpha_3\,h\,h_{\xi\xi} + \alpha_4\,(h_\xi)^2, \tag{26}$$

where $a = \dfrac{6}{\sqrt{65}}$, $\quad b = \sqrt[4]{7\sqrt{\dfrac{13}{5}}}$, $\mu^2 = \dfrac{17\,\delta_1}{8\sqrt{455}} \times \sqrt[4]{\dfrac{13}{5}}$, $\alpha_1 = \dfrac{43}{9100}$, $\quad \alpha_2 = \dfrac{199}{13650\sqrt{65}}$, $\alpha_3 = \dfrac{5}{6\sqrt{7}} \times \sqrt[4]{\dfrac{13}{5}}$ and $\alpha_4 = \dfrac{1559}{1950\sqrt{7}} \times \sqrt[4]{\dfrac{13}{5}}$.

We investigate the stability of the static solution of equation (26), when its linear part is given by

$$\frac{\partial h}{\partial \widehat{\tau}} = -h_{\xi\xi\xi\xi} - 2\mu^2\,h_{\xi\xi} - \gamma\,h. \tag{27}$$

Upon introducing the normal modes $h(\xi,\tau) = \mathrm{e}^{\sigma\tau + i\,k\,\xi}$, we obtain the following dispersion relation:

$$\sigma = -(k^2 - \mu^2)^2 - \gamma. \tag{28}$$

The static state solution is unstable whenever $\gamma < \mu^4$. To investigate the weakly nonlinear stability of the evolution equation we introduce the small parameter, $\epsilon \ll 1$, and conduct the perturbation analysis near the linear solution. We expand

$$\gamma = \mu^4 - \epsilon\,\gamma_1 - \epsilon^2\,\gamma_2, \quad \tau = \epsilon^2\,\eta,$$

$$h = \epsilon\, h_1 + \epsilon^2\, h_2 + \epsilon^3\, h_3 + \cdots$$

The order $O(\epsilon)$ of equation (26) is described by

$$(h_1)_{\xi\xi\xi\xi} + 2\mu^2\,(h_1)_{\xi\xi} + \gamma\, h_1 = 0. \tag{29}$$

The solution of equation (29) is $h_1 = \cos(\mu\xi)$. Because of the secular terms, we will apply the Poincaré - Lindstedt method [23] to obtain a uniformly valid periodic solution. Substitute $\nu = w\,\xi$ and expand $w = 1 + \epsilon\, w_1 + \epsilon^2\, w_2 + \cdots$ in equation (26) to obtain

$$w^4\, h_{\nu\nu\nu\nu} + 2\,\mu^2\, w^2\, h_{\nu\nu} + \gamma\, h = \alpha_1\, h^2 + \alpha_2\, h^3 + w^2\, \left(\alpha_3\, h\, h_{\nu\nu} + \alpha_4\, (h_\nu)^2\right). \tag{30}$$

Define: $\mathscr{L}(h) = h_{\nu\nu\nu\nu} + 2\mu^2\, h_{\nu\nu} + \gamma\, h$. The order $O(\epsilon)$ problem is described by

$$\mathscr{L}(h_1) = 0. \tag{31}$$

The solution of equation (31) is $h_1 = \cos(\mu\nu)$. Proceed to order $O(\epsilon^2)$, the problem is described by

$$\mathscr{L}(h_2) = \gamma_1\, \cos(\mu\nu) + \Gamma_1 + \Gamma_2\, \cos(2\mu\nu). \tag{32}$$

To remove the secular terms, we set $\gamma_1 = 0$, which means that there is no subcritical instability. The solution of the resulting equation is given by $h_2 = \dfrac{\Gamma_1}{\mu^4} + \dfrac{\Gamma_2}{9}\, \cos(2\,\mu\nu)$, where

$$\Gamma_1 = \frac{1}{2}\, \left(\alpha_1 - \mu^2\, \alpha_3 + \mu^2\, \alpha_4\right) \quad \text{and} \quad \Gamma_1 = \frac{1}{2}\, \left(\alpha_1 - \mu^2\, \alpha_3 - \mu^2\, \alpha_4\right).$$

Proceed to the next order $O(\epsilon^3)$, the problem is described by

$$\begin{aligned}\mathscr{L}(h_3) = \Bigg[&-4\mu^4\, w_1^2 + \gamma_2 - \frac{\Gamma_1}{\mu^4}(2\alpha_1 + \mu^2\, \alpha_3) \\ &+ \frac{\Gamma_2}{36}(-4\alpha_1 - 27\alpha_2 - 10\mu^2\alpha_3 + 16\mu^2\alpha_4)\Bigg]\, \cos(\mu\nu) - \left[\frac{w_1}{18}(96\mu^4\Gamma_2 + 9\mu^2\alpha_3)\right]\, \cos(2\mu\nu) \\ &+ \left[\frac{\Gamma_2}{36}(-4\alpha_1 - 9\alpha_2 - 10\mu^2\alpha_3 + 16\mu^2\alpha_4)\right]\, cos(3\mu\nu) - \frac{w_1\mu^2\, \alpha_3}{2}.\end{aligned} \tag{33}$$

Removing the secular term by setting

$$-4\mu^4\, w_1^2 + \gamma_2 - \frac{\Gamma_1}{\mu^4}(2\alpha_1 + \mu^2\, \alpha_3) + \frac{\Gamma_2}{36}(-4\alpha_1 - 27\alpha_2 - 10\mu^2\alpha_3 + 16\mu^2\alpha_4) = 0$$

yields

$$w_1 = \pm\sqrt{\frac{\gamma_2}{4\mu^4} - \frac{\Gamma_1}{4\mu^8}\,(2\alpha_1 + \mu^2\alpha_3) + \frac{\Gamma_2}{144\mu^4}\,(-4\alpha_1 - 27\alpha_2 - 10\mu^2\alpha_3 + 16\mu^2\alpha_4)}.$$

Upon solving the rest of equation (33), we get

$$h_3 = \left[\frac{w_1}{3\mu^2}(32\mu^2\Gamma_2 + 3\alpha_3)\right]\, \cos(2\mu\nu) + \left[\frac{\Gamma_2}{2304\mu^4}(-4\alpha_1 - 9\alpha_2 - 10\mu^2\alpha_3\right.$$

$$+16\mu^2\alpha_4)\big]\ \cos(3\mu\nu) + \frac{w_1\,\alpha_3}{2\mu^2}. \tag{34}$$

Thus, a uniformly valid steady state solution to equation (26) is given by

$$\begin{aligned} h &= \cos(\mu(1+\epsilon w_1)\xi)\,\epsilon + \left(\frac{\Gamma_1}{\mu^4} + \frac{\Gamma_2}{9}\cos(2\,\mu(1+\epsilon w_1)\xi)\right)\epsilon^2 \\ &+ \left(\left[\frac{w_1}{3\mu^2}(32\mu^2\Gamma_2 + 3\alpha_3)\right]\cos(2\mu(1+\epsilon w_1)\xi)\right) \\ &+ \left[\frac{\Gamma_2}{2304\mu^4}(-4\alpha_1 - 9\alpha_2 - 10\mu^2\alpha_3 + 16\mu^2\alpha_4)\right]\cos(3\mu(1+\epsilon w_1)\xi)) + \frac{w_1\,\alpha_3}{2\mu^2}\Bigg)\,\epsilon^3. \end{aligned} \tag{35}$$

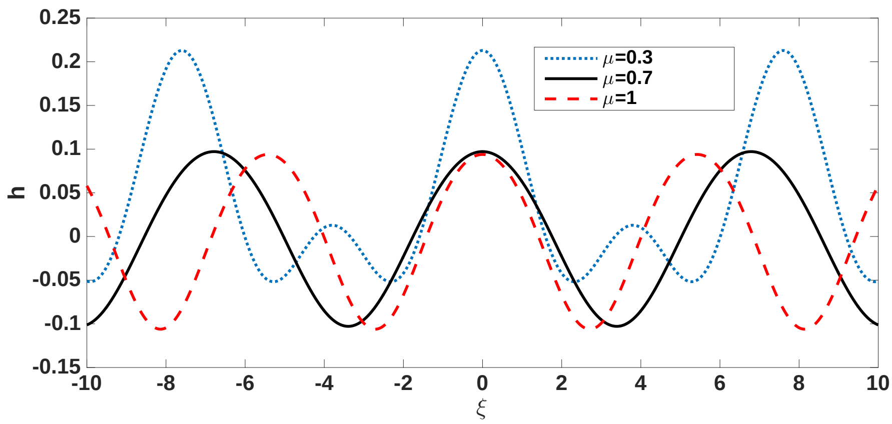


Figure 3: The plot of h as a function of ξ with $\mu = 0.3$ (dotted line), $\mu = 0.7$ (solid line), $\mu = 1$ (dashed line) and $\gamma_2 = 10$.

7 Conclusion

An SIRS model for the spread of a disease has been considered. This model involves the spatial diffusion so that the effect of landscape can be captured. The domain of the model is not infinite, it is complemented with two horizontal boundaries. The stability threshold condition is obtained so that whenever the reproduction number $\mathcal{R} > 1$, a supercritical instability region is depicted. See Theorem 3.1.

Because of the infinite wavelength, the weakly nonlinear stability analysis was conducted by applying the long wavelength asymptotic analysis method and the proposed model is reduced to a single evolution equation (12). Full pattern formation analysis of equation (12) is carried out and the subcritical down hexagons are depicted.

It is found that there exists a stable uniform solution, namely $F = 1$. Upon retrieving the original variables, we have $f = \mu\,F = \dfrac{\sqrt{1-\lambda}}{\epsilon}$, which yields the following expression of infected, susceptible and recovered:

$$I = \frac{\sqrt{1-\lambda}}{\epsilon}, \quad S = 1 + \frac{\sqrt{1-\lambda}}{2\epsilon}(Z^2 - Z), \quad R = \frac{\sqrt{1-\lambda}}{2\epsilon}(Z - Z^2).$$

Figure (4) shows the plot of I as a function of the reproducing number $\mathcal{R}$ and the space dependent functions S and R.

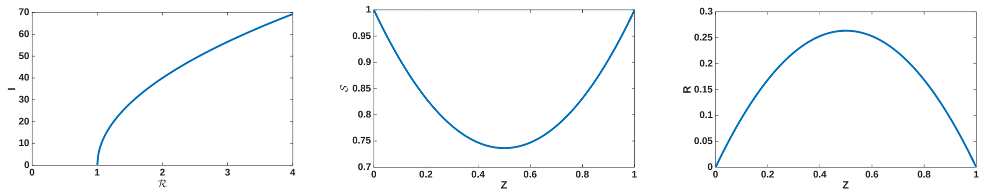


Figure 4: The plot of the infected, I, as a function of the reproduction number, $\mathcal{R}$, (left), the plot of the susceptible, S, as a function of z, (middle), and the plot of the recovered, R, as a function of z, (right).

Moreover, the Poincaré-Lindstedt method is applied to obtain a uniformly periodic valid steady state solution which is depicted in Figure 3.

References

[1] A. G. McKendrick. Applications of mathematics to medical problems. *Proc. Edinb. Math. Soc.* **13** (1926) 98–130.

[2] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* **115** (1927) 700–721.

[3] W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics, ii - the problem of endemicity. *Proc. of the Royal Society of Edinburgh. Section A. Mathematics* **138** (1932) 55–83.

[4] W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics, iii - further studies of the problem of endemicity. *Proc. of the Royal Society of Edinburgh. Section A. Mathematics* **141** (1933) 94–122.

[5] F. Sanchez, J. G. Calvo, E. Segura and Z. Feng. A partial differential equation model with age-structure and nonlinear recidivism: Conditions for a backward bifurcation and a general numerical implementation. *Computers & Mathematics with Applications* **78** (2019) 3916–3930.

[6] A. Vivas-Barber and S. Lee. Integro-Differential Age-Structured System for Influenza Transmission Dynamics. *Mathematics and Statistics* **7** (2019) 14–24.

[7] W. Guo, M. Ye, X. Li, A. Meyer-Baese and Q. Zhang. A theta-scheme approximation of basic reproduction number for an age-structured epidemic system in a finite horizon. *Mathematical Biosciences and Engineering* **16** (2019) 4107–4121.

[8] M. A. Safi and M. H. DarAssi. Mathematical analysis of an age-structured HSV-2 model. *Journal of Computational Methods in Sciences and Engineering* **19** (2019) 841–856.

[9] M. Ali, M. Rafiq and M. O. Ahmad. Numerical analysis of a modified SIR epidemic model with the effect of time delay. *Punjab Univ. J. Math.* **51** (2019) 79–90.

[10] M. Jawaz, N. Ahmed, D. Baleanu, M. Rafiq and M. Rehman. Positivity Preserving Technique for the Solution of HIV/AIDS Reaction Diffusion Model With Time Delay. *Front. Phys.* **7** (2020) 229.

[11] A. Kumar and Nilam. Mathematical analysis of a delayed epidemic model with nonlinear incidence and treatment rates. *J. Eng. Math.* **115** (2019) 1–20.

[12] A. Kumar, K. Goel and Nilam. A deterministic time-delayed SIR epidemic model: mathematical modeling and analysis. *Theory Biosci* **139** (2020) 67–76.

[13] M. H. DarAssi, M. A. Safi and B. Al-Hdaibat. A delayed SEIR epidemic model with pulse vaccination and treatment. *Nonlinear Studies* **25** (2018) 521–534.

[14] A. M. Fofana and A. Hurford. Mechanistic movement models to understand epidemic spread. *Phil. Trans. R. Soc. B* **372** (2017) 20160086.

[15] N. C. Chi, E. Av́ilaVales and G. G.Aimeida. Analysis of a HBV Model with Diffusion and Time Delay. *Journal of Applied Mathematics* **2012** (2012) 578561.

[16] R. Xu and Z. Ma. An HBV model with diffusion and time delay. *Journal of Theoretical Biology* **257** (2009) 499–509.

[17] E. F. Lambin, A. Tran, S. O. Vanwambeke, C. Linard and V. Soti. Pathogenic landscapes: Interactions between land, people, disease vectors, and their hosts. *International Journal of Health Geographics* **9** (2010) 54.

[18] Q. X Liu and Z. Jin. Formation of spatial patterns in epidemic model with constant removal rate of the infectives. *J. Stat. Mech.* (2007) P05002.

[19] L. Hadji. Long Wavelength Analysis of a Model for the Geographic Spread of a Disease. *Applications & Applied Mathematics* **9** (2014) 467–478.

[20] J. D. Murray. *Mathematical Biology II: Spatial Models and Biomedical Applications.* Springers-Verlag, New-York, 2003.

[21] P. G. Drazin and W. H. Reid. *Hydrodynamic Stability.* Cambridge University Press, United Kingdom, 2004.

[22] S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability.* Oxford: Clarendon Press, 1961.

[23] P. G. Drazin. *Nonlinear Systems.* Cambridge University Press, 1992.

Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 68–75

Boundedness and Dynamics of a Modified Discrete Chaotic System with Rational Fraction

N. Djafri [1], T. Hamaizia [2*] and F. Derouiche [3]

[1] *1EPSECSG Constantine*
[2] *Department of Mathematics, Faculty of Exact Sciences, University of Constantine 1, Algeria*
[3] *Department of Mathematics, Faculty of Exact Sciences, University of Oum El-Bouaghi, Algeria*

Received: October 12, 2020; Revised: December 25, 2020

Abstract: A modified 2-D discrete chaotic system with rational fraction is introduced in this paper, it has more complicated dynamical structures than the Hénon map and Lozi map. Some dynamical behaviors, value domain, fixed point, period-doubling bifurcation, the route to chaos, and Lyapunov exponents spectrum are further investigated using both theoretical analysis and numerical simulation. In particular, the map under consideration is a simple rational discrete bounded map capable of generating multi-fold strange attractors via period-doubling bifurcation routes to chaos. This new discrete chaotic system has extensive application in many fields such as optimization chaos and secure communication.

Keywords: *2-D rational chaotic map; new chaotic attractor; coexisting attractors.*

Mathematics Subject Classification (2010): 34D45, 70K55, 34D20.

1 Introduction

A discrete-time dynamical system is given by a map $T : X \to X$ from a space X into itself; we are interested in the asymptotic behavior of sequences $(x(n))$ defined by $x(n+1) = T(x(n))$, depending on the initial condition $x(0)$. Interest in dynamical systems sprang up in the 1960s-70s when it was shown that: (a) very simple dynamical systems can have an extremely complex "chaotic" behavior, which appears to be "random"; (b) such "chaotic" behavior can paradoxically be "stable"; (c) the behavior of some dynamical systems is so "chaotic" and "random" that it is best studied statistically. One of these models is the Lozi map [1, 2]. Moreover, it is possible to change the form of

* Corresponding author: mailto:el.tayyeb@umc.edu.dz

the Lozi map for obtaining others chaotic attractors [1, 3–6]. In [7], a one-dimensional discrete chaotic system with rational fraction was proposed. In [8], the authors extended the first one-dimensional discrete chaotic system with two-dimensional and in a recent work given in [9] the dynamics of a new simple 2-D rational discrete mapping was studied. In particular, an example of coexistence of several chaotic attractors was presented and discussed. In this paper, we propose the new discrete chaotic system with rational fraction given by

$$f(x,y) = \begin{pmatrix} y + 1 - a.(\frac{1}{0.1+x^2}) \\ b.x \end{pmatrix}. \tag{1}$$

The map (1) is obtained by changing the term $|x|$ in the nonlinear Lozi mapping by the fraction $(\frac{1}{0.1+x^2})$, the discrete iterative systems with rational fraction was discovered in the study of evolutionary algorithm, this type of applications is used in secure communications using the notions of chaos [10, 11].

2 Analytical Results

The new chaotic attractors described by map (1) have several important properties such as: (i) The map (1) is defined for all points in the plane. (ii) The associated function $f(x,y)$ of the map (1) is of class $C^\infty(\mathbb{R}^2)$, and it has no vanishing denominator. (iii) The system (1) and the Lozi system are not topologically equivalent, because the Lozi system is a piecewise linear, but the model (1) is a nonlinear system.

3 Fixed Points and Their Stability

In this section, we begin by studying the existence of fixed points of the f mapping and determine their stability type. Indeed, we have

$$\begin{cases} x = bx + 1 - a\left(\dfrac{1}{0.1+x^2}\right), \\ y = bx. \end{cases} \tag{2}$$

Hence, $x = bx + 1 - a\left(\dfrac{1}{0.1+x^2}\right)$, then $[(1-b)x - 1](0.1 + x^2) + a = 0$, so

$$(1-b)x^3 - x^2 + 0.1(1-b)x - 0.1 + a = 0. \tag{3}$$

First, eliminate the term x^2 by substituting $x = X - (-\frac{1}{3(1-b)})$ which yields the reduced cubic equation $X^3 + PX + q = 0$, where

$$P = 0.1 - \frac{1}{3(1-b)^2} \tag{4}$$

and

$$q = -\frac{2}{27(1-b)} + \frac{a-0.1}{1-b} + \frac{0.1}{3(1-b)^2}. \tag{5}$$

The reduced cubic equations with the negative discriminate $27q^2 + 4p^3$ will have 3 real roots only if $P < 0$.

Proposition 3.1 *The f mapping will have 3 fixed points only if* $b \in]1 - \sqrt{\frac{10}{3}}, 1[\cup]1, 1 + \sqrt{\frac{10}{3}}[$.

