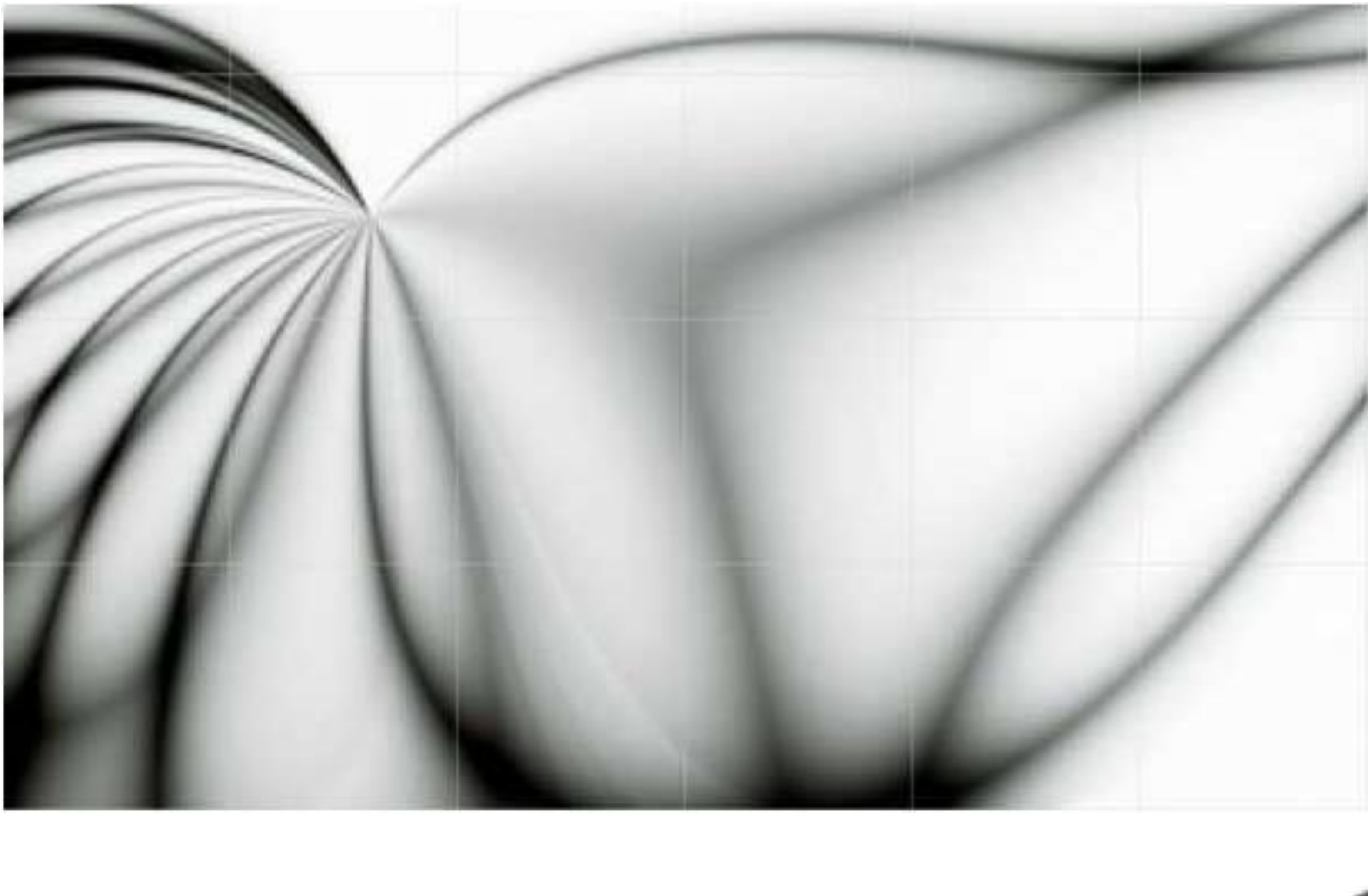
ISSN:2146-0957
eISSN:2146-5703

Volume:8 Number:1
January 2018

An International Journal of Optimization and Control:

Theories & Applications



www.ijocta.com

www.ijocta.com
info@ijocta.com

An International Journal of Optimization and Control: Theories & Applications

Volume: 8, Number: 1
January 2018

Publisher & Owner *(Yayımcı & Sahibi)*:

Prof. Dr. Ramazan YAMAN
Balikesir University, Faculty of Engineering, Department of Industrial Engineering, Cagis Campus, 10145, Balikesir, Turkey
Balıkesir Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Çağış Kampüsü, 10145, Balıkesir, Türkiye

ISSN: 2146-0957
eISSN: 2146-5703

Press *(Basımevi)*:

Bizim Dijital Matbaa (SAGE Publishing), Kazım Karabekir Street, Kültür Market, No:7 / 101-102, İskitler, Ankara, Turkey
Bizim Dijital Matbaa (SAGE Yayıncılık), Kazım Karabekir Caddesi, Kültür Çarşısı, No:7 / 101-102, İskitler, Ankara, Türkiye

Date Printed *(Basım Tarihi)*:

January 2018
Ocak 2018

Responsible Director (Sorumlu Müdür):

Prof. Dr. Ramazan YAMAN

IJOCTA is an international, bi-annual, and peer-reviewed journal indexed/ abstracted by *(IJOCTA, yılda iki kez yayımlanan ve aşağıdaki indekslerce taranan/dizinlenen uluslararası hakemli bir dergidir)*:
Cabell's Directories, DOAJ, EBSCO Databases, JournalSeek, Google Scholar, Index Copernicus, International Abstracts in Operations Research, JournalTOCs, Mathematical Reviews (MathSciNet), ProQuest, Scopus, Ulakbim Engineering and Basic Sciences Database (Tubitak), Ulrich's Periodical Directory, and Zentralblatt Math.

iThenticate and ijocta.balikesir.edu.tr are granted by Balikesir University.

An International Journal of Optimization and Control:

Theories & Applications

Volume: 8 Number: 1
January 2018

CONTENTS

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.1, pp.1-7 (2018)
https://doi.org/10.11121/ijocta.01.2018.00540

RESEARCH ARTICLE

Novel solution methods for initial boundary value problems of fractional order with conformable differentiation

Mehmet Yavuz*

Department of Mathematics-Computer Sciences, Necmettin Erbakan University, Turkey
mehmetyavuz@konya.edu.tr

ARTICLE INFO

Article history:
Received: 11 September 2017
Accepted: 2 October 2017
Available Online: 8 October 2017

Keywords:
Conformable fractional derivative
Approximate-analytical solution
Adomian decomposition method
Modified homotopy perturbation method

AMS Classification 2010:
35R11, 49M27

ABSTRACT

In this work, we develop a formulation for the approximate-analytical solution of fractional partial differential equations (PDEs) by using conformable fractional derivative. Firstly, we redefine the conformable fractional Adomian decomposition method (CFADM) and conformable fractional modified homotopy perturbation method (CFMHPM). Then, we solve some initial boundary value problems (IBVP) by using the proposed methods, which can analytically solve the fractional partial differential equations (FPDE). In order to show the efficiencies of these methods, we have compared the numerical and exact solutions of the IBVP. Also, we have found out that the proposed models are very efficient and powerful techniques in finding approximate solutions for the IBVP of fractional order in the conformable sense.

1. Introduction

Fractional differential equations have an important role in modelling and describing certain problems such as diffusion processes, chemistry, engineering, economic, material sciences and other areas of application. Zhang [1] used a finite difference method for the fractional PDEs. Ibrahim [2] interpreted holomorphic solutions for nonlinear singular fractional differential equations. Odibat and Momani [3, 4] applied several different types of methods to fractional PDEs and compared the results they obtained.

On the other hand, several researchers [5-17] have applied the homotopy perturbation/analysis methods (HPM/HAM) and Adomian decomposition method (ADM) to solve different kinds of fractional ordinary differential equations (ODEs), fractional partial differential equations (ODEs), integral equations (IEs) and integro-differential equations (IDEs). Among them Javidi and Ahmad [18] proposed a numerical method which is based on the homotopy perturbation method and Laplace transform for fractional PDEs. In [19], LHPM which is a combination of the HPM and Laplace Transform (LT) has been employed for solving one-dimensional partial differential equations.