Proof. Let

$$\begin{cases} x = b.x + 1 - a.(\frac{1}{0.1+x^2}), \\ y = b.x. \end{cases}$$

Then we have $x = b.x + 1 - a(\frac{1}{0.1+x^2})$. Hence, $[(1-b)x - 1](0.1 + x^2) + a = 0$. That is, $(1-b)x^3 - x^2 + 0.1(1-b)x - 0.1 + a = 0$. In this case, we have $q^2 < -\frac{4}{27}p^3 \Leftrightarrow q^2 < -\frac{4}{27}p \times p^2$ only if $p < 0$. That is, $0.1 - \frac{1}{3(1-b)^2} < 0; (b - 1 \neq 0)$. so $\frac{1}{3(1-b)^2} > 0.1$, then $(1-b)^2 < \frac{10}{3}$. Thus $1 - \sqrt{\frac{10}{3}} < b < 1 + \sqrt{\frac{10}{3}}$.

The Jacobian matrix of the map (1) is

$$J = \begin{pmatrix} \frac{2ax}{(0.1+x^2)^2} & 1 \\ b & 0 \end{pmatrix}.$$

We have $|J| = -b$. So, if $b > -1$, then the system is dissipative. J has the following characteristic polynomial:

$$P(\lambda) = \lambda^2 - \frac{2ax}{(0.1 + x^2)^2}\lambda - b,$$

so, the eigenvalues are

$$\lambda_{1,2} = \frac{ax}{(0.1 + x^2)^2} \pm \sqrt{\frac{(ax)^2}{(0.1 + x^2)^4} + b}.$$

It is easy to check that the smallest absolute values are always less than 1. Then we deduce that the fixed points are of saddle type.

4 Determination of Bounded and Unbounded Orbits

We remark that the variations of the right-hand side of system (1) depend mainly on the fraction which is a smooth function. In what follows, we shall prove the boundedness of system (1) using a comparison criterion. It is possible to rewrite system (1) in the form (6) below:

$$x_{n+1} = 1 - (\frac{a}{0.1 + x_n^2}) + bx_{n-1}. \tag{6}$$

Now, by successive substitution of the terms of the sequence $(x_n)_n$ we can prove that this sequence is bounded for all $b < 1$ as shown by the following result.

Theorem 4.1 *For every $n > 1$, and all values of a and b, and for all values of the initial conditions $(x_0, x_1) \in \mathbb{R}^2$, the sequence $(x_n)_n$ satisfies the following conditions:*

(a) *If $b \neq 1$, then*

$$x_n = \begin{cases} \frac{b^{\frac{n-1}{2}}-1}{b-1} + b^{\frac{n-1}{2}}x_1 - a\sum\limits_{m=1}^{m=\frac{n-1}{2}} \frac{b^{m-1}}{0.1+x^2_{n-(2m-1)}} & \textit{if } n \textit{ is odd,} \\ \frac{b^{\frac{n}{2}}-1}{b-1} + b^{\frac{n}{2}}x_0 - a\sum\limits_{m=1}^{m=\frac{n}{2}} \frac{b^{m-1}}{0.1+x^2_{n-(2m-1)}} & \textit{if } n \textit{ is even;} \end{cases} \tag{7}$$

(b) *If b=1, then*

$$x_n = \begin{cases} \frac{n-1}{2} + x_1 - a \sum\limits_{m=1}^{m=\frac{n-1}{2}} \frac{1}{0.1+x^2_{n-(2m-1)}} & \textit{if } n \textit{ is odd,} \\ \frac{n}{2} + x_0 - a \sum\limits_{m=1}^{m=\frac{n}{2}} \frac{1}{0.1+x^2_{n-(2m-1)}} & \textit{if } n \textit{ is even.} \end{cases} \tag{8}$$

.

Theorem 4.2 *The sequence $(x_n)_n$ given in (1) satisfies the following inequality:*

$$\forall a, b \in \mathbb{R}, \forall n > 1, |1 - x_n + bx_{n-2}| \leq |10a|. \tag{9}$$

Proof. We have for every $n > 1$: $x_n = 1 - a.(\frac{1}{0.1+x^2_{n-1}}) + b.x_{n-2}$, then one has

$$|-x_n + 1 + bx_{n-2}| = |\frac{a}{0.1 + x^2_{n-1}}| \leq |10a|. \tag{10}$$

Since $x^2_n > 0$, one has $0.1 + x^2_n > 0.1$, so $\frac{1}{0.1+x^2_n} < 10$, then we get $|\frac{a}{0.1+x^2_n}| < |10a|$.

5 Existence of Bounded and Unbounded Orbits

In the following theorem, we give sufficient conditions for bounded and unbounded orbits of the system (1).

Theorem 5.1 *For all $a \in \mathbb{R}$ and all initial conditions $(x_0; x_1) \in \mathbb{R}^2$:*

i. *The orbits of the map (1) are bounded in the following subregions of $\mathbb{R}^4$:*

$$\Gamma_1 = \left\{(a, b, x_0, x_1) \in \mathbb{R}^4 / |b| < 1\right\}. \tag{11}$$

ii. *The map (1) possesses unbounded orbits in the following subregions of $\mathbb{R}^4$:*

$$\Gamma_2 = \left\{(a, b, x_0, x_1) \in \mathbb{R}^4 / |b| > 1, \textit{and both } |x_0|, |x_1| > \frac{|10a| + 1}{|b| - 1}\right\} \tag{12}$$

and

$$\Gamma_3 = \left\{(a, b, x_0, x_1) \in \mathbb{R}^4 / |b| = 1, \textit{and } |10a| < 1\right\}. \tag{13}$$

Proof. **I)** From equation (1) and the fact that $(\frac{1}{0.1+x^2_n})$ is a bounded function for all $x \in \mathbb{R}$, one has the following inequalities for all $n > 1$:

$$|x_n| \leq 1 + |10a| + |bx_{n-2}|. \tag{14}$$

If we replace the successive terms $x_{n-2}, x_{n-4}, x_{n-6}...$, in the term x_n, then the last term is obtained:

$$|x_n| \leq (1 + |10a|) + |b|(1 + |10a|) + |b|^2(1 + |10a|) + |b|^3|x_{n-6}|. \tag{15}$$

Since $|b| < 1$, the use of (15) and induction about some integer k using the sum of a geometric growth formula permits us to obtain the following inequalities for every $n > 1$, $k > 0$:

$$|x_n| \leq (1 + |10a|)(\frac{1 - |b|^k}{1 - |b|}) + |b|^k|x_{n-2k}|. \tag{16}$$

Thus, one has the following two cases:
1) if $n = 2m+1$, then $(x_n)_n$ satisfies the following inequalities:

$$|x_{2m+1}| \leq (1+|10a|)(\frac{1-|b|^m}{1-|b|}) + |b|^m|x_1| = w_m; \tag{17}$$

2) if $n = 2m$, then $(x_n)_n$ satisfies the following inequalities:

$$|x_{2m}| \leq (1+|10a|)(\frac{1-|b|^m}{1-|b|}) + |b|^m|x_0| = v_m. \tag{18}$$

Thus, since $|b| < 1$, the sequences $(w_m)_m$ and $(v_m)_m$ are bounded, and one has

$$\begin{cases} w_n \leq \frac{1+|10a|}{1-|b|} + ||x_1| - \frac{1+|10a|}{1-|b|}| & \text{for all } m \in \mathbb{N}, \\ v_n \leq \frac{1+|10a|}{1-|b|} + ||x_0| - \frac{1+|10a|}{1-|b|}| & \text{for all } m \in \mathbb{N}. \end{cases} \tag{19}$$

Thus, the previous formulas give the following bounds for the sequence $(x_n)_n$:

$$|x_m| \leq max(\frac{1+|10a|}{1-|b|} + ||x_1| - \frac{1+|10a|}{1-|b|}|, \frac{1+|10a|}{1-|b|} + ||x_0| - \frac{1+|10a|}{1-|b|}|). \tag{20}$$

Finally, for all values of a and all values of b satisfying $|b| < 1$ and all initial conditions $(x_0; x_1) \in \mathbb{R}^2$, one concludes that all orbits of the map (1) are bounded, i.e., in the sub-region of $\mathbb{R}^4$

$$\Gamma_1 = \left\{(a, b, x_0, x_1) \in \mathbb{R}^4 / |b| < 1\right\}.$$

Hence the proof (i) is completed.

II) (a) For every $n > 1$ we have $x_n = 1 - a.(\frac{1}{0.1+x_{n-1}^2}) + b.x_{n-2}$, then $|b.x_{n-2} - a.(\frac{1}{0.1+x_{n-1}^2})| = |x_n - 1|$ and $||b.x_{n-2}| - |a.(\frac{1}{0.1+x_{n-1}^2})|| \leq |x_n - 1|$, (we use the triangular inequality), this implies that

$$|b.x_{n-2}| - |a.(\frac{1}{0.1+x_{n-1}^2})| \leq |x_n| + 1. \tag{21}$$

Since $|(\frac{1}{0.1+x_{n-1}^2})| \leq 10$, this implies that $|a(\frac{1}{0.1+x_{n-1}^2})| \leq 10|a|$.
$|b.x_{n-2}| - |a(\frac{1}{0.1+x_{n-1}^2})| \geq |b.x_{n-2}| - 10|a|$. Finally, one has from (21) that

$$|b.x_{n-2}| - (10|a|+1) \leq |x_n|. \tag{22}$$

Then, by induction, as in the previous section, one has

$$|x_n| \geq \begin{cases} (\frac{|10a|+1}{|b|-1} + |x_1|)|b|^{\frac{n-1}{2}} + \frac{|10a|+1}{|b|-1} & \text{if } n \text{ is odd}, \\ (\frac{|10a|+1}{|b|-1} + |x_0|)|b|^{\frac{n}{2}} + \frac{|10a|+1}{|b|-1} & \text{if } n \text{ is even}. \end{cases} \tag{23}$$

Thus, if $|b| > 1$ and both $|x_0|, |x_1| > (\frac{|10a|+1}{|b|-1})$, one has $lim_{n\to+\infty}|x_n| = +\infty$.

(b) For $b = 1$, one has

$$|x_n| \geq \begin{cases} (1-|10a|)(\frac{n-1}{2}) + |x_1| & \text{if } n \text{ is odd}, \\ (1-|10a|)(\frac{n}{2}) + |x_0| & \text{if } n \text{ is even}. \end{cases} \tag{24}$$

Hence, if $|10a| < 1$, then one has $lim_{n\to+\infty}|x_n| = +\infty$.

For $b = -1$, one has from Theorem 1 the following inequalities:

$$x_n \leq \begin{cases} -(\frac{n-1}{2}) + x_1 + |a \sum\limits_{m=1}^{m=\frac{n-1}{2}} \frac{(-1)^{m-1}}{0.1+x^2_{n-(2m-1)}}| & \text{if } n \text{ is odd,} \\ -(\frac{n}{2}) + x_0 + |a \sum\limits_{m=1}^{m=\frac{n}{2}} \frac{(-1)^{m-1}}{0.1+x^2_{n-(2m-1)}}| & \text{if } n \text{ is even.} \end{cases} \tag{25}$$

Because $|(\frac{a(-1)^{m-1}}{0.1+x^2_{n-(2m-1)}})| \leq |10a|$, then one has

$$x_n \leq \begin{cases} (|10a| - 1)(\frac{n-1}{2}) + |x_1| & \text{if } n \text{ is odd} \\ (|10a| - 1)(\frac{n}{2}) + |x_0| & \text{if } n \text{ is even.} \end{cases} \tag{26}$$

Thus, if $|10a| < 1$, then one has $lim_{n\to+\infty}|x_n| = +\infty$. Note that there is no similar proof for the following subregions of $\mathbb{R}^4$ defined by

$$\Gamma_4 = \left\{(a, b, x_0, x_1) \in \mathbb{R}^4/|b| > 1, \text{and both } |x_0|, |x_1| \leq \frac{|10a| + 1}{|b| - 1}\right\} \tag{27}$$

and

$$\Gamma_5 = \left\{(a, b, x_0, x_1) \in \mathbb{R}^4/|b| = 1, \text{and } |a| \geq 1\right\}. \tag{28}$$

Hence, the proof (ii) is completed.

6 Some Observed New Attractors

As already mentioned in the introduction, we study the two-dimensional discrete chaotic system with two parameters a and b, obtained via a direct modification in the Lozi map, where the absolute value term is replaced by the rational fraction defined on all $\mathbb{R}$ and being differentiable continuous. This fact is the central idea which makes the solutions of system (1) bounded for some values of b. This new map generates chaotic attractors with multiple "multifold" that evolves around three points as shown in Fig.1.

7 Numerical Simulations and Route to Chaos

In this section, the dynamic behavior of the map (1) is studied numerically. We shall illustrate some observed chaotic attractors. The bifurcation diagram is a way for a discrete dynamical system to make a transition from regular behavior to chaos [12]. To demonstrate the chaotic dynamics, the largest Lyapunov exponent should be the first thing to be considered, because any system containing at least one positive Lyapunov exponent is defined to be chaotic. From Fig.2 and Fig.3, it is clear that the bifurcation diagram well coincides with the spectrum of Lyapunov exponents. Fig.2 shows that the system (1) can evolve into periodic and chaotic behaviors. Indeed, when a varies from 0.108 to 0.347, it can be seen that there is a positive Lyapunov exponent over a wide range of parameters, implying that the system is chaotic over this range. When a increases in the region [0.118, 0.220], the system (1) converges to a stable fixed point. In the interval section (0.108, 0.112), the trajectory of the system will turn to a stable limit cycle. In addition, the system shows a periodic motion of some windows in the chaotic region, i.e., [0.245, 0.255]. Finally, it is clear that the system is chaotic for $a \in (0.220, 0.245) \cup (0.255, 0.347)$.

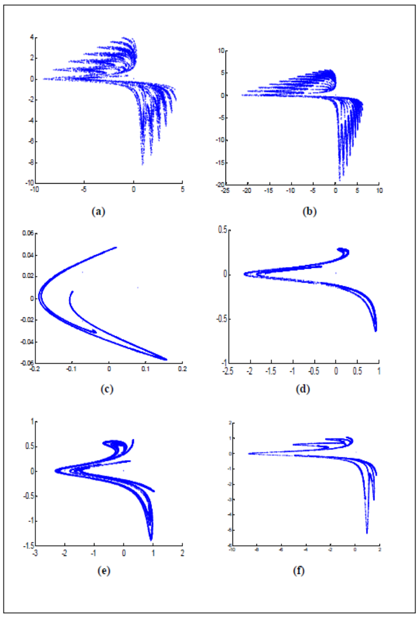


Figure 1: Attractors of the map (1) with (a) $a = 0.9$, $b = 0.9$, (b) $a = 0.9$, $b = 2$, (c) $a = 0.3$, $b = 0.115$, (d) $a = 0.3$, $b = 0.3$, (e) $a = 0.6$, $b = 0.4$, (f) $a = 0.6$, $b = 0.9$.

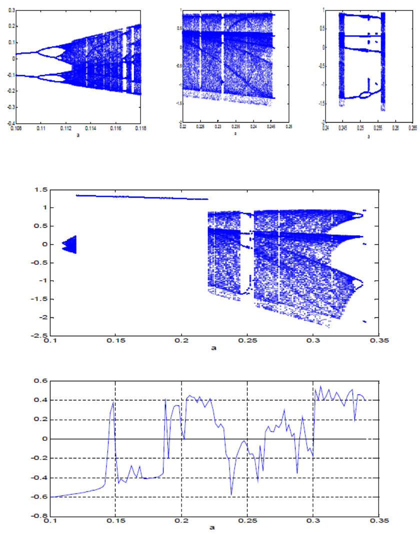
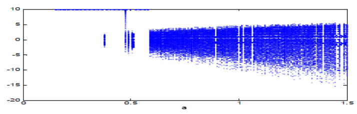


Figure 2: (a) The bifurcation diagram for the map (1) obtained for $b = 0.3$ and $0,108 < a < 0.347$. (b) Variation of the Lyapunov exponents of map (1) versus the parameter $0,108 < a < 0.347$ with $b = 0.3$.



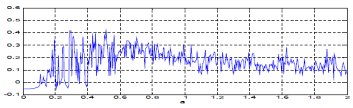


Figure 3: (a) The bifurcation diagram for the map (3) obtained for $b = 0.9$ and $0, 13 < a < 2$. (b) Variation of the Lyapunov exponents of map (3) versus the parameter $0, 13 < a < 2$ with $b = 0.9$.

8 Conclusion

In this paper we have presented a modified two-dimensional discrete chaotic system with rational fraction, obtained via direct modification of the Hénon mapping. The detailed dynamical behaviors of this map (which is useful for the evolutionary algorithm and secure communication) are further investigated using both theoretical analysis and numerical simulation.

References

[1] R. Lozi. Un attracteur etrange (?) du type attracteur de Hénon. *J. hys.* **39** (C5) (1978) 9–10.

[2] M. Misiurewicz. Strange attractors for the Lozi mappings. *Nonlinear Dyn.* **357** (1) (1980) 348–358.

[3] M. Hénon. A two dimensional mapping with a strange attractor. *Communications in Mathematical Physics* **50** (1976) 69–77.

[4] M. Benedicks and L. Carleson. The dynamics of the Hénon maps. *Annals of Mathematics* **133** (1) (1991) 73–169.

[5] Y. Cao and Z. Liu. Oriontation-preserving Lozi map. *Chaos, Solutions and Fractals* **9** (11) (1998) 1857–1863.

[6] A. Aziz-Alaoui, C. Robert and C. Grebogi. Dynamics of a Hénon-Lozi map. Chaos, Solitions and Fractals **12** (11) (2001) 2323–2341.

[7] J.A. Lu, X. Wu, J. L and L. Kang. A new discrete chaotic system with rational fraction and its dynamical behaviors. *Chaos, Solitons and Fractals* **22** (2) (2004) 311–319.

[8] L. Chang, J. Lu and X. Deng. A new two-dimensional discrete chaotic system with rational fraction and its tracking and synchronization. *Chaos, Solitons and Fractals* **24** (4) (2005) 1135–1143.

[9] E. Zeraoulia and J.C. Sprott. On the dynamics of a new simple 2-D rational discrete mapping. *Int. J. Bifurc. Chaos.* **21** (1) (2011) 1–6.

[10] G. Chen and X. Dong. *From Chaos to Order: Methodology, Perspectives and Applications.* World Scientific, 1998.

[11] B.R. Hunt, J.A. Kennedy, T.Y. Li and H. Nusse. *The Theory of Chaotic Attractorss.* Springer, New-York. 2003.

[12] V. Avrutin, M. Schanz and S. Banerjee. Occurrence of multiple attractor bifurcations in the two-dimensional piecewise linear normal form map. *Nonlinear Dyn.* **67** (2012) 293–307. DOI 10.1007/s11071-011-9978-5.

Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 76–89

Tube-MPC Based on Zonotopic Sets for Uncertain System Stabilisation

W. Hamdi[1,2*], W. Bey[1] and N. Benhadj Braiek[1,2]

[1] *Laboratory of Advanced Systems (L.A.S), Polytechnic School of Tunisia, Carthage University, BP 743,2078 La Marsa, Tunisia.*
[2] *National Higher Engineering School of Tunis, Tunis University, Taha Hussein Avenue Montfleury, 1008 Tunis.*

Received: October 23, 2020; Revised: January 27, 2021

Abstract: This paper is dedicated to the model predictive control (MPC) for constrained discrete time systems with additive uncertainties to track a reference system. The tracking MPC problem distributes a control law appropriate for regulating constrained uncertain system to a given target system. Though, when the target operating point changes, the feasibility of the controller may be lost and the controller misses to track the reference. In this paper, a novel MPC for tracking variation system references is introduced. The main issue consists in minimizing a cost that penalizes the error between the state of the original system and the reference system state. The polyhedral invariant set for tracking is considired an extended terminal constraint. The properties of the proposed controller have been tested in two examples. Simulation results show that the proposed tracking MPC successfully achieves robust tracking in terms of control performance.

Keywords: *LTV systems; robust control; tracking MPC; polytopic invariant sets.*

Mathematics Subject Classification (2010): 93B51, 93C05.

* Corresponding author: mailto:hamdi.walid987@gmail.com

1 Introduction

An effective strategy for the practical implementation of the robust MPC is the tube-based MPC. The design of a robust control law guarantees the satisfaction of hard constraints and is addressed by means of calculating a sequence of state space regions, called an accessibility tube. The term tube is based on control techniques whose purpose is to maintain all the possible trajectories of an uncertain system inside a sequence of admissible regions using set-theory related tools. Such approach has been widely employed to robustify MPC [1–5].

The tube-based robust model predictive control (TMPC) is an advanced control algorithm that can deal with model uncertainty. The basic idea of the tube-based robust MPC is to maintain a state trajectory of an uncertain system inside a sequence of tubes [6]. The TMPC is motivated by the fact that a real state trajectory differs from a state trajectory of a nominal system due to uncertainty [17]. In 2001, [8] developed a tube-based robust model predictive controller for a linear time-invariant (LTI) system subject to bounded disturbance. The control law is obtained by solving an unconstrained LQR problem. The objective is to drive the state of an uncertain system to a terminal set while using the input as little as possible. Constraint fulfillment is guaranteed by replacing the original constraints with more stringent ones. A larger control horizon implies better control performance at the price of a higher computational load, so a suitable trade off is required. In 2004, [11] proposed a tube-based robust MPC using the time-varying control inputs instead of the LTI control law. A sequence of time-varying control inputs is obtained by solving an optimal control problem subject to additional constraints sets in order to guarantee robust stability. Since the control inputs are time-varying, the proposed MPC algorithm can achieve better control performances than the conventional tube-based MPC algorithm using the LTI control law.

Tube-based MPC approaches are motivated by the fact that the predicted evolution of a system obtained using a nominal model differs from the real evolution due to uncertainty. An MPC formulation that permits to consider this mismatch in the controller synthesis is the tube-based one, whose basis consists in computing the region around the nominal prediction that contains the state of the system under any possible uncertainties [8–10].

Gonzalez et al. [12] proposed a tube-based robust MPC for tracking of a linear time-varying (LTV) system subject to bounded disturbance. The proposed MPC algorithm requires an additional assumption that the time-varying parameter at each step within the prediction horizon is known a priori. Then a reachable set at each time step is calculated instead of a disturbance invariant set in order to reduce the conservativeness. Although the conservativeness is reduced, the computational problem is more severe because both the optimal control problem and the reachable set are computed on-line.

Bumroongsri and Kheawhom [13] proposed a strategy for the design of a tube-based output feedback MPC which is independent of the estimation method employed. They formulate a control policy by choosing a candidate estimate that is consistent with the reachable sets of the system under control. The proposed method can be combined with any estimation scheme as long as the assumed error bounds are satisfied. They show that the proposed method is recursively feasible, robustly exponentially stable, and performs better than other available strategies. The idea of the Tube MPC is motivated by robustness considerations for system dynamics affected by bounded disturbances.

In this work, a new strategy for formulation of an optimal problem of the robust tube MPC based on zonotopic invariant sets was established. This method contains two steps. The first one, off-line, calculates a sequence of state feedback control laws for global systems corresponding to a sequence of zonotopic invariant sets using the LMI technique proposed by [14]. For the nominal system, a feedback control law using the LQR problem is computed. The second step, on-line, at each sampling time, determines the smallest invariant tube containing the measured state and implements the computed state feedback control. Such a control is obtained from the last two control laws. Finally, the global control law is applied to the original process allowing to improve system control performances.

The paper is organized as follows. General problem setup is presented in Section 2. Then, in Section 3, the robust model predictive control is described. Polyhedral and zonotopic sets are introduced in Section 4. Main result is presented in Section 5. The implementation of the proposed algorithm is illustrated in two numerical examples in Section 6. Finally, the paper is concluded.

2 General Problem Setup

Consider the following discrete-time LTV system with disturbance:

$$x_{k+1} = A(k)x_k + B(k)u_k + w, \tag{1}$$

where $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^n$ is the control input, $w \in \mathbb{R}^n$ is the bounded disturbance. The system is subject to the state constraint $x \in X$, the control constraint $u \in U$, and the disturbance constraint $w \in W$, where $X \subset R^n$, $U \subset R^m$ and $W \subset R^n$ are convex polytopes and each set contains the origin as an interior point.

Remark 2.1 $[A(k),\ B(k)] \in conv\,\{[A_j, B_j], \forall j \in 1,\ 2,\ ...,\ L\}$, where $[A_j, B_j]$ are vertices of the convex hull and L is the number of vertices of the convex hull. The pair $[A_j, B_j]$ is controllable.

Let the nominal system be defined by

$$x^{'}_{k+1} = Ax^{'}_k + Bu^{'}_k, \tag{2}$$

where $x^{'} \in R^n$ and $u^{'} \in R^m$ are the state and control input of the nominal system, respectively. Now, we calculate the difference between the global and the nominal system:

$$\begin{aligned} x_{k+1} - x^{'}_{k+1} &= Ax_k + Bu_k + w - (Ax^{'}_k + Bu^{'}_k) \\ &= A(x_k - x^{'}_k) + B(u_k - u^{'}_k) + w. \end{aligned} \tag{3}$$

The objective is to robustly stabilize the system (1). The presence of a persistent disturbance w means that it is not possible to regulate the state x to the origin. The best that can be hoped for, is to regulate the state to a neighborhood of the origin. Then the proposed idea is to compensate the mismatch between the real and the nominal state, and to steer the nominal system as close as possible to the reference without constraints violation. For that purpose, we consider the following control law:

$$u_k = K(x_k - x^{'}_k) + u^{'}_k, \tag{4}$$

where K is the disturbance rejection gain whose goal is to compensate the system realisation at each sampling instant.

The system (3) is rewritten as

$$x_{k+1} - x'_{k+1} = (A + BK)(x_k - x'_k) + w. \tag{5}$$

Then, using the nominal dynamics, an effective invariant tube is defined in the state space of the nominal system and a deterministic finite horizon optimization problem is formulated and solved on-line resulting in an optimal sequence of nominal controls $u'_k = \left\{u'_{k/k}, u'_{k+1/k}, ...\right\}$. Finally, the control law (4) is implemented to the process (1).

3 Robust Model Predictive Control

In this section, we present an off-line step for the MPC problem developed by [15], which consists in determining a sequence of feedback control law. In order to build an invariant tubes based on zonotopes, from the gains K_i to be found by this algorithm we establish the zonotopic invariant sets.

By solving the optimization problem presented in (6)-(10), we obtain a state feedback control law $u_k = K_i x_k$ with a state feedback gain $K_i = Y_i Q_i^{-1}$ that can stabilize the system while satisfying the input and output constraints.

The optimization problem is shown in the following LMI:

$$\min_{\gamma_i, Y_i, Q_i} \quad \gamma_i \tag{6}$$

subject to

$$\begin{bmatrix} 1 & x_i^T \\ x_i & Q_i \end{bmatrix} \geqslant 0, \tag{7}$$

$$\begin{bmatrix} Q_i & QA_j^T + Y_i^T B_j^T & Q_i \Theta^{1/2} & Y_i^T R^{1/2} \\ A_j Q_i + B_j Y_i & Q_i & 0 & 0 \\ \Theta^{1/2} Q_i & 0 & \gamma_i I & 0 \\ R^{1/2} Y_i & 0 & 0 & \gamma_i I \end{bmatrix} \geqslant 0 \ \forall j = 1, 2, \ldots, L, \tag{8}$$

$$\begin{bmatrix} X & Y_i \\ Y_i^T & Q_i \end{bmatrix} \geqslant 0, \ X_{hh} \leqslant u_{h,\max}^2, \ h = 1, 2, \ldots, n_u, \tag{9}$$

$$\begin{gathered} \begin{bmatrix} S & C((A_j Q_i + B_j Y_i) \\ (A_j Q_i + B_j Y_i)^T C^T & Q_i \end{bmatrix} \geqslant 0, \ S_{rr} \leqslant y_{r,\max}^2, \\ r = 1, 2, \ldots, n_y, \quad \forall j = 1, 2, \ldots, L, \end{gathered} \tag{10}$$

where Q is a symmetric matrix. For each K_i, calculate the sequence of zonotopic invariant sets as shown in [14].

4 Polyhedral and Zonotopic Sets

Definition 4.1 (G-representation of a zonotope) Given a vector $c \in R^n$ and a set $G = \{g_1, \ldots, g_m\}$ of vectors of R^n, $m \geqslant n$, a zonotope Z of order m is defined as follows:

$$Z = \left\{ x \in R^n, \, x = c + \sum_{i=1}^{p} \gamma_i g_i; \, -1 \leqslant \gamma_i \leqslant 1 \right\}. \tag{11}$$

The vector c is called the center of the zonotope Z. The vectors $g_1, ..., g_m$ are called generators of Z.

Definition 4.2 (V-representation of a polytope). Given r vertices $v_i \in R^n, P = conv\,\{v_1, ..., v_r\}$ is a convex polytope, where *conv* is the convex hull operator. To obtain zonotopic sets from polyhedral ones, we have to perform the following three steps:
Step 1: Compute the vertices $v_i \in R^n$ (V-representation) of all N polytopes $S_i, i = 1, ..., N$.
Step 2: Obtain the minimum and maximum values of each polytope i:

$$\begin{aligned} m_{\min} &= \min(V_i^1, ..., V_i^r)\ , \\ m_{\max} &= \max(V_i^1, ..., V_i^r), \end{aligned} \tag{12}$$

where V_i^r is the i-th component of the vector V^j and r is the number of the vertices of each polytope.
Step 3: Compute a G-representation of the n-dimensional interval $[m_{\min}, m_{\max}]$:

$$[m_{\min}, m_{\max}] = \left\{ x = c + \sum_{i=1}^{P} \gamma_i.g_i\ , -1 \leqslant \gamma_i \leqslant 1 \right\}, \tag{13}$$

where

$$c = 0.5(m_{\min} + m_{\max}), \tag{14}$$

$$g_i^{(i)} = \begin{cases} 0.5(m_{\max} - m_{\min}), & if\ i = j, \\ 0, & otherwise. \end{cases} \tag{15}$$

5 Main Result

In this part, we propose a robust tube-based MPC controller via invariant zonotopic sets. The idea of the Tube MPC is motivated by robustness considerations for system dynamics affected by bounded disturbances instead of considering each possible disturbance sequence separately in the prediction. The effect of the bounded disturbances is over-approximated by a sequence of sets which contains all possible state trajectories. In order to prevent these sets from growing too quickly within the prediction horizon, feedback is assumed in the predictions. Here we do not consider system dynamics affected by disturbances at each time instance, but instead consider (uncertain) initial offsets, which lead to a similar uncertainty in the predictions. In this way, we use robust MPC methods in order to approximate the input generated by a (nominal) MPC controller for an infinite number of initial conditions by a single robust MPC controller.

Determination of the invariant tube: Consider the zonotopic sets

$$Z = \left\{ x \in R^n,\, x = c + \sum_{i=1}^{p} \gamma_i g_i;\ -1 \leqslant \gamma_i \leqslant 1 \right\}. \tag{16}$$

The predicted state trajectory when the initial state $x' = \left(x'_{0/k}, x'_{1/k}, \ldots, x'_{N-1/k}\right)$, where $x'_{k/k} \subseteq R^n$ is the k steps ahead prediction calculated from the set of initial conditions x_k.

An invariant tube is written in the form

$$T = \left(\left\{ x'_{0/k} \right\} \oplus Z, \left\{ x'_{1/k} \right\} \oplus Z, \dots, \left\{ x'_{N-1/k} \right\} \oplus Z \right), \tag{17}$$

where $\oplus$ is the Minkowski sum. Since $x_{k+1} - x'_{k+1}$ is bounded by T, we can control the nominal system $x'_{k+1} = Ax'_k + Bu'_k$ in such a way that the LTI system with disturbance $x_{k+1} = A(k)x_k + B(k)u_k + w$ satisfies the original state and control constraints $x \in X$ and $u \in U$, respectively. To achieve this, the tighter constraint sets for the nominal system are employed, $x'_i \in X \oplus T$, $u'_i \in U \oplus KT$ for $i \in \{0, \dots, N-1\}$.

It has been demonstrated, that the control law $u_k = K(x_k - x'_k) + K_{lqr}x'_k$ keeps the states x of the original system $x_{k+1} = A(k)x_k + B(k)u_k + w$ close to the state x' of the nominal system $x'_{k+1} = Ax'_k + Bu'_k$. It is clear that if we can regulate x' to the origin, then x must be regulated to a robust positively invariant set T whose center is at the origin.

Terminal cost and terminal invariant set: A common technique to ensure the asymptotic stability of MPC is to incorporate both a terminal cost and a set of terminal constraints [16, 17]. In this part, we are interested in the two problems related to the nominal system (2). Note that the stability properties for analogous control approaches have been analyzed in the literature. The goal of the final cost is to provide closed-loop stability. For this reason, it requires the use of a Lyapunov function with a stabilization control law. In our case, a procedure similar to [1] was followed for the nominal system (2).

In order to ensure stability, an additional terminal constraint is implemented, $x'_N \in X'_f \subset X \oplus T$, where X'_f is the terminal constraint set.
Hence, $(A+BK)X'_f \subset X'_f$, $X'_f \subset X \oplus T$, $KX'_f \subset U \oplus KT$, and $V_f(A+BK)x' + l(x', u') \leqslant V_f(x')$, $\forall x' \in X'_f$,

where $u_k = K(x_k - x'_k) + K_{lqr}x'_k$ is the stage cost, and $V_f(x') = \frac{1}{2}(x')^T P x'$ is the terminal cost, Θ, R and P are the positive definite weighting matrices.

Proposition 5.1 *If $x_k \in x'_k \oplus T$, $x'_k \in X \oplus T$ and $K_{lqr}x'_k \in U \oplus T$, with K_{lqr} provided by solving an LQR problem optimisation for the nominal system (2). Then the control law $u_k = K(x_k - x'_k) + K_{lqr}x'_k$ of the global system (1) ensures satisfaction of the original constraints $x \in X$, $u \in U$ for $\forall w \in W$ and $[A(k),\ B(k)] \in conv\,\{[A_j, B_j], \forall j \in 1, 2, \dots, L\}$.*

Proposition 5.1 states that the control law $u_k = K(x_k - x'_k) + u'_k$, where $u'_k = K_{lqr}x'_k$, ensures satisfaction of the original state and control constraints.

Theorem 5.1 *For the LTV system as shown in (1), given the control law $u_k = K(x_k - x'_k) + K_{lqr}x'_k$ with a state feedback gain $K = YQ^{-1}$ provided by solving the optimization problem presented in (6)-(10) and a state feedback gain K_{lqr} provided by solving an LQR optimisation problem for the nominal system (2), the invariant tubes as shown in (17) provide a set of states whereby the system will evolve to the origin without input and output constraints violation.*

Proof. The feedback gain $K_i = Y_i Q_i^{-1}$ used in the construction of the zonotopic invariant set Z, $Z = \left\{ x \in R^n,\ x = c + \sum_{i=1}^{p} \gamma_i g_i;\ -1 \leqslant \gamma_i \leqslant 1 \right\}$, is obtained by solving convex optimization problem with LMI constraints as shown in (6)-(10). The satisfaction of (8) for a state feedback gain K ensures that $\left([A_j + BK]\, x_k\right)^T \gamma Q^{-1} \left([A_j + BK]\, x_k\right) - x_k^T \gamma Q^{-1} x_k \leq \left[x_k^T \Theta x_k + u_k^T R u_k\right]$, $j = 1, ..., l$. Also, $V_k = x_k^T \gamma Q^{-1} x_k$ is a strictly decreasing Lyapunov function (negative derivative) and the closed-loop system is robustly stabilized by the state feedback gain K.

Hence, a set of initial states $T = \left(\left\{x'_{0/k}\right\} \oplus Z, \left\{x'_{1/k}\right\} \oplus Z, \ldots, \left\{x'_{N-1/k}\right\} \oplus Z\right)$ is constructed such that all predicted states remain inside $T(x_k \subset T)$, and approach to the origin without constraint violation. Moreover, the invariant tube T constructed is never an empty set ($T \notin \{\}$) because the given feedback gain K_{lqr} is a stabilizable gain.

Corollary 5.1 *The state of the LTV system with disturbance $x_{k+1} = Ax_k + Bu_k + w$ at each time step is restricted to lie within a tube whose center is the state of the nominal LTV system $x'_{k+1} = Ax'_k + Bu'_k$.*

So, in summary, with Theorem 1 and Proposition 1, the off-line tube robust MPC algorithm based on zonotopes for the LTV system with disturbance $x_{k+1} = Ax_k + Bu_k + w$ can be formulated as follows.

Off-line:

– Solve (6)-(10) using Yalmip toolbox MATLAB.
– Calculate the feedback gain $K = YQ^{-1}$.
– Construction of the corresponding zonotopic invariant sets:
$Z = \left\{x \in R^n,\, x = c + \sum_{i=1}^{p} \gamma_i g_i;\, -1 \leqslant \gamma_i \leqslant 1\right\}$.
– Calculate the feedback gain K_{lqr}, for the nominal system (2) using the LQR problem.

On-line: At each sampling time, calculate x'_{k+1} from $x'_{k+1} = (A_j + BK_{lqr})\, x'_k$, $j = 1, 2, ..., L$. Then obtain the invariant tubes $T = \left(\left\{x'_{0/k}\right\} \oplus Z, \left\{x'_{1/k}\right\} \oplus Z, \ldots, \left\{x'_{N-1/k}\right\} \oplus Z\right)$, and we determine the smallest tube invariant set containing the measured state and implement the corresponding state feedback control law $u_k = K_i(x_k - x'_k) + K_{lqr} x'_k$, $i = 1, 2, ..., N$, to the process.

6 Numerical Examples

6.1 Example 1

Let us consider an uncertain non-isothermal CSTR [12], where the exothermic reaction $A \rightarrow B$ takes place. The reaction is irreversible and the rate of reaction is first order with respect to component A. A cooling coil is used to remove heat that is released in the exothermic reaction. The uncertain parameters are: the reaction rate constant k_0 and the heat of reaction δH_{rxn}. The linearized model based on the component balance and the energy balance is given by the following state equations:

$$\left\{ \begin{array}{c} \dot{x}\,(t) = Ax\,(t) + Bu\,(t) + w, \\ y\,(t) = Cx\,(t), \end{array} \right. \tag{18}$$

where $\begin{bmatrix} C_A \\ T \end{bmatrix}$ is the state vector $x(t)$ and $\begin{pmatrix} C_{A,F} \\ F_C \end{pmatrix}$ is the input control vector $u(t)$. Matrices are defined by

$$A = \begin{bmatrix} 0.85 - 0.0986\alpha(k) & -0.0014\alpha(k) \\ 0.9864\alpha(k)\beta(k) & 0.0487 + 0.01403\alpha(k)\beta(k) \end{bmatrix}, w = [0 \quad 0]^T, \quad (19)$$
$$B = \begin{bmatrix} 0.15 & 0 \\ 0 & -0.912 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

where C_A is the concentration of A in the reactor, $C_{A,F}$ is the feed concentration of A, T is the reactor temperature, and F_C is the coolant flow. The operating parameters are: $F = 1m^3/min, V = 1m^3$, $k_0 = 10^9 - 10^{10}min^{-1}$, $E/R = 8330.1K$, $-?H_{rxn} = 10^7 - 10^8 cal/kmol$, $\rho = 10^6 g/m^3$, $UA = 5.3410^6 cal/(Kmin?)$ and $C_p = 1cal/(gK)$.

Let $\overline{C}_A = C_A - C_{A,eq}$, $\overline{C}_{A,F} = C_{A,F} - C_{A,F,eq}$ and $\overline{F}_C = F_C - F_{C,eq}$, where the subscript eq is used to denote the corresponding variable at the equilibrium condition. By discretization, using a sampling time of $15min$, and the discrete-time model with $\begin{bmatrix} \overline{C}_A(k) \\ \overline{T}(k) \end{bmatrix}$ as a state vector $\begin{bmatrix} \overline{C}_{A,F} \\ \overline{F}_C(k) \end{bmatrix}$ as a control vector, is given as follows:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + w, \\ y(k) = Cx(k) \end{cases} \quad (20)$$

$$\begin{cases} x(k+1) = \begin{bmatrix} \overline{C}_A(k+1) \\ \overline{T}(k+1) \end{bmatrix} \\ \quad = \begin{bmatrix} 0.85 - 0.0986\alpha(k) & -0.0014\alpha(k) \\ 0.9864\alpha(k)\beta(k) & 0.0487 + 0.01403\alpha(k)\beta(k) \end{bmatrix} \begin{bmatrix} \overline{C}_A(k) \\ \overline{T}(k) \end{bmatrix} \\ \quad + \begin{bmatrix} 0.15 & 0 \\ 0 & -0.912 \end{bmatrix} \begin{bmatrix} \overline{C}_{A,F} \\ \overline{F}_C(k) \end{bmatrix}, \\ y(k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \overline{C}_A(k) \\ \overline{T}(k) \end{bmatrix}, \end{cases} \quad (21)$$

where $1 \le \alpha(k) = k_0/10^9 \le 10$ and $1 \le \beta(k) = -\Delta H_{rxn}/10^7 \le 10$.

The two parameters $\alpha(k)$ and $\beta(k)$ are independent of each other. Then we consider the following polytopic uncertain model with four vertices:

$$\Omega = conv \left\{ \begin{matrix} \begin{bmatrix} 0.751 & -0.0014 \\ 0.986 & 0.063 \end{bmatrix}, \begin{bmatrix} 0.751 & -0.0014 \\ 9.864 & 0.189 \end{bmatrix} \\ \begin{bmatrix} 0.751 & -0.0014 \\ 0.986 & 0.063 \end{bmatrix}, \begin{bmatrix} 0.751 & -0.0014 \\ 9.864 & 0.189 \end{bmatrix} \end{matrix} \right\}. \quad (22)$$

These matrices are used to calculate four off-line feedback gains K_{lqr} for the nominal system $x'_{k+1} = Ax'_k + Bu'_k$.

The objective is to regulate the concentration $\overline{C}_A$ and the reactor temperature $\overline{T}$ to the origin by manipulating $\overline{C}_{A,F}$ and $\overline{F}_C$, respectively. These variables are constrained by $|\overline{C}_{A,F}| \le 0.5\, kmol/m^3$ and $|\overline{F}_C| \le 1.5\, m^3/min$.

The weighting matrices in the cost function are given as $\Theta = I$ and $R = 0.1I$.

Let us choose a sequence of states

$$x_i = \left\{ \begin{matrix} (0.0525, 0.0525), (0.0475, 0.0475), \\ (0.0425, 0.0425), (0.0375, 0.0375), \\ (0.0325, 0.0325), (0.0275, 0.0275). \end{matrix} \right\}. \quad (23)$$

The nominal feedback control laws $u' = K_{lqr}x'$ calculated with the LQR problem, are given by

$K_1 = \begin{bmatrix} -0.34 & 0 \\ 0.50 & 0.03 \end{bmatrix}$, $K_2 = \begin{bmatrix} -3.41 & 0 \\ 5.08 & 0.09 \end{bmatrix}$, $K_3 = \begin{bmatrix} 0.12 & 0.04 \\ 4.91 & 0.09 \end{bmatrix}$, $K_4 = \begin{bmatrix} -4.28 & 0.01 \\ 49.49 & 0.72 \end{bmatrix}$.

Six feedback gains calculated off-line using the LMI methods, are given by
$K_1 = \begin{bmatrix} -1.51 & -0.01 \\ 24.66 & 0.13 \end{bmatrix}$, $K_2 = \begin{bmatrix} -1.54 & -0.01 \\ 27.39 & 0.15 \end{bmatrix}$, $K_3 = \begin{bmatrix} -1.56 & -0.01 \\ 30.64 & 0.18 \end{bmatrix}$, $K_4 = \begin{bmatrix} -1.60 & -0.01 \\ 38.83 & 0.21 \end{bmatrix}$, $K_5 = \begin{bmatrix} -1.66 & -0.01 \\ 38.83 & 0.22 \end{bmatrix}$, $K_6 = \begin{bmatrix} -0.93 & -0.00 \\ 47.32 & 0.34 \end{bmatrix}$.

Figure 1 shows six invariant polytopes computed off-line corresponding to six feedback gains previously defined.

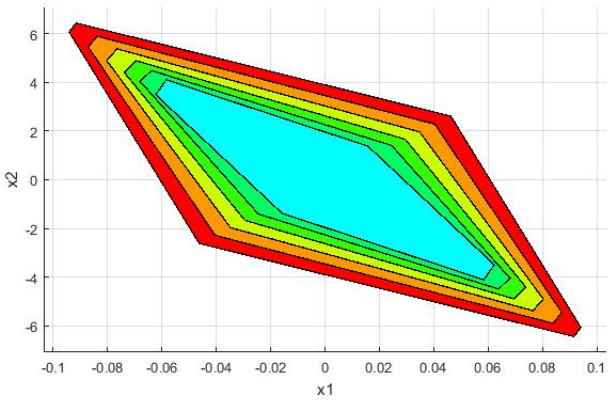


Figure 1: Polyhedral invariant sets constructed off-line.

Six zonotopes Z are defined by their centers:

$$c = \{2.9842,\ 3.1745,\ -1.3132,\ 1.3132,\ -3.1745,\ -2.9842\}. \tag{24}$$

The generators matrices are defined by
$G_1 = \begin{bmatrix} 3.1047 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $G_2 = \begin{bmatrix} 0 \\ 3.2739 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $G_3 = \begin{bmatrix} 0 \\ 0 \\ 1.2955 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $G_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1.2955 \\ 0 \\ 0 \end{bmatrix}$, $G_5 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 3.2739 \\ 0 \end{bmatrix}$, $G_6 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3.1047 \end{bmatrix}$.

In Figure 2, the terminal cost $V_f(x)$ and the invariant tubes are obtained with the optimal finite horizon value $N = 10$ function for the unconstrained problem and X_f

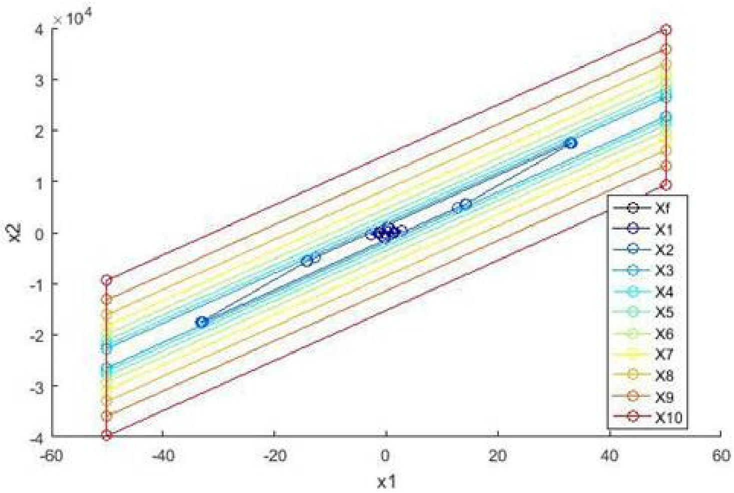


Figure 2: The terminal cost $V_f(x)$, 10 invariant tubes with $N = 10$, and the maximal set X_f.

is the associated maximal output admissible set. The regulated outputs are shown, respectively, in Figure 3 and Figure 4, it is seen that the proposed algorithm is able to steer faster the state of the uncertain CSTR to the neighborhood of the origin, and we have concluded that the use of invariant tubes based on zonotopes gives less conservative results as compared with invariant tubes based on polytope.

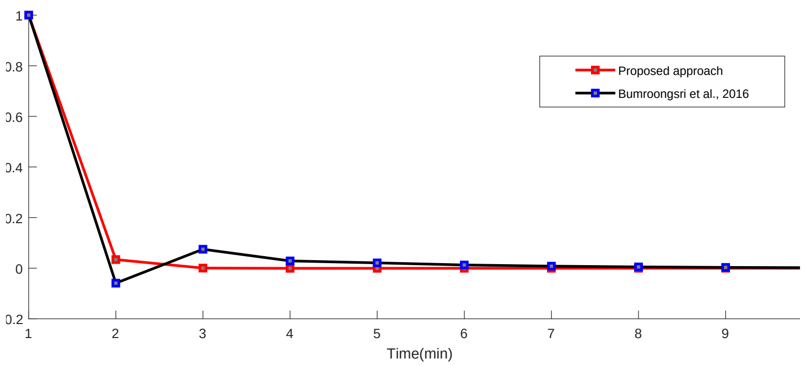


Figure 3: The concentration of A in the reactor of the regulated output.

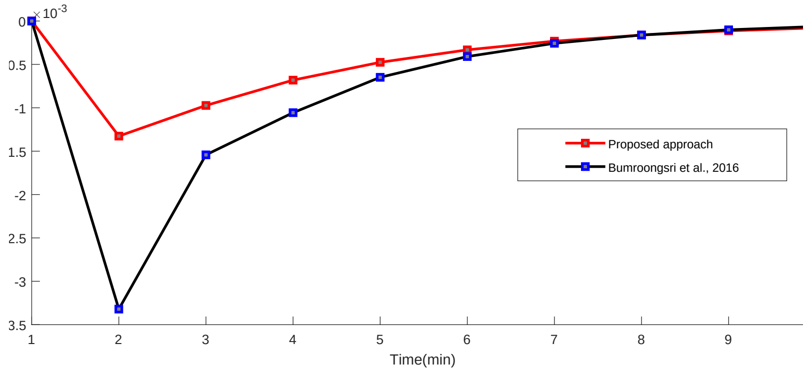


Figure 4: The reactor temperature of the regulated output.

6.2 Example 2

Consider the following LTV system with bounded disturbance:

$$x(k+1) = \begin{pmatrix} 1 & 1 \\ 0 & alpha \end{pmatrix} + \begin{pmatrix} 0.5 \\ 1 \end{pmatrix} u + w, \tag{25}$$

where $0.9 \leq \alpha \leq 1.1$. The state $x \in X$, where $X = \left\{x \in R^2 | \left[0 \quad 1\right] x \leq 2\right\}$, the control $u \in U$, where $U = \{u \in R| \, |u| \leq 1\}$, and the disturbance $w \in W$, where $W = \left\{w \in R^2 | [-0.1 \quad 0.1]^T \leq w\right\}$.

The weighting matrices in the cost function are given as $\Theta = I$ and $R = 0.01$. The following nominal LTV system:

$$x^{'}(k+1) = \left[\begin{array}{cc} 1 & 1 \\ 0 & \alpha \end{array}\right] x^{'} + \left[\begin{array}{c} 0.5 \\ 1 \end{array}\right] u^{'} \tag{26}$$

is subject to a tighter state and control constraints, $x^{'} \in X \oplus T$ and $u^{'} \in U \oplus K_{lqr} T$. By solving the LMIs problem (6)-(10) we obtain seven feedback gains corresponding to seven invariant polytopic sets shown in Figure 5.

$K_1 = \left[\begin{array}{cc} -0.45 & -1.02 \end{array}\right]$, $K_2 = \left[\begin{array}{cc} -0.47 & -1.11 \end{array}\right]$, $K_3 = \left[\begin{array}{cc} -0.51 & -1.24 \end{array}\right]$, $K_4 = \left[\begin{array}{cc} -0.47 & -1.14 \end{array}\right]$, $K_5 = \left[\begin{array}{cc} -0.53 & -1.26 \end{array}\right]$,$K_6 = \left[\begin{array}{cc} -0.53 & -1.26 \end{array}\right]$, $K_7 = \left[\begin{array}{cc} -0.53 & -1.26 \end{array}\right]$.

Using these feedback gains we obtained seven zonotopes Z defined by their centers $c = \{-1.93, \, -0.49, \, 0.69, \, 1.93, \, 0, \, 0, \, 0\}$. The generators matrices are defined by

$$G_1 = \left[\begin{array}{c} 3.38 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right], \, G_2 = \left[\begin{array}{c} 0 \\ 2.79 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right], \, G_3 = \left[\begin{array}{c} 0 \\ 0 \\ 3.04 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right], \, G_4 = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 3.28 \\ 0 \\ 0 \\ 0 \end{array}\right], \, G_5 = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ -4.63 \\ 0 \\ 0 \end{array}\right],$$

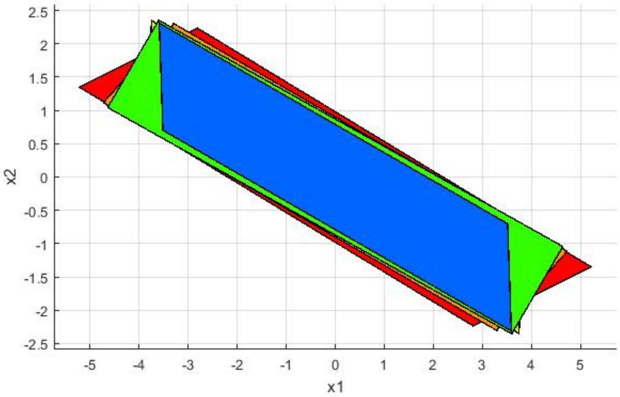


Figure 5: Polyhedral invariant sets.

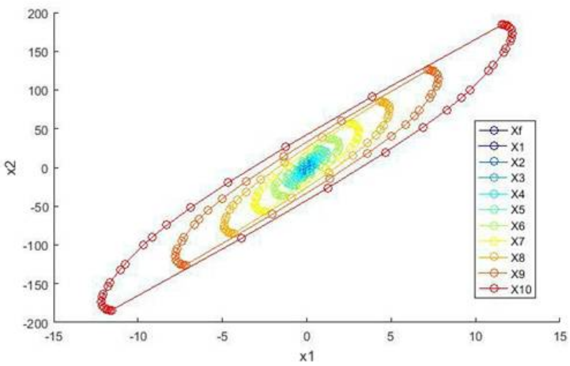


Figure 6: The terminal cost $V_f(x)$, ten invariant tubes with $N = 10$, and the maximal set X_f.

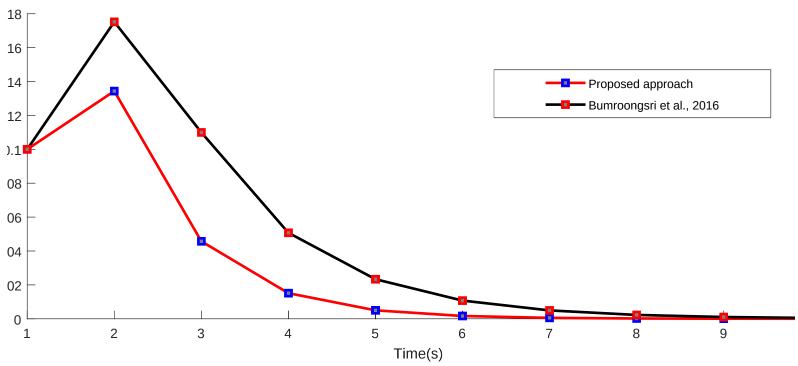


Figure 7: The regulated output using invariant tubes.

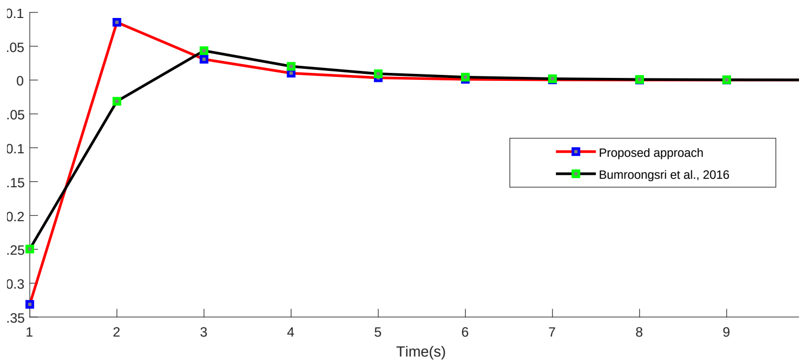


Figure 8: The control input.

$$G_6 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -2.34 \\ 0 \end{bmatrix}, \; G_7 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -4.63 \end{bmatrix}.$$

By solving an LQR problem to the nominal system (30) we obtain two feedback gains: $K_1 = [-0.65 - 1.25]$ and $K_2 = [-0.67 - 1.39]$. And using an optimal finite horizon $N = 10$ we obtain ten invariant tubes (Figure 6). Figure 7 and Figure 8 represent the closed-loop response of the system and the control input, respectively. The chosen horizon is $N = 10$ and the initial state $[0.10.2]^T$. It is seen that the proposed Tube MPC algorithm achieves better control performances.

7 Conclusion

In this paper, we have presented a new approach of uncertain discrete time system stabilization based on the robust tube MPC algorithm using zonotopic invariant sets. The proposed algorithm used an off-line solution of an optimal control optimization problem to determine a sequence of feedback gains for the global system. Then a sequence of feedback gains for the nominal system is computed using an LQR problem. Finally, a sequence of nested invariant tubes is constructed. At each sampling time, we determine the smallest zonotopic invariant set containing the measured state and implement the obtained global state feedback control law.

The proposed approach applied on two examples, provides better control performances and less computational cost.

References

[1] D. Limon, I. Alvarado, T. Alamo and E. F. Camacho. Robust tube-based MPC for tracking of constrained linear systems with additive disturbances. *Journal of Process Control* **20**(3) (2010) 248–260.