Recently, [20-22] introduced a new fractional derivative called conformable derivative operator (CDO) and by the help of this operator, the behaviors of many scientific problems have been solved and some solution methods have been developed. Many researchers [23-27] have studied on CDO in engineering, physical and applied mathematics problems. The aim of this study is to construct CADM and CMHPM by using conformable derivative. Many linear and nonlinear fractional PDEs can be solved with these methods. We have solved two fractional order PDEs with these mentioned methods and compared the numerical and approximate-analytical solutions in term of figures and tables. When looking at the results, it is obvious that these methods are very effective and accurate for solving fractional partial differential equations.

2. Some preliminaries

In this section, we give some basic concepts of conformable fractional derivative and its properties.

Definition 1. *Given a function* $f:[0,\infty)\to\mathbb{R}$. *Then the conformable derivative of* f *order* $\alpha\in(0,1]$ *is defined by [20]:*

$$CD_{*\alpha}(f)(t)=\lim_{\varepsilon\to 0}\frac{f\left(t+\varepsilon t^{1-\alpha}\right)-f(t)}{\varepsilon} \qquad (1)$$

*Corresponding author

for all $t > 0$.

Theorem 1. [20] *Let* $\alpha \in (0,1]$ *and* f, g *be* $\alpha-$ *differentiable at a point* $t > 0$. *Then;*

(i) $CD_{*\alpha}(af + bg) = aCD_{*\alpha}(f) + bCD_{*\alpha}(g)$ *for all* $a, b \in \mathbb{R}$,

(ii) $CD_{*\alpha}(t^k) = kt^{k-\alpha}$ *for all* $k \in \mathbb{R}$,

(iii) $CD_{*\alpha}(f(t)) = 0$ *for all constant functions* $f(t) = k$,

(iv) $CD_{*\alpha}(fg) = fCD_{*\alpha}(g) + gCD_{*\alpha}(f)$,

(v) $CD_{*\alpha}(f/g) = \dfrac{gCD_{*\alpha}(f) - fCD_{*\alpha}(g)}{g^2}$,

(vi) *If* $f(t)$ *is differentiable, then* $CD_{*\alpha}(f(t)) = t^{1-\alpha}\dfrac{d}{dt}f(t)$.

Definition 2. [20, 27] *Let* f *be an* $n-$*times differentiable at* t. *Then the conformable derivative of* f *order* α *is defined as:*

$$CD_{*\alpha}(f(t)) = \lim_{\varepsilon \to 0} \frac{f^{(\lceil\alpha\rceil-1)}\left(t + \varepsilon t^{(\lceil\alpha\rceil-\alpha)}\right) - f^{(\lceil\alpha\rceil-1)}(t)}{\varepsilon}$$

for all $t > 0$, $\alpha \in (n, n+1]$. *Here* $\lceil\alpha\rceil$ *is the smallest integer greater than or equal to* α.

Lemma 1. [20, 27] *Let* f *be an* $n-$*times differentiable at* t. *Then*

$$CD_{*\alpha}(f(t)) = t^{\lceil\alpha\rceil-\alpha} f^{\lceil\alpha\rceil}(t)$$

for all $t > 0$, $\alpha \in (n, n+1]$.

3. Conformable fractional adomian decomposition method

Consider the following nonlinear fractional partial differential equation:

$$L_{*\alpha}(u(x,t)) + R(u(x,t)) + N(u(x,t)) = v(x,t) \tag{2}$$

where $L_{*\alpha} = CD_{*\alpha}$ is a linear operator with conformable derivative of order α $(n < \alpha \le n+1)$, R is the other part of the linear operator, N is a non-linear operator and $v(x,t)$ is a non-homogeneous term. In Eq. (2), if we apply the linear operator to Lemma 1, we obtain the following equation [28]:

$$t^{\lceil\alpha\rceil-\alpha}\frac{\partial^{\lceil\alpha\rceil}u(x,t)}{\partial t^{\lceil\alpha\rceil}} + R(u(x,t)) + N(u(x,t)) = v(x,t).$$

Applying the inverse of linear operator $L_{*\alpha}^{-1} = \int_0^t \int_0^{\gamma_1} \cdots \int_n^{\gamma_{n-1}} \frac{1}{\gamma_n^{\lceil\alpha\rceil-\alpha}}(.)\,d\gamma_n d\gamma_{n-1}\cdots d\gamma_1$, to both sides of Eq. (2), we obtain

$$L_{*\alpha}^{-1}L_{*\alpha}(u(x,t)) + L_{*\alpha}^{-1}R(u(x,t)) + L_{*\alpha}^{-1}N(u(x,t)) = L_{*\alpha}^{-1}v(x,t). \tag{3}$$

The conformable ADM suggests the solution $u(x,t)$ be decomposed into the infinite series of components

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t). \tag{4}$$

The nonlinear function in Eq. (2) is decomposed as follows:

$$N(u) = \sum_{n=0}^{\infty} A_n,\ (u_0, u_1, \cdots, u_n), \tag{5}$$

where A_n is the so-called Adomian polynomials. These polynomials can be calculated for all forms of nonlinearity with respect to the algorithms developed by Adomian [29].

Substituting (4) and (5) into (3), we obtain

$$\sum_{n=0}^{\infty} u_n = u(x,0) + L_{*\alpha}^{-1}v - L_{*\alpha}^{-1}R\left(\sum_{n=0}^{\infty} u_n\right) - L_{*\alpha}^{-1}\left(\sum_{n=0}^{\infty} A_n\right). \tag{6}$$

By using Eq. (6), the iteration terms are obtained by the following way:

$$\begin{aligned} u_0 &= u(x,0) + L_{*\alpha}^{-1}v, \\ u_1 &= -L_{*\alpha}^{-1}Ru_0 - L_{*\alpha}^{-1}A_0, \\ &\vdots \\ u_{n+1} &= -L_{*\alpha}^{-1}Ru_n - L_{*\alpha}^{-1}A_n,\ n \ge 0. \end{aligned} \tag{7}$$

Then, the approximate-analytical solution of Eq. (2) is obtained by

$$\tilde{u}_k(x,t) = \sum_{n=0}^{k} u_n(x,t).$$

Finally, we obtain the exact solution of Eq. (2) as

$$u(x,t) = \lim_{k \to \infty} \tilde{u}_k(x,t).$$

4. Conformable fractional modified homotopy perturbation method

In this section, some basic solution steps and properties of modified homotopy perturbation method are given in the conformable sense (CMHPM) definition. We introduce a solution algorithm in an effective way for the nonlinear PDEs of fractional order. Firstly, we consider the following nonlinear fractional equation:

$$CD_{*t}^{\alpha}u(x,t) = L(u, u_x, u_{xx}) + N(u, u_x, u_{xx}) + v(x,t), \tag{8}$$

where $t > 0$, L is a linear operator, N is a nonlinear operator, v is a known analytical function and CD_{*t}^{α}, $m-1 < \alpha \le m$, is the Conformable fractional derivative of order α, subject to the initial conditions

$$u^k(x,0) = v_k(x),\ k = 0, 1, \ldots, m-1.$$

According to the homotopy technique, we can construct the following homotopy:

$$\frac{\partial^m u}{\partial t^m} - L(u,u_x,u_{xx}) - v(x,t) = p\left(\frac{\partial^m u}{\partial t^m} + N(u,u_x,u_{xx}) - CD_{*t}^{\alpha}u\right), \tag{9}$$

or evenly,

$$\frac{\partial^m u}{\partial t^m} - v(x,t) = p\left(\frac{\partial^m u}{\partial t^m} + L(u,u_x,u_{xx}) + N(u,u_x,u_{xx}) - CD_{*t}^{\alpha}u\right), \tag{10}$$

where $p\in[0,1]$. Here, the homotopy parameter p always changes from zero to unity. In case $p=0$, Eq. (9) becomes the linearized equation

$$\frac{\partial^m u}{\partial t^m} = L(u,u_x,u_{xx}) + v(x,t),$$

and Eq. (10) becomes the linearized equation

$$\frac{\partial^m u}{\partial t^m} = v(x,t).$$

If we take the homotopy parameter $p=1$, Eq. (9) or Eq. (10) turns out to be the original differential equation of fractional order (8). As the basic assumption is that the solution of Eq. (10) can be written by using a power series in p:

$$u = u_0 + pu_1 + p^2u_2 + p^3u_3 + \cdots.$$

At the end of the solution steps, we approximate the solution as:

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t).$$

5. Numerical examples

In this section of the study, we show the effectiveness and appropriateness of the CADM and CMHPM by applying them to two different problems.