[2] P. Trodden, and A. Richards. Distributed model predictive control of linear systems with persistent disturbances. *International Journal of Control* **83**(8) (2010) 1653–1663.

[3] K. Jouili and N. B. Braiek. Design a Lyapunov control under input constraint for stabilisation of a CSTR. *International Journal of Automation and Control* **10**(1) (2016) 52–72.

[4] W. Bey, Z. Kardous and N. B. Braiek. Stabilisation of constrained uncertain systems by multi-parametric optimisation. *International Journal of Automation and Control* **10**(4) (2016) 407–416.

[5] W. Hamdi, W. Bey and N. B. Braiek. Constrained Uncertain System Stabilization with Enlargement of Invariant Sets. *Complexity*, Hindawi (11) (2020) 1468109.

[6] J. B. Rawlings and D. Q. Mayne. *Model Predictive Control: Theory and Design*. Nob Hill Pub. (2009).

[7] D. Q. Mayne and W. Langson. Robustifying model predictive control of constrained linear systems. *Electronics Letters* **37**(23) (2001) 1422–1423.

[8] L. Chisci, J. A. Rossiter and G. Zappa. Systems with persistent disturbances: predictive control with restricted constraints. *Automatica* **37**(7) (2001) 1019–1028.

[9] G. A. Leonov and M. M. Shumafov. Stabilization of Controllable Linear Systems. *Nonlinear Dynamics and Systems Theory* **9**(1) (2009) 37–50.

[10] R. Mtar, M. M. Belhaouane, H. B. Ayadi and N. B. Braiek. An LMI criterion for the global stability analysis of nonlinear polynomial systems. *Nonlinear Dynamics and Systems Theory* **9**(2) (2009) 171–183.

[11] W. Langson, I. Chryssochoos, S. V. Rakovic and D. Q. Mayne. Robust model predictive control using tubes. *Automatica* **40**(1) (2004) 125–133.

[12] R. Gonzalez, M. Fiacchini, T. Alamo, J. L. Guzman and F. Rodriguez. Online robust tube-based MPC for time-varying systems: A practical approach. *International Journal of Control* **84**(6) (2011) 1157–1170.

[13] P. Bumroongsri, and S. Kheawhom. An off-line formulation of tube-based robust MPC using polyhedral invariant sets. *Chemical Engineering Communications* **203**(6) (2016) 736–745.

[14] W. Hamdi, W. Bey and N. B. Braiek. Stabilization of constrained uncertain systems by an off-line approach using zonotopes. *ASTESJ journal* **3**(1) (2018) 281–287.

[15] P. Bumroongsri and S. Kheawhom. An off-line robust MPC algorithm for uncertain polytopic discrete-time systems using polyhedral invariant sets. *Journal of Process Control* **22**(6) (2012) 975–983.

[16] J. Hanema, M. Lazar and R. Toth. Stabilizing tube-based model predictive control: Terminal set and cost construction for LPV systems. *Automatica* **85** (2017) 137–144.

[17] D. Q. Mayne, E. C. Kerrigan and P. Falugi. Robust model predictive control: advantages and disadvantages of tube-based methods. *IFAC Proceedings* **44**(1) (2011) 191–196.

Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 90–99

Weak Solutions for Anisotropic Nonlinear Discrete Dirichlet Boundary Value Problems in a Two-Dimensional Hilbert Space

I. Ibrango [1*], B. Koné [2], A. Guiro [3] and S. Ouaro [4]

[1] *Laboratoire de Mathématiques et Informatique, UFR, Sciences et Techniques, Université Nazi Boni, 01 BP 1091 Bobo-Dioulasso 01, Burkina Faso,*
[2] *Laboratoire de Mathématiques et Informatique, UFR, Sciences Exactes et Appliquées, Université Joseph Ki-Zerbo, 03 BP 7021 Ouagadougou 03, Burkina Faso.*
[3] *Laboratoire de Mathématiques et Informatique, UFR, Sciences et Techniques, Université Nazi Boni, 01 BP 1091 Bobo-Dioulasso 01, Burkina Faso.*
[4] *Laboratoire de Mathématiques et Informatique, UFR, Sciences Exactes et Appliquées, Université Joseph Ki-Zerbo, 03 BP 7021 Ouagadougou 03, Burkina Faso.*

Received: November 23, 2020; Revised: January 29, 2021

Abstract: Using a minimization method we study the existence of weak solutions for a family of nonlinear discrete Dirichlet boundary value problems where the solution lies in a discrete $(T_1 \times T_2)$-Hilbert space. The originality of this work is the study done on a two-dimensional Hilbert space.

Keywords: *discrete boundary value problem; critical point; weak solution; two dimensional discrete Hilbert space; electrorheological fluids.*

Mathematics Subject Classification (2010): 93A10; 35B38; 35P30; 34L05.

1 Introduction

In the last few years, great attention has been paid to the study of fourth-order nonlinear difference equations. These equations have been widely used to study discrete models in many fields such as computer science, economics, neural network, ecology, cybernetics, etc. For background and recent results, we refer the reader to [3]– [12], [13] and the references therein.

* Corresponding author: mailto:ibrango2006@yahoo.fr

The main purpose of the present paper is to extend the study of difference equations in two dimensions. These models are of independent interest since their mathematical structure has a different nature. In the literature, to our knowledge, no scientific study has concerned these types of problems which are nevertheless discrete variants of the anisotropic or isotropic partial differential equations and are usually studied in connection with numerical analysis.

We study a p-Laplacian difference equation on the subset of integers. So, for $i, j \in \mathbb{N}$ with $i \leq j$, we define $\mathbb{N}[i, j]$ as the discrete interval $\{i, i+1, \dots, j\}$ and we investigate the existence of solutions for the following nonlinear discrete Dirichlet boundary value problem:

$$\begin{cases} -\Delta\big(a(k-1, h-1, \Delta u(k-1, h-1))\big) = f(k, h), \ (k, h) \in \mathbb{N}[1, T_1] \times \mathbb{N}[1, T_2], \\ u(k, h) = 0, \quad \forall (h, k) \in \Gamma, \end{cases} \tag{1}$$

where

$$\Gamma = (\{0, T_1 + 1\} \times \mathbb{N}[0, T_2 + 1]) \cup (\mathbb{N}[0, T_1 + 1] \times \{0, T_2 + 1\})$$

is the boundary of the domain $\mathbb{N}[0, T_1+1] \times \mathbb{N}[0, T_2+1]$; $\Delta u(k, h) = u(k+1, h+1) - u(k, h)$ is the forward difference operator and

$$a : \mathbb{N}[1, T_1] \times \mathbb{N}[1, T_2] \times \mathbb{R} \longrightarrow \mathbb{R}, \quad f : \mathbb{N}[1, T_1] \times \mathbb{N}[1, T_2] \longrightarrow \mathbb{R}$$

are functions to be defined later.

Our goal is to use a minimization method in order to establish some existence results of solutions of (1). The idea of the proof is to transfer the problem of the existence of solutions for (1) into the problem of existence of a minimizer for some associated energy functional. This method was successfully used by Bonanno et al. [2] for the study of an eigenvalue nonhomogeneous Neumann problem, where, under an appropriate oscillating behavior of the nonlinear term, they proved the existence of a determined open interval of positive parameters for which the problem under consideration admits infinitely many weak solutions that strongly converge to zero, in an appropriate Orlicz-Sobolev space.

The remaining part of this paper is organized as follows. Section 2 is devoted to mathematical preliminaries. The main existence result is stated and proved in Section 3. In the last section of this paper we study an extension of the problem (1).

2 Mathematical Preliminaries

We define the $(T_1 \times T_2)$-dimensional Hilbert space

$$H = \{u : \mathbb{N}[0, T_1 + 1] \times \mathbb{N}[0, T_2 + 1] \longrightarrow \mathbb{R} \quad \text{such that} \quad u(k, h) = 0, \quad \forall\ (h, k) \in \Gamma\}$$

with the inner product

$$\langle u, v \rangle = \sum_{k=1}^{T_1} \sum_{h=1}^{T_2} u(k, h) v(k, h)$$

and the associated norm defined by

$$\|u\| = \left(\sum_{k=1}^{T_1} \sum_{h=1}^{T_2} |u(k, h)|^2 \right)^{1/2}.$$

However, we introduced another norm on the space H, namely

$$|u|_m = \left(\sum_{k=1}^{T_1} \sum_{h=1}^{T_2} |u(k,h)|^m \right)^{1/m}, \quad \forall\, m \geq 2.$$

Due to equivalence of $\|.\|$ and $|.|_m$ there exist constants $C_2 \geq C_1 > 0$ such that

$$C_1\|u\| \leq |u|_m \leq C_2\|u\|, \quad \forall u \in H. \tag{2}$$

For the data f and a we impose the following conditions:

$$f \in H, \quad a(k,h,.) : \mathbb{R} \longrightarrow \mathbb{R} \text{ is continuous } \forall (k,h) \in \mathbb{N}[1,T_1] \times \mathbb{N}[1,T_2] \tag{3}$$

and there exists a mapping $A : \mathbb{N}[1,T_1] \times \mathbb{N}[1,T_2] \times \mathbb{R} \longrightarrow \mathbb{R}$ which satisfies

$$a(k,h,\xi) = \frac{\partial}{\partial \xi} A(k,h,\xi) \qquad \text{and} \qquad A(k,h,0) = 0 \quad \forall (k,h) \in \mathbb{N}[1,T_1] \times \mathbb{N}[1,T_2]. \tag{4}$$

We also assume that there exists a positive constant C_3 such that

$$|a(k,h,\xi)| \leq C_3 \left(1 + |\xi|^{p(k,h)-1}\right). \tag{5}$$

The following relations hold true for all $(k,h) \in \mathbb{N}[1,T_1] \times \mathbb{N}[1,T_2]$:

$$\left(a(k,h,\xi) - a(k,h,\eta)\right)(\xi - \eta) > 0, \ \forall\, \xi, \eta \in \mathbb{R} \text{ with } \xi \neq \eta \tag{6}$$

and

$$|\xi|^{p(k,h)} \leq a(k,h,\xi)\xi \leq p(k,h)A(k,h,\xi), \ \forall\, \xi \in \mathbb{R}. \tag{7}$$

Example 2.1 We can give the following function:

$$A(k,h,\xi) = \frac{1}{p(k,h)}\left(\left(1 + |\xi|^2\right)^{p(k,h)/2} - 1\right),$$

where

$$a(k,h,\xi) = \left(1 + |\xi|^2\right)^{(p(k,h)-2)/2}\xi, \quad \forall\ (k,h) \in \mathbb{N}[1,T_1] \times \mathbb{N}[1,T_2], \ \xi \in \mathbb{R}$$

and conditions on the function a are checked.

In this paper, we assume that the function

$$p : \mathbb{N}[1,T_1] \times \mathbb{N}[1,T_2] \longrightarrow (1, +\infty). \tag{8}$$

We will use the following notations:

$$p^- = \min_{k \in \mathbb{N}[1,T_1]} \left(\min_{h \in \mathbb{N}[1,T_2]} p(k,h) \right) \quad \text{and} \quad p^+ = \max_{k \in \mathbb{N}[1,T_1]} \left(\max_{h \in \mathbb{N}[1,T_2]} p(k,h) \right). \tag{9}$$

The discrete Wirtinger type inequalities can be generalized in two dimensions as follows.

Lemma 2.1 *For any function $u \in H$, the following inequality holds:*

$$\begin{aligned} 4\sin^2\left(\frac{\pi}{2(T_1+1)}\right)\sum_{h=1}^{T_2}\sum_{k=1}^{T_1}|u(k,h)|^2 &\leq \sum_{h=1}^{T_2}\sum_{k=1}^{T_1}|\Delta u(k-1,h-1)|^2 \\ &\leq 4\cos^2\left(\frac{\pi}{2(T_1+1)}\right)\sum_{h=1}^{T_2}\sum_{k=1}^{T_1}|u(k,h)|^2. \end{aligned}$$

Proof. Let $u \in H$. For h fixed in $\mathbb{N}[0, T_2+1]$, since $u(0,h) = 0 = u(T_1+1,h)$, the discrete Wirtinger type inequalities hold (see Theorem 12.6.1, page 860 in [1]). So, just apply the sum for $h = 0, \ldots, T_1+1$ and make a variable change. $\square$

We need the following auxiliary result throughout our paper.

Lemma 2.2 *For any function $u \in H$ with $\|u\| > 1$, there exist constants $C_4, C_5 > 0$ such that*

$$\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1}|\Delta u(k-1,h-1)|^{p(k-1,h-1)} \geq C_4\|u\|^{p^-} - C_5. \tag{10}$$

Proof. Fix $u \in H$ with $\|u\| > 1$.
Let

$$v : \mathbb{N}[0, T_1+1] \longrightarrow \mathbb{R},\ k \mapsto v(k) = u(k,h)$$

and

$$q : \mathbb{N}[0, T_1] \longrightarrow (1, +\infty),\ k \mapsto q(k) = p(k,h) \text{ with } h \text{ fixed in } \mathbb{N}[0, T_2+1].$$

According to Lemma 1 in [9] we have

$$\sum_{k=1}^{T_1+1}|\Delta v(k-1)|^{q(k-1)} \geq T_1^{(2-q^-)/2}\|v\|^{q^-} - T_1.$$

Then, there exist two constants $C_4, C_5 > 0$ such that

$$\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1}|\Delta u(k-1,h-1)|^{p(k-1,h-1)} \geq C_4\|u\|^{p^-} - C_5. \ \square$$

3 Main Results

In this section we study the existence of weak solution that we state in the following theorem.

Theorem 3.1 *Assume that (3)-(8) are satisfied. Then there is at least one weak solution for problem (1).*

By a weak solution for problem (1) we understand a function $u \in H$ such that

$$\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1}a(k-1,h-1,\Delta u(k-1,h-1))\Delta v(k-1,h-1) = \sum_{k=1}^{T_1}\sum_{h=1}^{T_2}f\big(k,h\big)v(k,h) \tag{11}$$

for any $v \in H$. The energy functional $J : H \longrightarrow \mathbb{R}$ corresponding to problem (1) is defined by the formula

$$J(u) = \sum_{k=1}^{T_1+1} \sum_{h=1}^{T_2+1} A(k-1, h-1, \Delta u(k-1, h-1)) - \sum_{k=1}^{T_1} \sum_{h=1}^{T_2} f\big(k, h\big)u(k, h). \tag{12}$$

This energy functional is vastly different from the energy functions defined before this work. Thus we indicate its properties. It is easy to see that the functional J is continuous, Gateaux differentiable and its Gateaux derivative J' at u reads

$$\langle J'(u), v\rangle = \sum_{k=1}^{T_1+1} \sum_{h=1}^{T_2+1} a(k-1, h-1, \Delta u(k-1, h-1))\Delta v(k-1, h-1) - \sum_{k=1}^{T_1} \sum_{h=1}^{T_2} f\big(k, h\big)v(k, h) \tag{13}$$

for all $v \in H$. If $u \in H$ is a critical point to J, namely $\langle J'(u), v\rangle = 0$ for all $v \in H$, we observe that

$$\sum_{k=1}^{T_1+1} \sum_{h=1}^{T_2+1} a(k-1, h-1, \Delta u(k-1, h-1))\Delta v(k-1, h-1) - \sum_{k=1}^{T_1} \sum_{h=1}^{T_2} f\big(k, h\big)v(k, h) = 0.$$

Since v is any in H, we see that the critical point u to J satisfies the problem (1).

The following results prove Theorem 3.1.

Lemma 3.1 *The functional J is coercive and bounded from below.*

Proof. We will only prove that the energy functional is coercive since the boundedness from below of J is a consequence of coerciveness.

$$\begin{aligned}
J(u) &= \sum_{k=1}^{T_1+1} \sum_{h=1}^{T_2+1} A(k-1, h-1, \Delta u(k-1, h-1)) - \sum_{k=1}^{T_1} \sum_{h=1}^{T_2} f\big(k, h\big)u(k, h) \\
&\geq \sum_{k=1}^{T_1+1} \sum_{h=1}^{T_2+1} \frac{1}{p(k-1, h-1)} |\Delta u(k-1, h-1)|^{p(k-1,h-1)} - \sum_{k=1}^{T_1} \sum_{h=1}^{T_2} |f(k, h)||u(k, h)| \\
&\geq \frac{1}{p^+} \sum_{k=1}^{T_1+1} \sum_{h=1}^{T_2+1} |\Delta u(k-1, h-1)|^{p(k-1,h-1)} \\
&- \left[\sum_{k=1}^{T_1} \sum_{h=1}^{T_2} |f(k, h)|^2\right]^{\frac{1}{2}} \left[\sum_{k=1}^{T_1} \sum_{h=1}^{T_2} |u(k, h)|^2\right]^{\frac{1}{2}} \\
&\geq \frac{C_4}{p^+}\|u\|^{p^-} - C_5 - C_6\|u\|.
\end{aligned} \tag{14}$$

Hence, since $p^- > 1$, the functional J is coercive. □

Lemma 3.2 *The functional J is weakly lower semi-continuous.*

Proof. For any $u \in H$, let

$$I(u) = \sum_{k=1}^{T_1+1} \sum_{h=1}^{T_2+1} A(k-1, h-1, \Delta u(k-1, h-1)).$$

According to the convexity of A, the function I is convex. Thus it is enough to show that I is lower semi-continuous.

Let us fix $u \in H$ and $\varepsilon > 0$. Since I is convex, we have $I(v) - I(u) \geq \langle I'(u), v-u\rangle$ for any $v \in H$. Therefore

$$\begin{aligned}
I(v) &\geq I(u) + \sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} a\Big(k-1,h-1,\Delta u(k-1,h-1)\Big) \times \\
&\qquad \Big(\Delta v(k-1,h-1) - \Delta u(k-1,h-1)\Big) \\
&\geq I(u) - \sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} \Big|a(k-1,h-1,\Delta u(k-1,h-1))\Big| \times \\
&\qquad \Big|\Delta v(k-1,h-1) - \Delta u(k-1,h-1)\Big| \\
&\geq I(u) - \sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} \Big|a(k-1,h-1,\Delta u(k-1,h-1))\Big| \times \\
&\qquad \Big|(v(k,h) - v(k-1,h-1)) - (u(k,h) - u(k-1,h-1))\Big| \\
&\geq I(u) - (\Lambda(u) + (\Phi(u)),
\end{aligned}$$

where

$$\Lambda(u) = \sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} |a(k-1,h-1,\Delta u(k-1,h-1))||v(k,h) - u(k,h)|$$

and

$$\Phi(u) = \sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} |a(k-1,h-1,\Delta u(k-1,h-1))||v(k-1,h-1) - u(k-1,h-1)|.$$

We use the Schwartz inequality to get

$$\begin{aligned}
\Lambda(u) &\leq \left[\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} |a(k-1,h-1,\Delta u(k-1,h-1))|^2\right]^{\frac{1}{2}} \\
&\quad \times \left[\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} |v(k,h) - u(k,h)|^2\right]^{\frac{1}{2}} \\
&\leq \left[\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} |a(k-1,h-1,\Delta u(k-1,h-1))|^2\right]^{\frac{1}{2}} \|v-u\|
\end{aligned}$$

and

$$\Phi(u) \leq \left[\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} |a(k-1,h-1,\Delta u(k-1,h-1))|^2\right]^{\frac{1}{2}} \|v-u\|.$$

Consequently, we have

$$\begin{aligned} I(v) &\geq I(u) - \left[1 + 2\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} |a(k-1,h-1,\Delta u(k-1,h-1))|^2\right]^{\frac{1}{2}} \|v-u\| \\ &\geq I(u) - \varepsilon \end{aligned}$$

for all $v \in H$ with $\|v-u\| < \sigma = \dfrac{\varepsilon}{K(T_1,T_2,u)}$, where

$$K(T_1,T_2,u) = \left[1 + 2\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} |a(k-1,h-1,\Delta u(k-1,h-1))|^2\right]^{\frac{1}{2}}.$$

We conclude that the functional I is lower semi-continuous. This implies that the functional J is also semi-continuous. □

Proof of Theorem 3.1. Since J is proper, weakly lower semi-continuous and coercive on H, using the relation between critical points of J and problem (1), we deduce that J has a minimizer which is a weak solution of (1). □

4 Uniqueness of Solution

In this section we examine the uniqueness of the weak solution for the problem (1). To do this, let us consider u, $v \in H$ being two solutions to the problem. By choosing $u-v \in H$ as a test function, according to the notion of weak solution, we obtain

$$\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} a(k-1,h-1,\Delta u(k-1,h-1))\Delta\left(u-v\right)(k-1,h-1) = \sum_{k=1}^{T_1}\sum_{h=1}^{T_2} f\big(k,h\big)\left(u-v\right)(k,h)$$

and

$$\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} a(k-1,h-1,\Delta v(k-1,h-1))\Delta\left(u-v\right)(k-1,h-1) = \sum_{k=1}^{T_1}\sum_{h=1}^{T_2} f\big(k,h\big)\left(u-v\right)(k,h).$$

By subtracting the two equalities above, we have

$$\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} \left(a(k-1,h-1,\Delta u(k-1,h-1)) - a(k-1,h-1,\Delta v(k-1,h-1))\right) \times \Delta\left(u-v\right)(k-1,h-1) = 0.$$

Therefore, according to the assumption (6), necessarily

$$\Delta u(k-1,h-1) = \Delta v(k-1,h-1), \quad \text{for all} \quad (k,h) \in \mathbb{N}[1,T_1+1] \times \mathbb{N}[1,T_2+1],$$

so, using Lemma 2.1

$$\begin{aligned} \|u-v\|^2 &= \sum_{k=1}^{T_1}\sum_{h=1}^{T_2} |u(k,h)-v(k,h)|^2 \\ &\leq \left(4\sin^2\left(\frac{\pi}{2(T_1+1)}\right)\right)^{-1} \sum_{h=1}^{T_2}\sum_{k=1}^{T_1} |\Delta u(k-1,h-1) - \Delta v(k-1,h-1)|^2 \\ &\leq 0, \end{aligned}$$

which means that

$$u = v.$$

5 An Extension

In this section we are going to show that the existence result obtained for problem (1) can be extended to the problem

$$\begin{cases} -\Delta\big(a(k-1,h-1,\Delta u(k-1,h-1))\big) = f(k,h,u(k,h)), & (k,h) \in \mathbb{N}[1,T_1] \times \mathbb{N}[1,T_2], \\ \\ u(k,h) = 0, \quad (h,k) \in \Gamma. \end{cases} \tag{15}$$

We shall replace the hypothesis on the source term f by the following. For each couple $(k,h) \in \mathbb{N}[0,T_1] \times \mathbb{N}[0,T_2]$, the function $f(k,h,.) : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous and there exists a constant $C_7 > 0$ and $r : \mathbb{N}[0,T_1] \times \mathbb{N}[0,T_2] \longrightarrow [2,+\infty)$ such that

$$|f\,(k,h,u(k,h))\,| \le C_7 \left(1 + |u(k,h)|^{r(k,h)-1}\right), \tag{16}$$

where $2 \le r(k,h) < p^-$ for all $(k,h) \in \mathbb{N}[0,T_1] \times \mathbb{N}[0,T_2]$.

In what follows, we denote by

$$r^- = \min_{\{(k,h)\in\mathbb{N}[0,T_1]\times\mathbb{N}[0,T_2]\}} r(k,h) \qquad \text{and} \qquad r^+ = \max_{\{(k,h)\in\mathbb{N}[0,T_1]\times\mathbb{N}[0,T_2]\}} r(k,h).$$

We denote

$$F(k,h,\xi) = \int_0^\xi f(k,h,s)ds \quad \text{for} \quad (k,h,\xi) \in \mathbb{N}[0,T_1] \times \mathbb{N}[0,T_2] \times \mathbb{R}$$

and we deduce that there exists a constant $C_8 > 0$ such that

$$|F(k,h,u)| \le C_8 \left(1 + |u(k,h)|^{r(k,h)}\right). \tag{17}$$

By a weak solution, we mean a function $u \in H$ such that

$$\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1} a(k-1,h-1,\Delta u(k-1,h-1))\Delta v(k-1,h-1) = \sum_{k=1}^{T_1}\sum_{h=1}^{T_2} f\big(k,h,u(k,h)\big)v(k,h) \tag{18}$$

for any $v \in H$.

Let

$$L(u) = \sum_{k=1}^{T_1}\sum_{h=1}^{T_2} F\big(k,h,u(k,h)\big).$$

Then, for any $u, v \in H$,

$$\langle L'(u), v\rangle = \sum_{k=1}^{T_1}\sum_{h=1}^{T_2} f\big(k,h,u(k,h)\big)v(k,h).$$

It is easy to see that L' is completely continuous and thus, the functional L is weakly lower semi-continuous. Therefore, the energy functional J associated with problem (15), defined by

$$J(u)=\sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1}A(k-1,h-1,\Delta u(k-1,h-1))-\sum_{k=1}^{T_1}\sum_{h=1}^{T_2}F\big(k,h,u(k,h)\big) \tag{19}$$

is such that $J\in C^1(H,\mathbb{R})$ and is weakly lower semi-continuous with

$$\begin{aligned}\langle J'(u),v\rangle &= \sum_{k=1}^{T_1+1}\sum_{h=1}^{T_2+1}a(k-1,h-1,\Delta u(k-1,h-1))\Delta v(k-1,h-1)\\ &\quad -\sum_{k=1}^{T_1}\sum_{h=1}^{T_2}f\big(k,h,u(k,h)\big)v(k,h)\end{aligned}$$

for all $v\in H$. This implies that the weak solution of problem (15) coincides with the critical points of the functional J. It suffices now to show that the energy functional J is coercive to conclude that the problem (15) admits at least one weak solution.
According to hypothesis (17) and using the relation (2), we have

$$\begin{aligned}L(u) &= \sum_{k=1}^{T_1}\sum_{h=1}^{T_2}F\big(k,h,u(k,h)\big)\\ &\le \sum_{k=1}^{T_1}\sum_{h=1}^{T_2}C_8\left(1+|u(k,h)|^{r(k,h)}\right)\\ &\le C_8T_1T_2+C_8\sum_{k=1}^{T_1}\sum_{h=1}^{T_2}|u(k,h)|^{r(k,h)}\\ &\le C_9+C_8\sum_{k=1}^{T_1}\sum_{h=1}^{T_2}|u(k,h)|^{r^-}+C_8\sum_{k=1}^{T_1}\sum_{h=1}^{T_2}|u(k,h)|^{r^+}\\ &\le C_9+C_{10}\left(\|u\|^{r^-}+\|u\|^{r^+}\right).\end{aligned}$$

Therefore the inequality (14) becomes

$$J(u)\ge\frac{C_4}{p^+}\|u\|^{p^-}-C_5-\left(C_9+C_{10}\left(\|u\|^{r^-}+\|u\|^{r^+}\right)\right), \tag{20}$$

namely

$$J(u)\ge\frac{C_4}{p^+}\|u\|^{p^-}-C_{10}\left(\|u\|^{r^+}+\|u\|^{r^-}\right)-C_{11}, \tag{21}$$

where C_{10} and C_{11} are positive constants. Hence, since $p^->r+\ge r^-\ge 2$, the functional J is coercive.

Acknowledgement

The authors want to thank the anonymous referees for their valuables comments on the paper. They would also like to thank the PDE network in West Africa (Réseau EDP-MC).

References

[1] R. P. Agarwal. *Difference Equations and Inequalities: Theory, Methods and Applications*. Vol. 228 of Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker, New York, NY, USA, 2nd edition, 2000.

[2] G. Bonanno, G. M. Bisci and V. Radulescu. Arbitrarily small weak solutions for nonlinear eigenvalue problem in Orlicz-Sobolev spaces. *Monatshefte fur Mathematik* **165** (3-4) (2012) 305–318.

[3] X. Cai and J. Yu. Existence theorems for second-order discrete boundary value problems. *J. Math. Anal. Appl.* **320** (2006) 649–661.

[4] A. Castro and R. Shivaji. Nonnegative solutions for a class of radically symmetric nonpositone problems. *Proceedings of the American Mathematical Society*. **106** (1989) 735–740.

[5] Y. Chen, S. Levine and M. Rao. Variable exponent, linear growth functionals in image restoration. *SIAM Journal on Applied Mathematics* **66** (4) (2006) 1383–1406.

[6] L. Diening. Theoretical and numerical results for electrorheogical fluids. *[PhD. thesis], University of Freiburg*, 2002.

[7] A. Guiro, I. Ibrango and S. Ouaro. Weak Heteroclinic Solutions of Discrete Nonlinear Problems of Kirchhoff Type with Variable Exponents. *Nonlinear Dynamics and Systems Theory* **18** (1) (2018) 67–79.

[8] B. Koné and S. Ouaro. Weak solutions for anisotropic discrete boundary value problems. *J. Differ. Equ. Appl.* **16** (2) (2010) 1–11.

[9] M. Mihailescu, V. Radulescu and S. Tersian. Eigenvalue problems for anisotropic discrete boundary value problems. *J. Differ. Equ. Appl.* **15** (2009) 557–567.

[10] K. R. Rajagopal and M. Ruzicka. Mathematical modelling of electrorheological materials. *Continuum Mechanics and Thermodynamics*. **13** (2001) 59–78.

[11] M. Ruzicka. *Electrorheological Fluids: Modeling and Mathematical Theory*. Vol. 1748 of Lecture Notes in Mathematics, Springer,Berlin, Germany, 2000.

[12] G. Zhang and S. Liu. On a class of semipositone discrete boundary value problem. *J. Math. Anal. Appl.* **325** (2007) 175–182.

[13] V. Zhikov. Averaging of functionals in the calculus of variations and elasticity. *Mathematics of the USSR-Izvestiya*. **29** (1987) 33–66.

Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 100–113

Averaging Method and Boundary Value Problems for Systems of Fredholm Integro-Differential Equations

O. M. Stanzhytskyi [1]*, S. G. Karakenova [2], and R. E. Uteshova [3]

[1] *Taras Shevchenko National University of Kyiv, Ukraine.*
[2] *Al-Farabi Kazakh National University, Almaty, Kazakhstan.*
[3] *International Information Technology University, Almaty, Kazakhstan.*

Received: January 21, 2020; Revised: February 4, 2021

Abstract: In this paper, an analogue of Bogolyubov's first theorem of the averaging method for systems of Fredholm integro-differential equations is established. The averaging method is also applied to boundary value problems for such systems. It is shown that if a boundary value problem for an averaged system, which is a system of ordinary differential equations, has a solution, then the original problem is solvable as well.

Keywords: *Fredholm integro-differential equation; boundary-value problem; averaging method.*

Mathematics Subject Classification (2010): 47H10, 34B15, 34B27, 45G10, 45J99.

1 Introduction

In this paper, we study systems of Fredholm integro-differential equations

$$\frac{dx}{dt} = \dot{x} = \varepsilon X(t, x, \int_0^{\frac{T}{\varepsilon}} \varphi(t, s, x(s))ds) \tag{1}$$

subject to the Cauchy conditions

$$x(0) = x_0, \tag{1'}$$

or to the boundary conditions

$$F(x(0), x(\tfrac{T}{\varepsilon})) = 0, \tag{1''}$$

* Corresponding author: mailto:ostanzh@gmail.com

where $\varepsilon > 0$ is a small parameter, X and F are d-dimensional vector functions, φ is an m-dimensional vector function, $T > 0$ is a fixed number. We define the integral average $X_0(x)$ as

$$X_0(x) = \lim_{A\to\infty} \frac{1}{A} \int_0^A X\big(t, x, \varphi_1(t, x)\big)dt, \tag{2}$$

where $\varphi_1(t, x) = \int_0^t \varphi\big(t, s, x\big)ds$, and put the problems ($1'$) and ($1''$) in correspondence with the averaged problems

$$\dot{y} = \varepsilon X_0(y), \tag{3}$$

$$y(0) = x_0, \tag{$3'$}$$

$$F(y(0), y(\tfrac{T}{\varepsilon})) = 0, \tag{$3''$}$$

or, on the slow time scale $\tau = \varepsilon t$,

$$\frac{dy}{dt} = X_0(y), \quad F(y(0), y(T)) = 0. \tag{4}$$

The main results of the present paper are the justification of the averaging method for the Cauchy problem and the statement that if the problem (3) has a solution, then for small values of the parameter ε the problem (2) has a solution as well, in a small neighborhood of the solution of the boundary value problem (3). The exact statement of the problems and the results formulation are presented in the main part of the paper.

It should be noted that the averaging method has not lost its relevance and is widely used in the study of various problems, for example, optimal control [16,19], systems with a multi-valued right-hand side [13], and many others.

Integro-differential equations arise as mathematical models of various processes in natural sciences; for instance, in population dynamics [1], chemical kinetics, fluid dynamics [2, 12, 22], epidemiology [21]. Interaction of modeling objects with the environment leads to boundary value problems for integro-differential equations. These problems have been studied by many authors [3,4,6–8,17].

In [20], boundary value problems for systems of Volterra integro-differential equations are investigated by using the averaging method. It is shown that, for small ε, the existence of a solution of a boundary value problem for an averaged system (3) implies that of the original boundary value problem (1); the proximity between corresponding solutions is proven. The result of [20] is a generalization of the classical result [18] concerning boundary value problems for systems of ordinary differential equations.

Note that the averaging method has already been used for solving boundary value problems for systems of Volterra integro-differential equations (see [15] and the references therein). However, in these works only an estimate of proximity between the solutions of the exact and averaged problems was established. The very fact of the existence of a solution was only postulated.

The present work is devoted to the further development of the ideas [20] as applied to the studying boundary value problems for Fredholm equations. Analogues of the Bogolyubov first theorem for Fredholm equations, unlike those for Volterra equations, were obtained only in a very special case, when the right-hand part is a sum of an ordinary term and an integral part, and the integration is carried out over a finite interval (see [9, 11]). Again, the existence of a solution is only a postulate. However, this theorem plays an essential role in obtaining results analogous to those of [20].

The paper is organized as follows. In Section 2, the problem statement and the main results are formulated. Section 3 contains some auxiliary results which are also of independent interest. For a Cauchy problem for systems of Fredholm integro-differential equations, we prove the existence and uniqueness theorem and investigate the continuous dependence of solutions on initial data. Section 4 is devoted to the justification of the averaging method. In Section 5, the existence of a solution of the boundary value problem is proved.

2 Problem Statement and Main Results

Throughout the rest of this paper, we denote by $|\cdot|$ the norm of a vector in $\mathbb{R}^d$ and by $\|\cdot\|$ the matrix norm consistent with a vector norm.

The following theorem justifies the averaging method.

Theorem 2.1 *Let the following conditions hold:*

(1.1) $X(t,x,y)$ *is defined and continuous in a domain* $Q=\{t\ge 0, x\in\mathbb{R}^d, y\in\mathbb{R}^m\}$, *bounded by a constant* M *in this domain, and satisfies a Lipschitz condition with respect to the variables* x *and* y *in the following sense: there exists a function* $\alpha(t)\ge 0$ *such that*

$$|X(t,x,y)-X(t,x_1,y_1)|\le \alpha(t)\big(|x-x_1|+|y-y_1|\big); \tag{5}$$

(1.2) $\varphi(t,s,z)$ *is defined and continuous in* $Q_1=\{t\ge 0, s\ge 0, z\in\mathbb{R}^d\}$, *bounded by a constant* $M>0$, *and satisfies a Lipschitz condition in the following sense: there exists a function* $\mu(t,s)\ge 0$ *such that*

$$|\varphi(t,s,z)-\phi(t,s,z_1)|\le \mu(t,s)|z-z_1|. \tag{6}$$

Besides, there exists a constant $\mu_0>0$ *such that* $\mu(t,s)\le\mu_0$, $\int_0^\infty \mu(t,s)ds\le\mu_0$, *and*

$$\frac{1}{t}\int_0^t d\tau\int_0^\tau \mu(\tau,s)ds\to 0,\quad t\to\infty; \tag{7}$$

there also exists $\bar\varepsilon>0$ *such that for* $\varepsilon\in(0,\bar\varepsilon]$

$$\varepsilon\bigg(\int_0^{\frac{T}{\varepsilon}}\alpha(s)ds+\int_0^{\frac{T}{\varepsilon}}\alpha(s)\bigg(\int_0^{\frac{T}{\varepsilon}}\mu(\tau,s)d\tau\bigg)ds\bigg)<1; \tag{8}$$

(1.3) the limits (2) and

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\bigg(\int_\tau^\infty|\varphi(\tau,s,x)|ds\bigg)d\tau=0 \tag{9}$$

exist uniformly with respect to $x\in D$ *(*D *is a domain in* $\mathbb{R}^d$*); and*

(1.4) the averaged system (3) has a solution $y(\tau)=y(\varepsilon t)$ *that belongs to* D *together with some* ρ*-neighborhood, for* $\tau\in[0,T]$.

Then, for every $\eta>0$, *there exists* $\varepsilon_0=\varepsilon_0(\eta)\le\bar\varepsilon$ *such that for* $\varepsilon\in[0,\varepsilon_0]$ *the Cauchy problem* $x(0)=y(0)=x_0$ *for (1) has a unique solution* $x(t,\varepsilon)$ *defined on* $[0,\frac{T}{\varepsilon}]$, *and the following inequality holds:*

$$|y(\varepsilon t)-x(t,\varepsilon)|\le\eta,\quad t\in[0,\tfrac{T}{\varepsilon}]. \tag{10}$$

For boundary value problem $(1) - (1'')$, the following statement holds true.

Theorem 2.2 *Let conditions (1.1)-(1.3) hold. Suppose, in addition, that the averaged boundary value problem* (3)-$(3'')$ *has a solution* $y = y(\tau) = y(\varepsilon t)$ *belonging to* D *together with some* ρ*-neighborhood, in which* $X_0(x)$, $F(x, y)$ *have continuous partial derivatives* $\frac{\partial X_0(x)}{\partial x}$, $\frac{\partial F}{\partial x}$, *and* $\frac{\partial F}{\partial y}$, *and*

$$det \frac{\partial F_0(x_0)}{\partial x_0} \neq 0, \tag{11}$$

where $x_0 = y(0), F_0(x_0) = F(x_0, y(T, x_0))$.