Example 1. We consider the linear time-fractional initial boundary value problem [30]

$$\frac{\partial^\alpha u}{\partial t^\alpha} = \frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial x} + u, \quad t>0, \quad x\in R, \quad 0<\alpha\le 1, \tag{11}$$

with the initial condition

$$u(x,0) = x \tag{12}$$

and the boundary conditions

$$u_x(x,0) = 1, \quad u(0,t) = 0. \tag{13}$$

Firstly, we will solve this problem by using the proposed conformable Adomian decomposition method of fractional order. Let $L_{*\alpha} = CD_{*\alpha} = \frac{\partial^\alpha}{\partial t^\alpha}$ be a linear operator, then if we apply the operator to Eq. (11) we have

$$CD_{*\alpha}u(x,t) = \frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial x} + u. \tag{14}$$

By using the Lemma 1, we can write the Eq. (14) as

$$t^{1-\alpha}\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial x} + u. \tag{15}$$

Now, we apply the inverse of operator $L_{*\alpha}$ which is $L_{*\alpha}^{-1} = \int_0^t \frac{1}{\zeta^{1-\alpha}}(.)\,d\zeta$ to both sides of Eq. (15), we get

$$u(x,t) = u(x,0) + L_{*\alpha}^{-1}\left(\frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial x} + u\right).$$

According to the iteration terms (7) and the initial condition (12), we can write the iterations and the decomposition series terms as:

$$\begin{aligned}
u_0 &= u(x,0) = x,\\
u_1 &= L_{*\alpha}^{-1}\left(\frac{\partial^2 u_0}{\partial x^2} + x\frac{\partial u_0}{\partial x} + u_0\right) = 2x\frac{t^\alpha}{\alpha},\\
u_2 &= L_{*\alpha}^{-1}\left(\frac{\partial^2 u_1}{\partial x^2} + x\frac{\partial u_1}{\partial x} + u_1\right) = 4x\frac{t^{2\alpha}}{2!\alpha^2},\\
u_3 &= L_{*\alpha}^{-1}\left(\frac{\partial^2 u_2}{\partial x^2} + x\frac{\partial u_2}{\partial x} + u_2\right) = 8x\frac{t^{3\alpha}}{3!\alpha^3},\\
&\vdots\\
u_n &= L_{*\alpha}^{-1}\left(\frac{1}{2}x^2\frac{\partial^2 u_{n-1}}{\partial x^2}\right) = 2^n x\frac{t^{n\alpha}}{n!\alpha^n}.
\end{aligned} \tag{16}$$

So, by using the decomposition series in Eq. (16), the approximate solution of Eq. (11) obtained by Adomian decomposition method in conformable sense is

$$\tilde{u}_k(x,t) = \sum_{n=0}^{k} u_n(x,t) = \sum_{n=0}^{k} 2^n x\frac{t^{n\alpha}}{n!\alpha^n}.$$

From the last equation we obtain the approximate analytical solution of the problem as

$$u(x,t) = \lim_{k\to\infty}\tilde{u}_k(x,t) = xe^{\frac{(2t)^\alpha}{\alpha}}.$$

Then the exact solution of the Eq. (11) subject to the initial condition (12) and the boundary conditions (13) for special case of $\alpha=1$, is obtained as

$$u(x,t) = xe^{2t}.$$

Secondly, we solve the Eq. (11) by using the modified homotopy perturbation method in conformable sense. If we consider the initial condition (12) and according to the homotopy (9), we can obtain the following set of linear partial differential equations:

$$\begin{aligned}
&\frac{\partial u_0}{\partial t} = 0, \quad u_0(x,0) = x,\\
&\frac{\partial u_1}{\partial t} = \frac{\partial u_0}{\partial t} + \frac{\partial^2 u_0}{\partial x^2} + x\frac{\partial u_0}{\partial x} + u_0 - CD_{*\alpha}u_0, \quad u_1(x,0) = 0,\\
&\frac{\partial u_2}{\partial t} = \frac{\partial u_1}{\partial t} + \frac{\partial^2 u_1}{\partial x^2} + x\frac{\partial u_1}{\partial x} + u_1 - CD_{*\alpha}u_1, \quad u_2(x,0) = 0,\\
&\vdots
\end{aligned} \tag{17}$$

By solving the Eq. (17) according to u_0, u_1, u_2 and u_3, the first several components of the modified homotopy perturbation solution for Eq. (11) are derived as follows:

$$u_0(x,t)=x,$$

$$u_1(x,t)=2xt,$$

$$u_2(x,t)=x\left(2t+2t^2-\frac{2t^{2-\alpha}}{2-\alpha}\right),$$

$$u_3(x,t)=x\left(2t+4t^2+\frac{4t^3}{3}-\frac{4t^{2-\alpha}}{2-\alpha}\right)$$
$$-x\left(\frac{4t^{3-\alpha}}{(2-\alpha)(3-\alpha)}+\frac{4t^{3-\alpha}}{3-\alpha}-\frac{2t^{3-2\alpha}}{3-2\alpha}\right),$$

$\vdots$

and so on, in this way the rest of components of the homotopy can be obtained. Then the approximate solution of Eq. (11) is given by

$$u(x,t)=u_0(x,t)+u_1(x,t)+u_2(x,t)+u_3(x,t)+\cdots$$
$$=x\left(1+6t+6t^2+\frac{4t^3}{3}-\frac{6t^{2-\alpha}}{2-\alpha}-\frac{4t^{3-\alpha}}{(2-\alpha)(3-\alpha)}\right).$$
$$+x\left(-\frac{4t^{3-\alpha}}{3-\alpha}+\frac{2t^{3-2\alpha}}{3-2\alpha}+\cdots\right)$$

Then the exact solution of the Eq. (11) subject to the initial condition (12) and the boundary conditions (13) for special case of $\alpha=1,$ is obtained with CMHPM as

$$u(x,t)=xe^{2t}.$$

The following Figure 1 shows CMHPM, CADM and exact solutions for various values of α. According to the Figure 1, it can be say that the numerical results found are very close to the exact solution results.

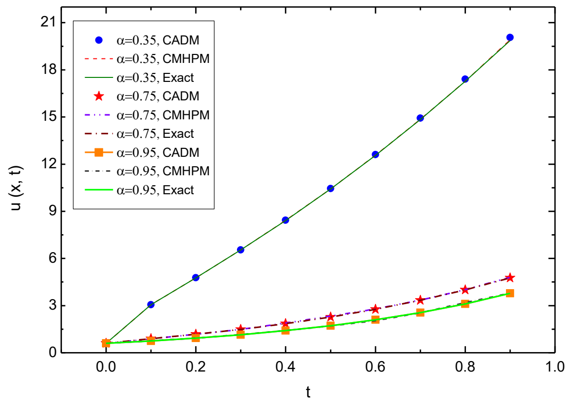


Figure 1. Comparison the numerical solutions and the exact solutions at $x=0.6$ for various values of α.