Then there exists $\varepsilon_0 > 0$ *such that, for* $\varepsilon \in (0, \varepsilon_0)$, *boundary value problem* $(1') - (1'')$ *has a solution* $x(t, \varepsilon)$, *and one can specify a function* $\xi = \xi(\varepsilon) \to 0$, $\varepsilon \to 0$, *such that*

$$|x(t, \varepsilon) - y(\varepsilon t)| \leq \xi(\varepsilon), \quad t \in [0, \tfrac{T}{\varepsilon}]. \tag{12}$$

3 Cauchy Problem for Fredholm Integro-Differential Equations

In this section, we consider the Cauchy problem

$$\dot{x} = X(t, x, \int_0^T \varphi\big(t, s, x(s)\big) ds), \quad x(0) = x_0, \tag{13}$$

where $[0, T]$ is a fixed interval.

Theorem 3.1 *Let the following conditions be satisfied:*

(2.1) *the function* $X(t, x, y)$ *is defined in a domain* $Q = \{t \in [0, T], x \in \mathbb{R}^d, y \in D\}$ *(*D *is a domain in* $\mathbb{R}^m$*) and satisfies a Lipschitz condition*

$$|X(t, x, y) - X(t, x_1, y_1)| \leq \alpha(t)(|x - x_1| + |y - y_1|), \tag{14}$$

as well as a linear growth condition with respect to x, y*; that is, there exists a constant* $M > 0$ *such that, for* $t \in [0, T], x \in \mathbb{R}^d, y \in D$

$$|X(t, x, y)| \leq M(1 + |x| + |y|); \tag{15}$$

(2.2) *the function* $\varphi(t, s, z)$ *is defined and continuous in a domain* $Q_1 = \{t \in [0, T], s \in [0, T], z \in \mathbb{R}^d\}$, *bounded by a constant* M_1 *in* Q_1, *and, with respect to* z, *satisfies a Lipschitz condition*

$$|\varphi(t, s, z) - \varphi(t, s, z_1)| \leq \mu(t, s)|z - z_1|; \tag{16}$$

(2.3) *the inequality*

$$\int_0^T \alpha(t) dt + \int_0^T \alpha(t) \bigg(\int_0^T \mu(t, s) ds \bigg) dt < 1 \tag{17}$$

holds;

(2.4) *the region* D *contains a closed ball* $\bar{B}_{TM_1}(0)$ *of radius* TM_1, *centered at the origin.*

Then, for all $x_0 \in \mathbb{R}^d$, *the Cauchy problem (13) has a unique solution* $x(t, x_0)$ $(x(0, x_0) = x_0)$ *on* $[0, T]$, *which depends continuously on the initial data* x_0.

Remark 3.1 The behavior of systems of kind (13) is substantially different from that of similar systems of Volterra type. In [5, p.71], an example is provided for the following equation

$$\dot{x} = Ax + \frac{1}{2\pi}\int_0^{2\pi} Bx(s)ds + f(t), x \in \mathbb{R}^2, \quad t \in (0; 2\pi), \tag{18}$$

with matrices $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $f(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \end{pmatrix}$.

It was shown that (18) is solvable only if the following condition is met:

$$\int_0^{2\pi} \big(- f_1(t)\sin t + (1-\cos t) f_2(t)\big)dt = 0.$$

Note that equation (18) does not satisfy conditions of Theorem 3.1.

Proof. The proof of Theorem 3.1 falls into three parts.

1. Uniqueness. Let the Cauchy problem have two solutions $x(t)$ and $y(t)$ on $[0, T]$, such that $\sup\limits_{t\in[0,T]} |x(t) - y(t)| = \gamma > 0$. Note that $x(t)$ and $y(t)$ satisfy the equations

$$x(t) = x_0 + \int_0^t X\left(\tau, x(\tau), \int_0^T \varphi(\tau, s, x(s))ds\right)d\tau \tag{19}$$

and

$$y(t) = x_0 + \int_0^t X\left(\tau, y(\tau), \int_0^T \varphi(\tau, s, y(s))ds\right)d\tau, \tag{20}$$

respectively. Thus, by (15) and (16), we get the estimate

$$|x(t) - y(t)| \le \int_0^t \alpha(\tau)\big|x(\tau) - y(\tau)\big|d\tau + \int_0^t \alpha(\tau)\left(\int_0^T \mu(\tau, s)|x(s) - y(s)|ds\right)d\tau \le$$

$$\le\left(\int_0^T \alpha(\tau)d\tau + \int_0^T \alpha(\tau)\left(\int_0^T \mu(\tau, s)ds\right)d\tau\right)\sup_{t\in[0,T]}|x(t) - y(t)|,$$

which contradicts (17).

2. Existence. We construct a system of functions $\{x_n(t)\}$ in the following way: $x_0(t) \equiv x_0$ and $x_n(t)$ is a solution of the following Cauchy problem for a system of differential equations

$$\dot{x}_n = X\left(t, x_n, \int_0^T \varphi(t, s, x_{n-1}(s))ds\right), \quad x_n(0) = x_0.$$

Let us show that this sequence is defined correctly. Indeed, we have

$$\dot{x}_1 = X\left(t, x_1, \int_0^T \varphi(t, s, x_0)ds\right), \quad x_1(0) = x_0. \tag{21}$$

Since $\varphi(t, s, x_0)$ is continuous jointly in its variables, then $g(t) = \int\limits_0^T \varphi(t, s, x_0)ds$ is continuous with respect to t as well. Moreover, $\left|\int\limits_0^T \varphi(t, s, x_0)ds\right| \le M_1 T$; hence,

$\int\limits_0^T \varphi(t,s,x_0)ds \in D$. It follows that the function $Y(t,x) = X(t,x,g(t))$ is defined, continuous with respect to $t \in [0,T], x \in \mathbb{R}^d$, and satisfies a linear growth condition with respect to $x \in \mathbb{R}^d$.

So, the Cauchy problem (21) has the global solution $x_1(t)$ on $[0,T]$. In the same manner, we can see that the whole sequence $\{x_n(t)\}$ is defined on $[0,T]$ as well.

We proceed to show that $x_n(t)$ is uniformly convergent on $[0,T]$. We have

$$x_n(t) = x_0 + \int_0^t X\bigg(\tau, x_n(\tau), \int_0^T \varphi(\tau,s,x_{n-1}(s))ds\bigg)d\tau, \tag{22}$$

$$x_{n-1}(t) = x_0 + \int_0^t X\bigg(\tau, x_{n-1}(\tau), \int_0^T \varphi(\tau,s,x_{n-2}(s))ds\bigg)d\tau.$$

We thus get

$$\big|x_n(t) - x_{n-1}(t)\big| \le \int_0^T \alpha(\tau)d\tau \sup_{t\in[0,T]} |x_n(t) - x_{n-1}(t)| +$$

$$+ \int_0^T \alpha(\tau)\bigg(\int_0^T \mu(\tau,s)ds\bigg)d\tau \sup_{t\in[0,T]} |x_{n-1}(t) - x_{n-2}(t)|.$$

Then

$$\sup_{t\in[0,T]} |x_n(t) - x_{n-1}(t)| \le \frac{\int_0^T \alpha(\tau)\bigg(\int_0^T \mu(\tau,s)ds\bigg)d\tau}{1 - \int_0^T \alpha(\tau)d\tau} \sup_{t\in[0,T]} |x_{n-1}(t) - x_{n-2}(t)|. \tag{23}$$

But it follows from (17) that

$$\frac{\int_0^T \alpha(\tau)\bigg(\int_0^T \mu(\tau,s)ds\bigg)d\tau}{1 - \int_0^T \alpha(\tau)d\tau} = A < 1.$$

Thus, in view of (22), we can conclude that the sequence $x_n(t)$ uniformly converges to a limit function $x^*(t)$ on $[0,T]$. We can now easily obtain from (14) and (16) that

$$\int_0^t X\bigg(\tau, x_n(\tau), \int_0^T \varphi(\tau,s,x_{n-1}(s))ds\bigg)d\tau \to \int_0^t X\bigg(\tau, x^*(\tau), \int_0^T \varphi(\tau,s,x^*(s))ds\bigg)d\tau$$

as $n \to \infty$. From this it follows that $x^*(t)$ is a solution of the Cauchy problem (13).

3. Continuous dependence on initial data. Assume that continuous dependence does not hold. Then there exist $\varepsilon > 0$, a sequence of initial data x_n converging to x_0 as $n \to \infty$, and a sequence $\{t_n\}$, $t_n \in (0,T]$ such that

$$|x(t_n,x_n) - x(t_n,x_0)| = \varepsilon. \tag{24}$$

Here $x(t,x_n)$ is a solution of the system (13) subject to initial data $x(0,x_n) = x_n$. Let us show that the sequence $x(t,x_n)$ is compact in $C([0,T])$. Indeed, from (15) we obtain

$$|x(t,x_n)| \le |x_n| + \int_0^t M|1 + |x_n(\tau)|d\tau + \int_0^t |\varphi(\tau,s,x_n(s))ds|d\tau \le$$

$$\leq |x_n| + MT + M\int_0^t |x_n(\tau)|d\tau + T^2 MM_1.$$

By Gronwall's lemma, we get

$$|x(t,x_n)| \leq \Big(|x_n| + MT + T^2MM_1\Big)e^{MT} \leq C \tag{25}$$

due to the boundedness of the sequence $\{x_n\}$.

Further, for $t_1 < t_2$, $t_1, t_2 \in [0,T]$,

$$|x(t_2,x_n) - x(t_1,x_0)| \leq \int_{t_1}^{t_2} M(1 + C + TM_1), \tag{26}$$

whence it follows that the sequence $\{x(t,x_n)\}$ is equicontinuous. Consequently, $\{x(t,x_n)\}$ contains a uniformly convergent on $[0,T]$ subsequence $\{x(t,x_{n_k})\}$. It is clear that this subsequence can be chosen so that the number sequence $\{t_{n_k}\}$ converges simultaneously to some $t^* \in [0,T]$. Thus, $x(t,n_k) \rightrightarrows x^*(t)$, $n_k \to \infty$, and

$$x(t,n_k) = x_{n_k} + \int_0^t X\Big(\tau, x_{n_k}(\tau), \int_0^T \varphi(\tau,s,x_{n_k}(s)ds\Big)d\tau. \tag{27}$$

Letting $n_k \to \infty$ in (27), we obtain

$$x^*(t) = x_0 + \int_0^t X\Big(\tau, x^*(\tau), \int_0^T \varphi(\tau,s,x^*(s)ds\Big)d\tau.$$

Hence $x^*(t)$ is a solution of the Cauchy problem (13) as well. Let us show that $x^*(t)$ does not coincide identically with $x(t,x_0)$. Taking into account that $x(t,x_{n_k})$ converges uniformly to $x^*(t)$, which is continuous, from the inequality

$$|x(t_{n_k},x_{n_k}) - x^*(t^*)| \leq |x(t_{n_k},x_{n_k}) - x^*(t_{n_k})| + |x^*(t_{n_k}) - x^*(t^*)|$$

we conclude that $x(t_{n_k},x_{n_k}) \to x^*(t^*)$, $n_k \to \infty$. Therefore, passing to the limit, as $n_k \to \infty$, in (24), we get

$$|x^*(t^*) - x(t^*,x_0)| = \varepsilon. \tag{28}$$

Note that $t^* \neq 0$, since otherwise (24) would not hold for large n. Thus (28) contradicts the uniqueness of a solution of the Cauchy problem. The proof is complete. □

4 Averaging Method for Systems (1)

In this section, we prove Theorem 2.1 on a justification of the averaging method.

4.1 Averaging Lemma

Assume the condition (1.4) of Theorem 2.1 to be fulfilled. Fix $K > 0$.

Definition 4.1 We say that a function $a(t,\varepsilon)$ belongs to a class A_K if:

(i) $a(t,\varepsilon)$ is defined for $\varepsilon > 0$, $t \geq 0$, and takes on values in a ρ-neighborhood of $y(\tau)$, which is a solution of the averaged Cauchy problem (3)–(3');

(ii) for $t \geq 0, s \geq 0$, and $\varepsilon > 0$, the following inequality holds:

$$|a(t,\varepsilon) - a(s,\varepsilon)| \leq K\varepsilon|t-s|. \tag{29}$$

Lemma 4.1 *Let the conditions of Theorem 2.1, except (8), be fulfilled. Then, for every $\eta > 0$, there exists $\varepsilon_0 = \varepsilon_0(\eta, K)$ such that, for $\varepsilon \in (0, \varepsilon_0]$, the system*

$$\dot{x} = \varepsilon X\left(t, x, \int_0^{\frac{T}{\varepsilon}} \varphi\big(t, s, a(s,\varepsilon)\big)ds\right) \tag{30}$$

has a solution $x(t)$, $x(0) = y(0) = x_0$, defined on $[0, \frac{T}{\varepsilon}]$, and the inequality

$$|y(\varepsilon t) - x(t)| \leq \eta, \quad t \in [0, \tfrac{T}{\varepsilon}], \tag{31}$$

holds.

Remark 4.1 In the above lemma, ε_0 does not depend on x_0 and is uniform throughout the class A_K.

Proof. As in the proof of Theorem 3.1, we can show that, for all ε, the Cauchy problem

$$\dot{x} = \varepsilon X\left(t, x, \int_0^{\frac{T}{\varepsilon}} \varphi\big(t, s, a(s,\varepsilon)\big)ds\right), \quad x(0) = x_0, \tag{32}$$

has a solution $x(t,\varepsilon)$ defined on $[0, \frac{T}{\varepsilon}]$.

Fix $\eta > 0$. Let us estimate the difference between $x(t)$ and $y(t)$ (for the convenience of notation, we will omit the dependence on ε). We have

$$\begin{aligned}
\big|x(t) - y(t)\big| &= \varepsilon \int_0^t \left[X\left(\tau, x(\tau), \int_0^{\frac{T}{\varepsilon}} \varphi(\tau, s, a(s))ds - X\left(\varepsilon, x(\tau), \int_0^{\tau} \varphi(\tau, s, a(s,\varepsilon)\right)ds\right] d\tau + \\
&+ \varepsilon \int_0^t \left[X\left(\tau, x(\tau), \int_0^{\tau} \varphi(\tau, s, a(s))ds\right) - X_0(y(\tau))\right] d\tau = \\
&= I_1(t) + \varepsilon \int_0^t \left[X\left(\tau, x(\tau), \int_0^{\tau} \varphi(\tau, s, a(s))ds\right) - X\left(\tau, y(\tau), \int_0^{\tau} \varphi(\tau, s, y(s))ds\right)\right] d\tau + \\
&+ \varepsilon \int_0^{\tau} \left[X\left(\tau, y(\tau), \int_0^{\tau} \varphi(\tau, s, y(s))ds\right) - X\left(\tau, y(\tau), \int_0^{\tau} \varphi(\tau, s, y(\tau))ds\right)\right] d\tau + \\
&+ \varepsilon \int_0^{\tau} \left[X(\tau, y(\tau), \int_0^{\tau} \varphi(\tau, s, y(\tau))ds\right) - X_0(y(\tau))\right] d\tau = I_1 + I_2 + I_3 + I_4.
\end{aligned} \tag{33}$$

Let us now estimate each term of (33) separately. Due to the Lipschitz condition in (1.1), we have

$$\begin{aligned}
\big|I_1(t)\big| &\leq \varepsilon \int_0^t L\left|\int_0^{\frac{T}{\varepsilon}} \varphi\big(\tau, s, a(s)\big)ds - \int_0^{\tau} \varphi(\tau, s, a(s))ds\right| d\tau \leq \\
&\leq \varepsilon L \int_0^t \left(\int_{\tau}^{\frac{T}{\varepsilon}} \big|\varphi\big(\tau, s, a(s)\big)\big|ds\right) d\tau.
\end{aligned} \tag{34}$$

Let us divide $[0, \frac{T}{\varepsilon}]$ into n subintervals of equal length by points t_i, $t_0 = 0 < t_1 < \dots < t_n = \frac{T}{\varepsilon}$. Then

$$\varepsilon L \int_0^t \left(\int_\tau^{\frac{T}{\varepsilon}} \big|\varphi(\tau, s, a(s))\big| ds \right) d\tau = \varepsilon L \int_0^t \left(\sum_i^{n-1} \int_{t_i}^{t_{i+1}} \big|\varphi(\tau, s, a(s)) - \varphi(\tau, s, a(t_i))\big| ds + \right.$$

$$\left. + \sum_i^{n-1} \int_{t_i}^{t_{i+1}} \big|\varphi(\tau, s, a(s)) - \varphi(\tau, s, a(t_i))\big| ds \right) d\tau = I_{11}(t) + I_{12}(t). \tag{35}$$

Here, for every $\varepsilon \in [0, t]$, the summation is performed over such indices i that $[t_i, t_{i+1})$ cover the interval $[\tau, \frac{T}{\varepsilon}]$.

We now proceed to estimate each term of (35) separately. To estimate I_{11}, note that, by virtue of (1.2) and (29),

$$\big|\varphi(\tau, s, a(s)) - \varphi(\tau, s, a(t_i))\big| \le \mu_0 \varepsilon K |s - t_i| \le \varepsilon \mu_0 \varepsilon K \frac{T}{\varepsilon n} = \mu_0 \frac{KT}{n}. \tag{36}$$

Then

$$|I_{11}(t)| \le \varepsilon L \int_0^t \left(\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mu_0 \frac{KT}{n} ds \right) d\tau \le \varepsilon L \frac{T^2}{\varepsilon} \frac{\mu_0 K}{n} = \frac{\mu_0 K T^2 L}{n}. \tag{37}$$

We now turn to estimation of $I_{11}(t)$. Observe that, by virtue of (9), there exists such a function $\beta(t)$, continuous and decreasing monotonically to zero as $t \to \infty$, that

$$\int_0^t \left(\int_\tau^{\frac{\tau}{\varepsilon}} \Big| \varphi\big(\tau, s, a(s_i)\big) \Big| ds \right) d\tau \le t\beta(t). \tag{38}$$

If t belongs to any subinterval $[t_i, t_{i+1}]$ except the first one, then it follows from (38) that

$$|I_{12}(t)| \le \varepsilon L t \beta(t) \le LT\beta(\tfrac{T}{\varepsilon n}). \tag{39}$$

For fixed n, the right-hand side of (39) approaches zero as $\varepsilon \to 0$.

If $t \in [0, t_1]$, we obtain by virtue of (38) and Dini's theorem that

$$|I_{12}(t)| \le \varepsilon \beta(t) \le \sup_{\tau \in [0,T]} \tau \beta(\tfrac{\tau}{\varepsilon}) \to 0, \quad \varepsilon \to 0. \tag{40}$$

Let us estimate the term $I_2(t)$ of (33). We have

$$|I_2(t)| \le \varepsilon L \int_0^t |x(\tau) - y(\tau)| d\tau + \varepsilon L \int_0^t \left(\int_0^\tau |\varphi(\tau, s, a(s)) - \varphi(\tau, s, y(s))| ds \right) d\tau.$$

But, under the conditions of Lemma 4.1, $a(s, \varepsilon)$ belongs to a ρ-neighborhood of $y(\tau)$, which is a bounded on $[0, T]$ solution of the averaged problem. Therefore, for all $\varepsilon > 0$, $s \ge 0$, the function $a(s, \varepsilon)$ is bounded by a constant $R = R(\rho, y(\tau))$ independent of that function. Consequently,

$$|I_2(t)| \le \varepsilon L \int_0^t \left(\int_0^\tau \mu(\tau, s) |a(s) - y(s)| ds \right) d\tau \le 2R\varepsilon \int_0^t \left(\int_0^\tau \mu(\tau, s) ds \right) d\tau.$$

Similarly to the previous case, it follows from (7) that there exists a function $\beta_1(t)$, monotonically approaching zero as $t \to \infty$, such that

$$|I_2(t)| \le 2R\varepsilon t \beta_1(t) \le 2RT\beta_1(t). \tag{41}$$

Hence, for chosen $\eta > 0$, there exists T_0 such that

$$|I_2(t)| \leq \tfrac{\eta}{4} \tag{42}$$

for $t \geq T_0$.