Example 2. Now let us consider the following time-fractional diffusion equation [31]

$$\frac{\partial^\alpha u}{\partial t^\alpha}=\frac{\partial^2 u}{\partial x^2},\quad t>0,\quad x\in R,\quad 0<\alpha\le 1, \tag{18}$$

with the initial condition

$$u(x,0)=\sin x. \tag{19}$$

subject to the boundary conditions

$$u_x(x,0)=\cos x,\quad u(0,t)=0. \tag{20}$$

Solve the problem by using CADM. Let us apply the linear operator to Eq. (18), then we obtain

$$CD_{*\alpha}u(x,t)=\frac{\partial^2 u(x,t)}{\partial x^2},\quad t>0,\quad x\in R,\quad 0<\alpha\le 1, \tag{21}$$

Also, we can write the Eq. (21) as

$$t^{1-\alpha}\frac{\partial u(x,t)}{\partial t}=\frac{\partial^2 u(x,t)}{\partial x^2},\quad t>0,\quad x\in R,\quad 0<\alpha\le 1, \tag{22}$$

Applying the inverse of operator $L_{*\alpha}$ to both sides of Eq. (22), we have

$$u(x,t)=u(x,0)+L_{*\alpha}^{-1}\left(\frac{\partial^2 u(x,t)}{\partial x^2}\right).$$

Using Eq. (7) and the initial condition (19), we can obtain the iterations in conformable sense as:

$$\begin{aligned}
u_0&=u(x,0)=\sin x,\\
u_1&=L_{*\alpha}^{-1}\left(\frac{\partial^2 u_0}{\partial x^2}\right)=-\sin x\frac{t^\alpha}{\alpha},\\
u_2&=L_{*\alpha}^{-1}\left(\frac{\partial^2 u_1}{\partial x^2}\right)=\sin x\frac{t^{2\alpha}}{2!\alpha^2},\\
u_3&=L_{*\alpha}^{-1}\left(\frac{\partial^2 u_2}{\partial x^2}\right)=-\sin x\frac{t^{3\alpha}}{3!\alpha^3},\\
&\vdots\\
u_n&=L_{*\alpha}^{-1}\left(\frac{\partial^2 u_{n-1}}{\partial x^2}\right)=\sin x(-1)^n\frac{t^{n\alpha}}{n!\alpha^n}.
\end{aligned} \tag{23}$$

Then, by using the obtained values in Eq. (23) the approximate solution of Eq. (18) is obtained as

$$\tilde{u}_k(x,t)=\sum_{n=0}^{k}u_n(x,t)=\sum_{n=0}^{k}\sin x(-1)^n\frac{t^{n\alpha}}{n!\alpha^n}.$$

Using the last equation we obtain the approximate analytical solution of the proposed problem

$$u(x,t)=\lim_{k\to\infty}\tilde{u}_k(x,t)=\sin x e^{-\frac{t^\alpha}{\alpha}}.$$

The exact solution of the Eq. (18) with the initial condition (19) for special case of $\alpha=1,$ is found as

$$u(x,t)=\sin x e^{-t}$$

which is the same solution with [31]. Now, let us consider the solution of problem (18) with CMHPM. In order to obtain the solution, we use the homotopy and following set of linear partial differential equations:

$$\begin{aligned}&\frac{\partial u_0}{\partial t}=0,\ u_0(x,0)=\sin x,\\&\frac{\partial u_1}{\partial t}=\frac{\partial u_0}{\partial t}+\frac{\partial^2 u_0}{\partial x^2}-CD_{*\alpha}u_0,\ u_1(x,0)=0,\\&\frac{\partial u_2}{\partial t}=\frac{\partial u_1}{\partial t}+\frac{\partial^2 u_1}{\partial x^2}-CD_{*\alpha}u_1,\ u_2(x,0)=0,\\&\vdots\end{aligned}\tag{24}$$

By solving Eq. (24) according to u_0, u_1 and u_2, the first three components of the modified homotopy perturbation solution for Eq. (18) are obtained as follows:

$$\begin{aligned}u_0(x,t)&=\sin x,\\u_1(x,t)&=-t\sin x,\\u_2(x,t)&=\sin x\left(-t+\frac{t^2}{2}+\frac{t^{2-\alpha}}{2-\alpha}\right),\\u_3(x,t)&=\sin x\left(-t+t^2-\frac{t^3}{6}+\frac{2t^{2-\alpha}}{2-\alpha}\right),\\&\quad-\sin x\left(\frac{t^{3-\alpha}}{(2-\alpha)(3-\alpha)}+\frac{t^{3-\alpha}}{3-\alpha}+\frac{t^{3-2\alpha}}{3-2\alpha}\right)\\&\vdots\end{aligned}$$

and so on, in this manner the rest of components of the homotopy can be obtained. The approximate solution of problem (18) is given by

$$\begin{aligned}u(x,t)&=u_0(x,t)+u_1(x,t)+u_2(x,t)+\cdots\\&=\sin x\left(1-3t+\frac{3t^2}{2}-\frac{t^3}{6}+\frac{3t^{2-\alpha}}{2-\alpha}\right)\\&\quad-\sin x\left(\frac{t^{3-\alpha}}{(2-\alpha)(3-\alpha)}+\frac{t^{3-\alpha}}{3-\alpha}+\frac{t^{3-2\alpha}}{3-2\alpha}\right)\end{aligned}$$

Then, for the special value of $\alpha=1$, the exact solution of the Eq. (18) subject to the initial condition (19) is obtained with CMHPM as $u(x,t)=\sin xe^{-t}$ which is the same solution with obtained CADM one.

Table 1. Absolute errors $|\tilde{u}_k(x,t)-u(x,t)|$ obtained with CADM for Example 2.

x	α	t			
		0.1	**0.3**	**0.5**	**0.7**
0.1	$\alpha=0.20$	6.08E-04	1.80E-03	3.45E-04	4.80E-02
	$\alpha=0.45$	5.24E-03	1.80E-02	5.06E-02	3.56E-02
	$\alpha=0.80$	6.49E-01	1.80E-02	1.28E-01	5.69E-01
0.4	$\alpha=0.20$	6.82E-04	9.03E-04	3.33E-05	3.15E-04
	$\alpha=0.45$	5.42E-04	6.07E-03	8.43E-05	8.62E-04
	$\alpha=0.80$	3.96E-02	5.80E-03	1.45E-02	5.69E-03
0.7	$\alpha=0.20$	5.39E-03	6.03E-04	1.75E-05	9.03E-05
	$\alpha=0.45$	4.44E-02	5.24E-02	3.08E-02	5.56E-03
	$\alpha=0.80$	8.43E-02	3.94E-01	7.80E-02	3.78E-02
1.0	$\alpha=0.20$	4.32E-05	4.80E-03	3.35E-04	3.07E-03
	$\alpha=0.45$	3.74E-04	6.92E-02	5.42E-02	9.10E-02
	$\alpha=0.80$	6.20E-02	3.42E-01	5.06E-02	6.05E-02

According to Table 1, we can say about the solution of Eq. (18) that the absolute error values are very small for various values α, x and t.

In addition, in the following Figure 2 and Figure 3, the graphs of solution functions of Eq. (18) with respect to the CADM and the exact solution for $\alpha=0.70$ are shown, respectively.

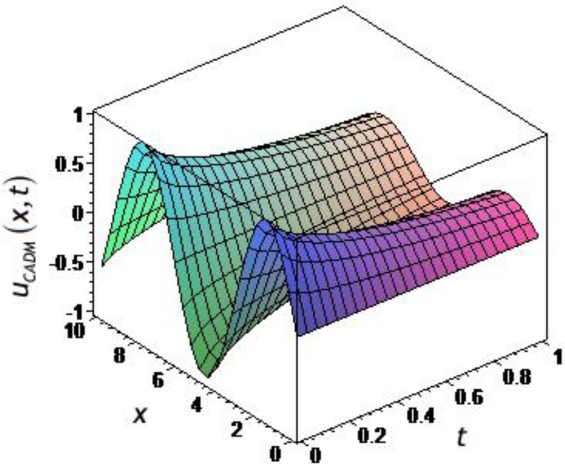


Figure 2. CADM solution with $\alpha=0.70$ for Example 2.

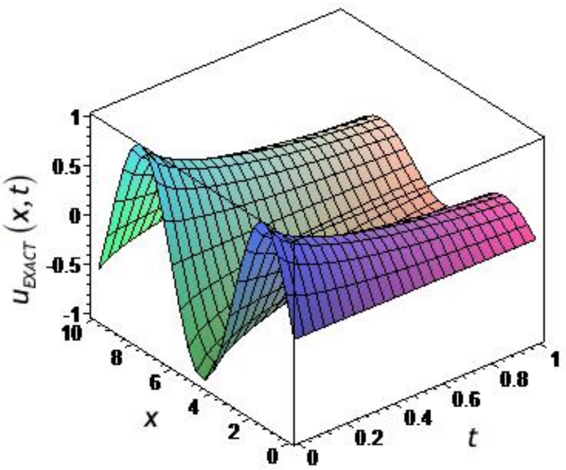


Figure 3. Exact solution with $\alpha=0.70$ for Example 2.

In the following Figure 4 and Figure5, the sketches of solution functions of Eq. (18) with respect to the CMHPM and the exact solution for $\alpha=0.30$ are shown, respectively.

According to the Figure 2, Figure 3, Figure 4 and Figure 5, we can say that the numerical results obtained from CADM and CMHPM are very close to the exact solution values.

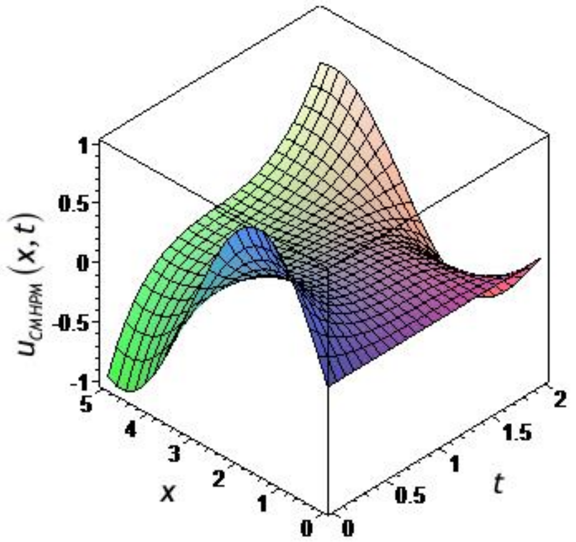


Figure 4. CMHPM solution with $\alpha = 0.30$ for Example 2.

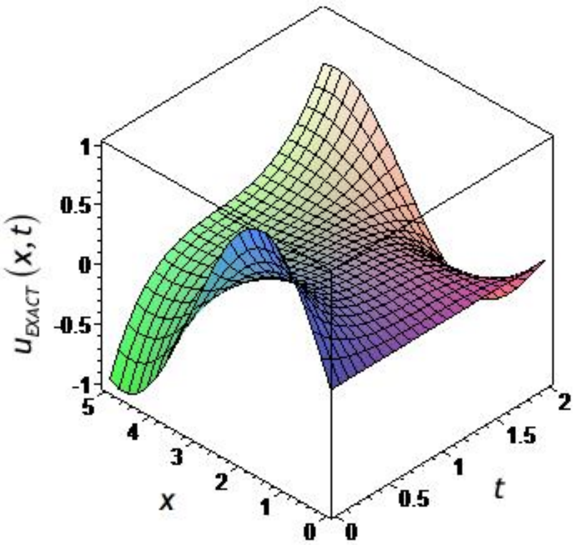


Figure 5. Exact solution with $\alpha = 0.30$ for Example 2.

6. Conclusion

We have found out approximate solutions with two numerical methods for time-fractional linear partial differential equations. These methods are based on conformable derivative (CD) which is extremely popular in the last years. In this study, firstly, by using the CD, we have redefined ADM and MHPM. Then we have demonstrated the efficiencies and accuracies of the proposed methods by applying them to two different problems. It is found that the approximate solutions generated by our methods are in complete agreement with the corresponding exact solutions. Besides, in view of their usability, our methods are applicable to many initial-boundary value problems and linear-nonlinear partial differential equations of fractional order.

References

[1] Zhang, Y. (2009). A finite difference method for fractional partial differential equation. *Applied Mathematics and Computation*, 215(2), 524-529.

[2] Ibrahim, R.W. (2011). On holomorphic solutions for nonlinear singular fractional differential equations. *Computers & Mathematics with Applications*, 62(3), 1084-1090.

[3] Odibat, Z. and Momani, S. (2008). A generalized differential transform method for linear partial differential equations of fractional order. *Applied Mathematics Letters*, 21(2), 194-199.

[4] Odibat, Z. and Momani, S. (2008). Numerical methods for nonlinear partial differential equations of fractional order. *Applied Mathematical Modelling*, 32(1), 28-39.

[5] Bildik, N. and Bayramoglu, H. (2005). The solution of two dimensional nonlinear differential equation by the Adomian decomposition method. *Applied mathematics and computation*, 163(2), 519-524.

[6] Bildik, N., Konuralp, A., Bek, F.O., & Küçükarslan, S. (2006). Solution of different type of the partial differential equation by differential transform method and Adomian's decomposition method. *Applied Mathematics and Computation*, 172(1), 551-567.

[7] Daftardar-Gejji, V. and Jafari, H. (2005). Adomian decomposition: A tool for solving a system of fractional differential equations. *Journal of Mathematical Analysis and Applications*, 301(2), 508-518.

[8] Elbeleze, A.A., Kılıçman, A., & Taib, B.M. (2013). Homotopy perturbation method for fractional Black-Scholes European option pricing equations using sumudu transform. *Mathematical Problems in Engineering*, 2013.

[9] El-Sayed, A. and Gaber, M. (2006). The adomian decomposition method for solving partial differential equations of fractal order in finite domains. *Physics Letters A*, 359(3), 175-182.

[10] El-Wakil, S.A., Abdou, M.A., & Elhanbaly, A. (2006). Adomian decomposition method for solving the diffusion–convection–reaction equations. *Applied Mathematics and Computation*, 177(2), 729-736.

[11] Gülkaç, V. (2010). The homotopy perturbation method for the Black–Scholes equation. *Journal of Statistical Computation and Simulation*, 80(12), 1349-1354.

[12] Momani, S. and Odibat, Z. (2007). Numerical approach to differential equations of fractional order. *Journal of Computational and Applied Mathematics*, 207(1), 96-110.

[13] Momani, S. and Odibat, Z. (2007). Homotopy perturbation method for nonlinear partial differential equations of fractional order. *Physics Letters A*, 365(5), 345-350.

[14] Evirgen, F. and Özdemir, N. (2012). A fractional order dynamical trajectory approach for optimization problem with HPM. *In*: D. Baleanu, Machado, J.A.T., Luo, A.C.J., eds. *Fractional Dynamics and Control*, Springer, 145-155.

[15] Yavuz, M., Ozdemir, N., & Okur, Y.Y. (2016). Generalized differential transform method for fractional partial differential equation from finance, Proceedings, International Conference on Fractional Differentiation and its Applications, Novi Sad, Serbia, 778-785.

[16] Kurulay, M., Secer, A., & Akinlar, M.A. (2013). A new approximate analytical solution of Kuramoto-Sivashinsky equation using homotopy analysis method. *Applied Mathematics & Information Sciences*, 7(1), 267-271.

[17] Turut, V. and Güzel, N. (2013). Multivariate pade approximation for solving nonlinear partial differential equations of fractional order. *Abstract and Applied Analysis*, 2013.

[18] Javidi, M. and Ahmad, B. (2013). Numerical solution of fractional partial differential equations by numerical Laplace inversion technique. *Advances in Difference Equations*, 2013(1), 375.

[19] Madani, M., Fathizadeh, M., Khan, Y., & Yildirim, A. (2011). On the coupling of the homotopy perturbation method and Laplace transformation. *Mathematical and Computer Modelling*, 53(9), 1937-1945.

[20] Khalil, R., Al Horani, M., Yousef, A., & Sababheh, M. (2014). A new definition of fractional derivative. *Journal of Computational and Applied Mathematics*, 264, 65-70.

[21] Anderson, D. and Ulness, D. (2015). Newly defined conformable derivatives. *Adv. Dyn. Syst. Appl*, 10(2), 109-137.

[22] Atangana, A., Baleanu, D., & Alsaedi, A. (2015). New properties of conformable derivative. *Open Mathematics*, 13(1), 889-898.

[23] Çenesiz, Y., Baleanu, D., Kurt, A., & Tasbozan, O. (2017). New exact solutions of Burgers' type equations with conformable derivative. *Waves in Random and Complex Media*, 27(1), 103-116.

[24] Avcı, D., Eroglu, B.I., & Ozdemir, N. (2016). Conformable heat problem in a cylinder, Proceedings, International Conference on Fractional Differentiation and its Applications, 572-581.