Obviously, we can assume $T_0 \in \big[0, \frac{T}{\varepsilon}\big]$. The estimate (42) for $t \in \big[0, T_0\big]$ is obtained by choosing a small ε_0, taking into account that $\int_0^\infty \mu(t,s)ds \leq \mu_0$.

An estimate of $I_3(t)$ is obtained similarly to that of $I_2(t)$ due to the fact that the function $y(\tau)$ is bounded on $[0,T]$.

The term $I_4(t)$ is estimated in the same way as in the proof of Theorem 3.3 in [10], taking into account the first condition of (1.3). The method of estimation is similar to that for $I_1(t)$.

For given $\eta > 0$, we choose n and T_0 large enough to make the terms (39) sufficiently small to satisfy the estimate (42). Once such n and T_0 are fixed, we choose $\varepsilon_0 > 0$ such that, for $\varepsilon \leq \varepsilon_0$, the terms (39),(40) and (42) for $t \in [0, T_0]$ are sufficiently small. The application of Gronwall's inequality completes the proof. □

4.2 Proof of Theorem 2.1

Proof. Choose $\eta > 0$ such that $\eta < \frac{\rho}{2}$ and keep it fixed. Let us construct a functional sequence $\{x_n(t,\varepsilon)\}$ in the following way: $x_0(t) = x_0$ and $x_n(t,\varepsilon)$, for every $\varepsilon > 0$, are defined recurrently as solutions of the Cauchy problems

$$\dot{x}_n = \varepsilon X\bigg(t, x_n, \int_0^{\frac{T}{\varepsilon}} \varphi(t,s,x_{n-1}(s,\varepsilon))ds\bigg). \tag{43}$$

As in the proof of Theorem 3.1, by virtue of (8), we can show that, for all $0 < \varepsilon < \bar{\varepsilon}$, the sequence $\{x_n(t,\varepsilon)\}$ converges uniformly with respect to $t \in [0, \frac{T}{\varepsilon}]$ as $n \to \infty$, and its limit function $x(t,\varepsilon)$ is a unique solution of the Cauchy problem for equation (1), $x(0) = x_0$, on $[0, \frac{T}{\varepsilon}]$. Clearly, the following estimate is valid for functions $\{x_n(t,\varepsilon)\}$:

$$|x_n(t_2,\varepsilon) - x_n(t_1,\varepsilon)| \leq \varepsilon M|t_2 - t_1|. \tag{44}$$

Further, the system

$$\dot{x}_1(t,\varepsilon) = \varepsilon X\bigg(t, x_1(t,\varepsilon), \int_0^{\frac{T}{\varepsilon}} \varphi(t,s,x_0)ds\bigg), \tag{45}$$

$$x_1(0,\varepsilon) = x_0,$$

is a system of kind (30) in the Averaging Lemma with the function $a(t,\varepsilon) = x_0$, which obviously satisfies the conditions of Lemma 4.1.

Thus, for chosen $\eta > 0$, there exists $\varepsilon_0 \leq \bar{\varepsilon}$ such that, for $\varepsilon < \varepsilon_0$, the estimate

$$|y(\varepsilon t) - x_1(t,\varepsilon)| \leq \eta < \frac{\rho}{2}, \quad t \in [0, \tfrac{T}{\varepsilon}], \tag{46}$$

holds.

By (44), the function $x_1(t,\varepsilon)$ belongs to the class A_K introduced above, with $K = M$. Therefore, the system of equations for determining $x_2(t,\varepsilon)$ is a system of kind (30) with $a(t,\varepsilon) = x_1(t,\varepsilon)$. Hence, for $\varepsilon \leq \varepsilon_0$, the function $x_2(t,\varepsilon)$ satisfies inequality (46) as well.

Now, setting $a(t,\varepsilon) = x_{n-1}(t,\varepsilon)$ for every n, we can conclude that all functions $x_n(t,\varepsilon)$ satisfy (26) with $K = M$, and hence

$$|x_n(t,\varepsilon) - y(\varepsilon t)| \le \eta < \frac{\rho}{2}, \quad t \in [0, \tfrac{T}{\varepsilon}], \tag{47}$$

for all $\varepsilon \le \varepsilon_0 = \varepsilon_0(\eta, M)$. We can therefore choose one and the same ε_0 for all n. In (47), passing to the limit as $n \to \infty$ for every $\varepsilon \le \varepsilon_0$ and taking into account the convergence of $x_n(t,\varepsilon)$ to $x(t,\varepsilon)$, we obtain the assertion of Theorem 2.1. $\square$

5 Proof of Theorem 2.2

Proof. Let $y(\tau) = y(\varepsilon t)$ be a solution of the boundary value problem (3) – (3″). According to (1.4), for $t \in [0,T]$ this solution belongs to a domain D with some ρ-neighborhood.

Let $x_0 = y(0)$ be an initial value of this solution. We now seek a solution of (1) – (1″) in the form

$$x(t,\varepsilon) = x(t, x_0 + \bar{x}, \varepsilon), \tag{48}$$

where $\bar{x}$ is chosen in some neighborhood of zero. We consider a solution $y(\tau, x_0 + \bar{x})$, $y(0, x_0+\bar{x}) = x_0+\bar{x}$ of the averaged problem. It follows from condition (8) and definition (2) of the averaged system that the function $X_0(x)$ satisfies the Lipschitz condition with a constant $L \le \frac{1}{T}$ (according to the problem statement, T is fixed).

By Gronwall's lemma, the following estimate

$$|y(\tau) - y(\tau, x_0 + \bar{x})| \le |\bar{x}| e^{LT} \tag{49}$$

holds until $y(\tau, x_0 + \bar{x})$ reaches the boundary of D. Therefore, if

$$|\bar{x}| < \frac{\rho}{2} e^{-LT}, \tag{50}$$

then a solution $y(\tau, x_0 + \bar{x})$ exists for $\tau \in [0,T]$ and belongs to a $\frac{\rho}{2}$-neighborhood of $y(\tau)$. Hence $y(\tau, x_0 + \bar{x})$, together with its $\frac{\rho}{2}$-neighborhood, belong to D.

We determine an unknown parameter $\bar{x}$ in (48) from the equation

$$F(x_0 + \bar{x}, x(\tfrac{T}{\varepsilon}, x_0 + \bar{x}, \varepsilon)) = 0. \tag{51}$$

Note that Theorem 2.1 applies to the solution $x(t, x_0 + \bar{x}, \varepsilon)$. Therefore, for $\varepsilon > 0$ sufficiently small, $x(t, x_0 + \bar{x}, \varepsilon)$ exists on $[0, \frac{T}{\varepsilon}]$. Moreover, for any $\eta > 0$, there exists $\varepsilon_0(\eta) > 0$ such that, for $\varepsilon \in (0, \varepsilon_0)$, the following estimate is valid:

$$|x(t, x_0 + \bar{x}, \varepsilon) - y(\varepsilon t, x_0 + \bar{x})| \le \eta(\varepsilon) \to 0, \quad \varepsilon \to 0. \tag{52}$$

From this we see that, for $\varepsilon \in (0, \varepsilon_0)$, the mapping $F(x_0 + \bar{x}, x(\frac{T}{\varepsilon}, x_0 + \bar{x}, \varepsilon))$, with respect to $\bar{x}$, is well-defined in a ball $B_r(0)$, where $r \le \frac{\rho}{2} e^{-LT}$.

We note also that the points $x_0 + \bar{x}$ and $x(\frac{T}{\varepsilon}, x_0 + \bar{x}, \varepsilon)$ belong to the ρ-neighborhood of $y(\tau)$, for $\varepsilon \in (0, \varepsilon_0)$.

Then, by virtue of conditions (1.4) imposed upon the function $F(x,y)$, there exists a constant $N(r) > 0$ such that $\|\frac{\partial F}{\partial x}\| \le N(r)$ and $\|\frac{\partial F}{\partial y}\| \le N(r)$, for $\bar{x} \in B_r(0)$.

Let us represent $F(x_0 + \bar{x}, x(\frac{T}{\varepsilon}, x_0 + \bar{x}, \varepsilon))$ in the following way:

$$F(x_0 + \bar{x}, x(\tfrac{T}{\varepsilon}, x_0 + \bar{x}, \varepsilon)) = F(x_0 + \bar{x}, x(\tfrac{T}{\varepsilon}, x_0 + \bar{x}, \varepsilon)) -$$

$$-F(x_0+\bar{x}, y(T, x_0+\bar{x})) + F(x_0+\bar{x}, y(T, x_0+\bar{x}))-$$
$$-F(x_0, y(T, x_0)) = R_1(\bar{x}, \varepsilon) + M_1(\bar{x}, \varepsilon).$$

For $R_1(\bar{x}, \varepsilon)$, the estimate

$$|R_1(\bar{x}, \varepsilon)| \le |N(r)(x(\tfrac{T}{\varepsilon}, x_0+\bar{x}, \varepsilon) - y(T, x_0+\bar{x}))| \le N(r)\eta(\varepsilon) \to 0, \quad \varepsilon \to 0, \tag{53}$$

holds due to (52).

Under conditions of Theorem 2.2, solutions of the averaged problem depend smoothly on initial data, hence

$$M_1(\bar{x}, \varepsilon) = \Big(\frac{\partial F(x_0, y(T, x_0))}{\partial x} + \frac{\partial F(x_0, y(T, x_0))}{\partial y} \cdot \frac{\partial y(T, x_0)}{\partial x_0}\Big)\bar{x} +$$
$$+ \int_0^1 \Big(\frac{\partial F(x_0+s\bar{x}, y(T, x_0+s\bar{x}))}{\partial x} - \frac{\partial F(x_0, y(T, x_0))}{\partial x}\Big)\bar{x}ds +$$
$$+ \int_0^1 \Big(\frac{\partial F(x_0+s\bar{x}, y(T, x_0+s\bar{x}))}{\partial y} \cdot \frac{\partial y(T, x_0+s\bar{x})}{\partial z}\Big|_{z=x_0+s\bar{x}} -$$
$$- \frac{\partial F(x_0, y(T, x_0))}{\partial y} \cdot \frac{\partial y(T, x_0)}{\partial z}\Big|_{z=x_0}\Big)\bar{x}ds =$$
$$= \Big(\frac{\partial F(x_0, y(T, x_0))}{\partial x} + \frac{\partial F(x_0, y(T, x_0))}{\partial y} \cdot \frac{\partial y(T, x_0)}{\partial z}\Big|_{z=x_0}\Big)\bar{x} + R_2(\bar{x})\bar{x} + R_3(\bar{x})\bar{x}. \tag{54}$$

Let us consider each term of (54) separately. Using the notation of $F_0(x_0)$ in (11), the first term can be represented as

$$\Big(\frac{\partial F(x_0, y(T, x_0))}{\partial x} + \frac{\partial F(x_0, y(T, x_0))}{\partial y} \cdot \frac{\partial y(T, x_0)}{\partial z}\Big|_{z=x_0}\Big)\bar{x} = \frac{\partial F_0}{\partial x_0}\bar{x}.$$

Regarding $R_2(\bar{x})$, by the uniform continuity of partial derivatives and (49), for $|\bar{x}| \le r$, we get the estimate

$$|R_2(\bar{x})| \le \delta(r) \to 0, \quad r \to 0, \tag{55}$$

where $r \le \frac{\rho}{2}e^{-LT}$.

To estimate $R_3(\bar{x})$, note that the derivative $\frac{\partial y(T,z)}{\partial z}$ with respect to initial data satisfies a linear variational equation and hence is a continuous function of a parameter z. So, similarly as above, for $|\bar{x}| \le r$, we get the estimate

$$|R_3(\bar{x})| \le \delta_1(r) \to 0, \quad r \to 0. \tag{56}$$

Now, equation (51) for determining $\bar{x}$ can be represented in the form

$$|\bar{x}| = -\Big(\frac{\partial F_0}{\partial x_0}\Big)^{-1}\big(R_1(\bar{x}, \varepsilon) + \big(R_2(\bar{x}) + R_3(\bar{x})\big)\bar{x}\big),$$

or

$$\bar{x} = \Big(\frac{\partial F_0}{\partial x_0}\Big)^{-1} M(\bar{x}, \varepsilon), \tag{57}$$

where $M(\bar{x},\varepsilon)$ satisfies the inequality

$$|M(\bar{x},\varepsilon)| \le N(r)\eta(\varepsilon) + \delta_2(r)\bar{x}, \tag{58}$$

where $\eta(\varepsilon) \to 0$, $\varepsilon \to 0$, $\delta_2(r)$, $r \to 0$.

Let $C = \|(\frac{\partial F_0}{\partial x_0})^{-1}\|$. Choose r so that

$$\delta_2(r) \le \frac{1}{2}, \tag{59}$$

and then choose $\varepsilon_1 \le \varepsilon_0$ such that

$$\eta(\varepsilon) \le \frac{r}{2CN(r)}. \tag{60}$$

Then, for $|\bar{x}| \le r$, from (40) we obtain

$$\left\| \left(\frac{\partial F_0}{\partial x_0}\right)^{-1} M(\bar{x},\varepsilon)\right\| \le C(N(r)\eta(\varepsilon) + \delta_2(r)|\bar{x}|) \le \frac{r}{2} + \frac{r}{2} = r.$$

Thus, if (59) and (60) hold, $(\frac{\partial F_0}{\partial x_0})^{-1}M(\bar{x},\varepsilon)$ maps the ball $B_0(r)$ into itself. Note also that, by Theorem 3.1, there exists a solution $x(t, x_0 + \bar{x}, \varepsilon)$ that is unique on $\left[0, \frac{T}{\varepsilon}\right]$ and continuously depends on $\bar{x}$. Therefore the mapping $(\frac{\partial F_0}{\partial x_0})^{-1}M(\bar{x},\varepsilon)$ is well-defined and continuous, and, by Brouwer's theorem, it has a fixed point $\bar{x}^* = \bar{x}^*(\varepsilon, r)$, which is the initial value of the solution of the boundary value problem $(1) - (1'')$.

Let us now pick r, as a function of a parameter ε, so that $r(\varepsilon) \to 0, \varepsilon \to 0$. We then pick $\varepsilon_1 \le \varepsilon_0$ so that the function $\eta(\varepsilon)$ in (53) satisfies the inequality

$$\frac{\eta(\varepsilon)}{r(\varepsilon)} \le \frac{1}{2CN(r(\varepsilon))}.$$

Note that such a choice is possible, since a function $N(r(\varepsilon))$, by which the partial derivatives $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ are bounded in the ball $B_0(r)$, does not increase as $r(\varepsilon)$ decreases.

The estimate (12) now follows from (49) and (52), and the proof is complete. □

References

[1] J. F. M. Al-Omari and S. A. Gourley. A nonlocal reaction-diffusion model for a single species with stage structure and distributed maturation delay. *European J. Appl. Math.* **16** (2005) 37–51.

[2] A. Alownel, K. Al-Khaled and M. Al-Towiq. Reliable algorithms for solving integro-differential equations with applications. *Int. J. of Comp. Math.* **87** (7) (2010) 1538–1554.

[3] A. A. Boichuk and I. A. Golovatska. Weakly nonlinear systems of integro-differential equations. *Nonlinear Oscillations* **16** (3) (2013) 314–321.

[4] A. A. Boichuk and I. A. Golovatska. Boundary value problems for integro-differential equations. *Nonlinear Oscillations* **16** (4) (2013) 460–474.

[5] A. A. Boichuk and A. M. Samoilenko. *Generalized Inverse Operators and Fredholm Boundary-Value Problems.* USP, Utrecht, Boston, 2004.

[6] D. S. Dzhumabaev. Necessary and sufficient conditions for the solvability of linear boundary-value problems for the Fredholm integro-differential equations. *Ukrainian Math. J.* **66** (8) (2015) 1200–1219.

[7] D. S. Dzhumabaev. On one approach to solve the linear boundary value problems for Fredholm integro-differential equations. *J. Comput. Appl. Math.* **294** (2016) 342–357.

[8] D. S. Dzhumabaev. New general solutions to linear Fredholm integro-differential equations and their applications on solving the boundary value problems. *J. Comput. Appl. Math.* **327** (2018) 79–108.

[9] A. N. Filatov. Averaging in systems of integral and integro-differential equations. In: *Research in analytical mechanics*. Nauka, Tashkent, 1965, 135–179. [Russian]

[10] A. N. Filatov and L. V. Sharova. *Integral Inequalities and Theory of Nonlinear Oscillations.* Moscow: Nauka, 1976. [Russian]

[11] A. N. Filatov and A. G. Umarov. On averaging method in systems of integro-differential equations. *Reports of Academy of Sciences, Uzbek SSR* **327** (6) (1985) 12–14. [Russian]

[12] P. K. Kythe and P. Puri. *Computational Methods for Linear Integral Equations.* Univ. of New Orleans, New Orleans, 1992.

[13] A. A. Martynyuk. Analysis of a set of trajectories of generalized standard systems: averaging technique. *Nonlinear Dyn. Syst. Theory* **17** (1) (2017) 29–41.

[14] Yu. A. Mitropolsky. *Averaging Method in Nonlinear Mechanics.* Naukova Dumka, Kiyv, 1971. [Russian]

[15] Yu. A. Mitropolsky, D. D. Bainov and S. D. Milusheva. Application of averaging method in solving boundary value problems for ordinary differential equations and integro-differential equations. *Math. Physics* **25** (1979) 3–22. [Russian]

[16] T. V. Nosenko and O. M. Stanzhytskyi. Averaging method in some problems of optimal control. *Nonlinear Oscill.* **11** (2008) 539–547.

[17] A. M. Samoilenko, A. A. Boichuk and S. A. Krivosheya. Boundary value problems for systems of integro-differential equations with degenerate kernel. *Ukrainian Math. J.* **48** (11) (1996) 1576–1579.

[18] A. M. Samoilenko and R. I. Petrishin. Averaging method in some boundary value problems. *Diff. Equat.* **25** (6) (1989) 956–964.

[19] A. N. Stanzhitskii and T. V. Dobrodzii. Study of optimal control problems on the half-line by the averaging method. *Diff. Equat.* **47** (2) (2011) 264–277.

[20] A. N. Stanzhytskyi, S. T. Mynbayeva and N. A. Marchuk. Averaging in boundary value problems for systems of differential and integro-differential equations. *Ukrainian Math. J.* **72** (2020) 277–301.

[21] H. R. Thilme. A model for the spatio-spread of an epidemic. *J. Math. Biol.* **4** (1977) 337–351.

[22] A. M. Wazwaz. A comparison study between the modified decomposition method and the traditional methods for solving nonlinear integral equations. *Appl. Math. and Comp.* **181** (2) (2006) 1703–1712.