[25] Avcı, D., Eroğlu, B.B.İ., & Özdemir, N. (2017). Conformable fractional wave-like equation on a radial symmetric plate. *In*: A. Babiarz, Czornik, A., Klamka, J., Niezabitowski, M., eds. *Theory and Applications of Non-Integer Order Systems*, Springer, 137-146.

[26] Avci, D., Eroglu, B.B.I., & Ozdemir, N. (2017). Conformable heat equation on a radial symmetric plate. *Thermal Science*, 21(2), 819-826.

[27] Abdeljawad, T. (2015). On conformable fractional calculus. *Journal of computational and Applied Mathematics*, 279, 57-66.

[28] Acan, O. and Baleanu, D. (2017). A new numerical technique for solving fractional partial differential equations. *arXiv preprint arXiv:1704.02575*,

[29] Adomian, G. (1988). A review of the decomposition method in applied mathematics. *Journal of Mathematical Analysis and Applications*, 135(2), 501-544.

[30] Demir, A., Erman, S., Özgür, B., & Korkmaz, E. (2013). Analysis of fractional partial differential equations by Taylor series expansion. *Boundary Value Problems*, 2013(1), 68.

[31] Momani, S. and Odibat, Z. (2007). Comparison between the homotopy perturbation method and the variational iteration method for linear fractional partial differential equations. *Computers & Mathematics with Applications*, 54(7), 910-919.

Mehmet Yavuz *is a lecturer at the Department of Mathematics-Computer Sciences, Faculty of Science, Necmettin Erbakan University, Turkey. He received his B.Sc. (2009) degree from Department of Mathematics, Bulent Ecevit University, Turkey. He received his M.Sc. (2012) and Ph.D. (2016) degrees from Department of Mathematics, Balikesir University, Turkey. His research areas include fractional calculus, differential equations and financial mathematics.*

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.1, pp.8-16 (2018)
https://doi.org/10.11121/ijocta.01.2018.00461

RESEARCH ARTICLE

A hybrid approach for the regularized long wave-Burgers equation

Asuman Zeytinoglu [a], Murat Sari [b]*, Bilender P. Allahverdiev [a]

[a] *Department of Mathematics, Suleyman Demirel University, Turkey*
[b] *Department of Mathematics, Yildiz Technical University, Turkey*
asumanzeytinoglu@sdu.edu.tr, sarim@yildiz.edu.tr, bilenderpasaoglu@sdu.edu.tr

ARTICLE INFO

Article history:
Received: 13 March 2017
Accepted: 11 August 2017
Available Online: 10 October 2017

Keywords:
Regularized long wave-Burgers equation
High order finite difference scheme
Weighted essentially non-oscillatory scheme
Hybrid approximation
MacCormack method

AMS Classification 2010:
65M06, 35Q35, 35G31

ABSTRACT

In this paper, a new hybrid approach based on sixth-order finite difference and seventh-order weighted essentially non-oscillatory finite difference scheme is proposed to capture numerical simulation of the regularized long wave-Burgers equation which represents a balance relation among dissipation, dispersion and nonlinearity. The corresponding approach is implemented to the spatial derivatives and then MacCormack method is used for the resulting system. Some test problems discussed by different researchers are considered to apply the suggested method. The produced results are compared with some earlier studies, and to validate the accuracy and efficiency of the method, some error norms are computed. The obtained solutions are in good agreement with the literature. Furthermore, the accuracy of the method is higher than some previous works when some error norms are taken into consideration.

1. Introduction

In describing many models in a great deal of fields of science, nonlinear partial differential equations (PDEs) play a significant role. Hence, reaching exact or well approximate solutions of nonlinear PDEs is still important. These kinds of partial differential equations may not have an exact solution by reason of their nonlinearity. So, it is of interest to introduce a new method or develop an existing technique to obtain accurate numerical results. One of the popular nonlinear partial differential equations studied for its numerical solutions is the regularized long wave-Burgers (RLW-Burgers) equation also known as Benjamin-Bona-Mahony-Burgers (BBMB) equation. This equation describes the propagation of surface water waves in a channel [1]. The RLW-Burgers equation is considered as follows with physical boundary conditions $u \to 0$ as $x \to \pm\infty$:

$$\begin{aligned} &u_t - u_{xxt} - \alpha u_{xx} + \beta u_x + g(u)_x = 0, -\infty < x < \infty,\ t \geq 0 \\ &u(x,0) = \phi(x) \to 0, \quad x \to \mp\infty . \end{aligned} \tag{1}$$

The subscripts t and x are time and space derivatives, and denote the horizontal coordinate along the channel and the elapsed time, respectively. $\phi(x)$ is a known function as initial condition, α is a positive constant, $\beta \in \mathbb{R}$ and $g(u)$ is a C^2 smooth nonlinear function. Eq. (1) represents a balance relation among dissipation, dispersion and nonlinearity [2]. Due to the fact that this equation is important for understanding the nonlinear wave phenomena, many researchers [2-9] have studied on it for many years.

Since the numerical methods are good means of understanding these types of equations, the effort of finding a more accurate numerical approach is still in progress. Investigating an effective and accurate numerical method encourages us to produce a new hybrid approach based on some high order finite difference (FD) schemes for analyzing the behavior of the RLW-Burgers equation. One of these FD schemes is a seventh-order weighted essentially non-oscillatory (WENO7) [10, 11, 16] method. It can be clear from the literature that the WENO method based on ENO schemes is one of the popular numerical methods for PDEs in conservative form $u_t + f(u)_x = 0$. High order accuracy can be achieved in the smooth regions and discontinuities can be computed without spurious oscillations [12]. Some studies in recent years have introduced several versions of the WENO scheme derived for improving ENO properties [10-16]. However, some researchers have combined the WENO schemes with a high order method to overcome some drawbacks [17-19]. Inspired by these drawbacks in the corresponding studies, we prefer to

*Corresponding author

combine the corresponding WENO scheme with the sixth-order finite difference (FD6) scheme [20, 21] because the FD6 gives convergent approximations as well as being effective, reliable and easy to implement. To validate the accuracy and efficiency of the proposed method, some error norms are presented and the obtained results are compared with the previous works in the literature.

The arrangement of this paper is as follows: The suggested scheme in both space and time are introduced in Section 2. Five test problems including different α, β parameters and $g(u)$, $\phi(x)$ functions are solved to show the efficiency and accuracy of the proposed method, and the computed results are compared with others selected from the available literature in Section 3. Finally, the last section includes the summary of findings in the paper.

2. Construction of the method

One can rewrite problem (1) with the following form:

$$u_t = v, \tag{2}$$

$$v_{xx} - v = f(u)_x - \alpha u_{xx}\ ,\ f(u)_x = \left(\beta u + g(u)\right)_x . \tag{3}$$

As α, β and $g(u)$ change, Eq. (3) changes for each test problem. It can also be seen from the above system, there is no time derivative term in Eq. (3). The proposed approach is involved the FD6 and WENO7 finite difference formulations to the spatial derivatives, and the MacCormack discretization is taken into account for the time derivative. Details of the implementation of the present method are introduced in the following subsections.

2.1. Space discretization with the hybrid scheme

First of all, we divide the domain of problem $[a,b]$ into N subintervals such as $a = x_1 < x_2 < ... < x_N < x_{N+1} = b$ with the spatial step size $h = \Delta x = x_{i+1} - x_i$ for $i = 1, 2..., N$. Also, $(n+1)$-th time level is defined by $t^{n+1} = t^n + \Delta t$ where t^n is the initial time for $n = 0$. Thus, the numerical solution of u is represented by u_i^n at grid point (x_i, t^n). We use the FD6 scheme derived for the second order derivatives to discretize the terms v_{xx} and u_{xx} in Eq. (3). The FD6 scheme can briefly be introduced as follows:

v' and v'' in space, can be approximated by the following FD6 formulae used 7-point stencil

$$v_i' = \frac{1}{h}\sum_{j=-L}^{R} a_{j+L} v_{i+j}\ ,\quad v_i'' = \frac{1}{h^2}\sum_{j=-L}^{R} \tilde{a}_{j+L} v_{i+j} \tag{4}$$

for $1 \le i \le N+1$.

In Equations (4), $(N+1)$ denotes the number of grid points, a_k and $\tilde{a}_k$ $(k = 0,...,R+L)$ are unknown constants, *R* and *L* denote the number of grid points in the right and left hand side for the taken stencil, respectively. At internal points, *R* and *L* is equal while they are different for the boundary nodes. The coefficients a_k and $\tilde{a}_k$ can be determined with Taylor series expansions about the related point and they are given in Table 1.

Table 1. The coefficients a_k and $\tilde{a}_k$

i	Coefficients*	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
1	a_k	-147	360	-450	400	-225	72	-10
	$\tilde{a}_k$	812	-3132	5265	-5080	2970	-972	137
2	a_k	-10	-77	150	-100	50	-15	2
	$\tilde{a}_k$	137	-147	-255	470	-285	93	-13
3	a_k	2	-24	-35	80	-30	8	-1
	$\tilde{a}_k$	-13	228	-420	200	15	-12	2
Internal Nodes	a_k	-1	9	-45	0	45	-9	1
	$\tilde{a}_k$	2	-27	270	-490	270	-27	2
$N-1$	a_k	1	-8	30	-80	35	24	-2
	$\tilde{a}_k$	2	-12	15	200	-420	228	-13
N	a_k	-2	15	-50	100	-150	77	10
	$\tilde{a}_k$	-13	93	-285	470	-255	-147	137
$N+1$	a_k	10	-72	225	-400	450	-360	147
	$\tilde{a}_k$	137	-972	2970	-5080	5265	-3132	812

*Each given values of a_k and $\tilde{a}_k$ in the table must be divided by 60 and 180, respectively

For the term $f(u)_x$ in Eq. (3), the WENO7 scheme is implemented together with the FD6 scheme. The WENO schemes are based on ENO schemes and it was first suggested by Liu et al [22]. They provide high order accurate solutions in smooth regions and have a good convergence since they use a convex combination of all candidate stencils against the ENO schemes. In the literature, many researchers have focused on the WENO schemes in order to improve them. Taking inspiration from those studies, the present work discusses a combination of the WENO7 finite difference scheme with the FD6 scheme in computing highly accurate results. The mentioned WENO scheme is applied to internal nodes and the FD6 formulae given in above are implemented for near the boundaries. We can then introduce the WENO7 procedure with its main points herein below [10, 11, 16]:

The WENO schemes for discretization of the spatial derivatives in the following hyperbolic conservation law

$$u_t + f(u)_x = 0 \tag{5}$$

are successful in terms of the numerical approximation. A reconstruction procedure based on the local smoothness of numerical solution is used as the main point of the WENO finite difference scheme in order to produce high order accurate solutions. The term $f(u)_x$ is approximated by

$$f(u)_x\big|_{x=x_j} \approx \frac{1}{\Delta x}\left(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}\right), \tag{6}$$

where $\hat{f}_{j+\frac{1}{2}}$ represents the numerical flux. The WENO7 scheme uses 7 candidate stencils written as a set $S = \{x_{j-3},...,x_{j+3}\}$ for these numerical fluxes. It is divided into four subset as $S^m = \{x_{j-3+m},...,x_{j+m}\}$, $m = 0,1,2,3$. The numerical flux $\hat{f}_{j+\frac{1}{2}}$ is written using stencil sets S^m as

$$\hat{f}_{j+\frac{1}{2}} = \sum_{m=0}^{3} \omega_m \hat{f}^{(m)}_{j+\frac{1}{2}} \quad , \quad \hat{f}^{(m)}_{j+\frac{1}{2}} = \sum_{i=0}^{3} b_{mi} f_{j+m-i} \tag{7}$$

for $m = 0,1,2,3$. In Eq. (7), ω_m are called non-linear weights defined by

$$\omega_m = \frac{\alpha_m}{\sum_{k=0}^{3}\alpha_k} \quad , \quad \alpha_m = d_m\left(1+\left(\frac{\tau_7}{(\varepsilon+\beta_m)}\right)^q\right)$$

with $\tau_7 = |\beta_0 - \beta_3|$ and the linear weights $d_0 = 1/35,\ d_1 = 12/35,\ d_2 = 18/35,\ d_3 = 4/35$. The coefficients b_{mi} can be calculated with the approach inspired by Xie [23] using a fourth order polynomial

$$h(x) = A + B(x - x_{j+1/2}) + C(x - x_{j+1/2})^2 + D(x - x_{j+1/2})^3 + E(x - x_{j+1/2})^4$$

with the four candidate stencils above and are presented in Table 2. The coefficients required for $\hat{f}_{j-\frac{1}{2}}$ can be found using the same stencils in a similar way. In calculations, $q = 2$, and ε is used to avoid the division by zero and it is selected to be quite small, $\varepsilon = 10^{-10}$. The smoothness indicators, β_m, are given by

$$\begin{aligned}
\beta_0 = {} & f(u_{j-3})\left[547 f(u_{j-3}) - 3882 f(u_{j-2}) + 4642 f(u_{j-1}) - 1854 f(u_j)\right] + f(u_{j-2})\left[7043 f(u_{j-2}) - 17246 f(u_{j-1}) + 7042 f(u_j)\right] \\
& + f(u_{j-1})\left[11003 f(u_{j-1}) - 9402 f(u_j)\right] + f(u_j)\left[2107 f(u_j)\right], \\
\beta_1 = {} & f(u_{j-2})\left[267 f(u_{j-2}) - 1642 f(u_{j-1}) + 1602 f(u_j) - 494 f(u_{j+1})\right] + f(u_{j-1})\left[2843 f(u_{j-1}) - 5966 f(u_j) + 1922 f(u_{j+1})\right] \\
& + f(u_j)\left[3443 f(u_j) - 2522 f(u_{j+1})\right] + f(u_{j+1})\left[547 f(u_{j+1})\right], \\
\beta_2 = {} & f(u_{j-1})\left[547 f(u_{j-1}) - 2522 f(u_j) + 1922 f(u_{j+1}) - 494 f(u_{j+2})\right] + f(u_j)\left[3443 f(u_j) - 5966 f(u_{j+1}) + 1602 f(u_{j+2})\right] \\
& + f(u_{j+1})\left[2843 f(u_{j+1}) - 1642 f(u_{j+2})\right] + f(u_{j+2})\left[267 f(u_{j+2})\right], \\
\beta_3 = {} & f(u_j)\left[2107 f(u_j) - 9402 f(u_{j+1}) + 7042 f(u_{j+2}) - 1854 f(u_{j+3})\right] + f(u_{j+1})\left[11003 f(u_{j+1}) - 17246 f(u_{j+2}) + 4642 f(u_{j+3})\right] \\
& + f(u_{j+2})\left[7043 f(u_{j+2}) - 3882 f(u_{j+3})\right] + f(u_{j+3})\left[547 f(u_{j+3})\right].
\end{aligned}$$

Table 2. The coefficients b_{mi} for the WENO7 scheme

b_{mi}	$i=0$	$i=1$	$i=2$	$i=3$
$m=0$	$25/12$	$-23/12$	$13/12$	$-1/4$
$m=1$	$1/4$	$13/12$	$-5/12$	$1/12$
$m=2$	$-1/12$	$7/12$	$7/12$	$-1/12$
$m=3$	$1/12$	$-5/12$	$13/12$	$1/4$

For more details of the WENO finite difference scheme, interested readers are referred to as the literature such as [10-13, 16].

2.2. Time discretization with MacCormack method

After the implementation of the aforementioned main schemes to Eq. (3), the values of variable v are found and then the MacCormack method is used to find new values of u at the next time level from Eq. (2). This method is widely used for solving nonlinear PDEs representing fluid flows and provides accurate results [24]. Let us consider the following general form of governing equation

$$\frac{du_i}{dt} = Pu_i .$$

In this form, P represents a spatial differential operator, and each values on the right hand side of the above equation are already known through the method described in the previous subsection. In order to solve this semi-discrete equation, the MacCormack approach is then implemented via the following process:

$$\text{Pre. Step: } \overline{u_i^{n+1}} = u_i^n + \Delta t P u_i^n ,$$

$$\text{Cor. Step: } u_i^{n+1} = u_i^{n+1/2} + \frac{\Delta t}{2} P\overline{u_i^{n+1}},\ u_i^{n+1/2} = \frac{u_i^n + \overline{u_i^{n+1}}}{2} .$$

3. Numerical Illustrations

In this section, we implement the previous procedure to five test problems for producing numerical solutions of the RLW-Burgers equation. The accuracy of the numerical solutions is observed by using absolute error and the following error norms

$$L_2 = \sqrt{h\sum_{j=1}^{N}\left|u_j^{analytical} - u_j^{numerical}\right|^2} ,$$

$$L_\infty = \max\left|u_j^{analytical} - u_j^{numerical}\right| .$$

which measure the mean and maximum differences between the numerical and analytical solutions. To show the behaviors of corresponding problems, some figures are also plotted.

Example 1. As the first test problem, Eq. (1) with $\alpha = 1$, $\beta = 1$ and $g(u) = u^2/2$ is considered by the following initial condition

$$u(x,0) = \operatorname{sech}^2(x/4) .$$

Table 3 gives the obtained results using $h = 0.25$ and $\Delta t = 0.01$ in the interval $-12 \le x \le 12$. It can be clearly seen that the produced results are compatible with the results of Zarebnia and Parvaz [8]. Also, the solutions at various times are qualitatively presented in Figure 1. As is the case in the study of Zarebnia and Parvaz [8] and as naturally expected, the amplitude of wave slightly decreases as the time goes on (see Figure 1).

Table 3. Numerical results with the parameters $h = 0.25$ and $\Delta t = 0.01$ for Example 1

	Present method					
	t					
x	0.2	0.5	0.7	1	1.5	2
-10	0.024700	0.022144	0.020606	0.018518	0.015529	0.013038
-5	0.256292	0.224794	0.206475	0.182379	0.149529	0.123687
0	0.978142	0.933383	0.897619	0.838349	0.733532	0.631524
5	0.319370	0.380986	0.423415	0.487534	0.589834	0.676862
10	0.032031	0.041918	0.049799	0.063792	0.093622	0.132550
	[8]					
-10	0.022951	0.019822	0.017950	0.015435	0.011934	0.009165
-5	0.256278	0.224742	0.206391	0.182231	0.149239	0.123215
0	0.978102	0.933352	0.897596	0.83834	0.733537	0.631526
5	0.319376	0.380993	0.42342	0.487532	0.589817	0.676809
10	0.030420	0.039796	0.047263	0.060502	0.088659	0.12528

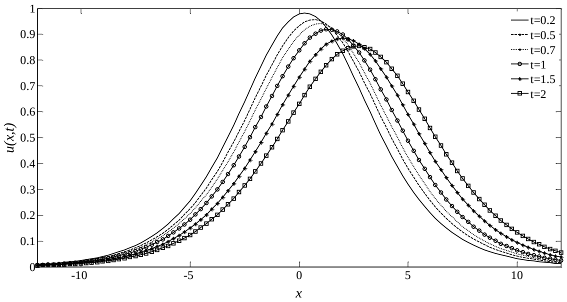


Figure 1. The behavior of the wave in Example 1.

Example 2. Consider Eq. (1) with the initial condition

$$u(x,0) = -\frac{9}{5} - \frac{6}{5}\tanh\left(\frac{x}{2}\right) + \frac{3}{5}\tanh^2\left(\frac{x}{2}\right)$$

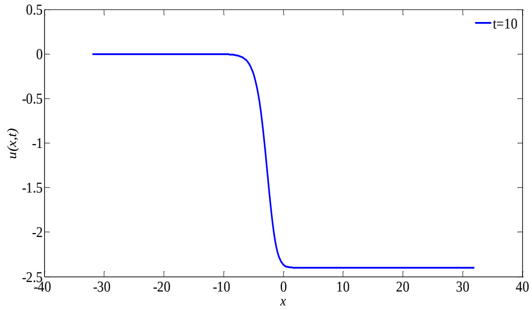
using $\alpha = 1$, $\beta = 1$ and $g(u) = u^2/2$. The analytical solution is given by

$$u(x,t) = -\frac{9}{5} - \frac{6}{5}\tanh\left(\frac{5x+t}{10}\right) + \frac{3}{5}\tanh^2\left(\frac{5x+t}{10}\right).$$

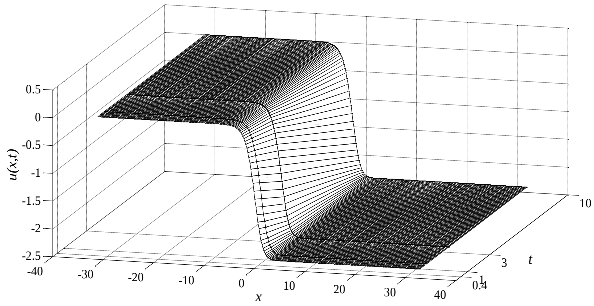
For this problem the parameters are chosen as $h = 0.2$ and $\Delta t = 0.01$ in the interval $-32 \le x \le 32$. In Table 4, the produced L_2, L_∞ errors are given. The obtained error values are quite good even for the larger time $t=10$. Furthermore, absolute errors at various points in the corresponding domain are presented and compared with the study of Alquran and Al-Khaled [4] for some time values in Table 5. It can be said that our results are at least three decimal digits better than the results of Alquran and Al-Khaled [4]. The qualitative behavior of solutions at $t=10$ and at various times are exhibited in Figure 2(a)-2(b), respectively.

Table 4. L_2 and L_∞ error norms for Example 2

t	L_2	L_∞
0.2	1.826363E-07	1.678531E-07
0.4	3.371454E-07	3.068812E-07
1	6.734666E-07	6.326276E-07
3	1.096361E-06	9.812815E-07
10	1.754436E-05	1.546722E-05



(a)



(b)

Figure 2. Numerical solutions for Example 2 using $h=0.2$ and $\Delta t=0.01$ for (a) $t=10$ and (b) various times.

Table 5. Absolute errors at various times for Example 2

	Present Method	[4]	Present Method	[4]	Present Method	Present Method	Present Method
$x \setminus t$	0.2	0.2	0.4	0.4	1	3	10
0.2	1.28E-07	6.76E-04	2.22E-07	5.02E-04	3.39E-07	7.71E-08	1.92E-07
0.4	6.42E-08	5.05E-04	1.00E-07	4.52E-04	9.16E-08	2.84E-07	2.11E-07
0.6	1.63E-09	4.82E-04	1.16E-08	6.02E-04	1.02E-07	3.27E-07	1.97E-07
0.8	4.01E-08	4.14E-04	8.19E-08	6.02E-04	1.97E-07	2.47E-07	1.65E-07
1	5.50E-08	3.21E-04	1.03E-07	2.02E-04	1.98E-07	1.12E-07	1.26E-07
1.2	4.85E-08	6.05E-05	8.56E-08	7.51E-05	1.39E-07	1.90E-08	8.94E-08
1.4	3.07E-08	5.85E-05	5.05E-08	5.98E-05	6.02E-08	1.14E-07	5.81E-08
1.8	4.73E-09	1.36E-05	1.39E-08	2.37E-05	5.65E-08	1.76E-07	1.66E-08
2.4	1.96E-08	1.25E-05	3.75E-08	1.01E-05	7.90E-08	1.08E-07	5.72E-09
3	1.20E-08	6.23E-06	2.19E-08	7.78E-06	4.08E-08	3.76E-08	7.24E-09
5	3.92E-10	4.82E-06	6.01E-10	6.42E-06	4.69E-10	2.01E-09	6.30E-10

Example 3. We now consider Eq. (1) with the parameters $\alpha=1$, $\beta=1$ and $g(u)=u^2/2$ with the initial condition

$$u(x,0)=\exp\left(-x^2\right).$$

In this example, the domain is taken to be $-30 \le x \le 30$ and the behavior of the problem is examined up to time $t=10$. We use $h=0.2$ and $\Delta t=0.1$ in the proposed scheme and the recorded values are presented in Table 6. Furthermore, the profile of the wave is plotted in Figure 3 from $t=0$ to $t=10$. It can be seen from the figure that the amplitude of wave and the position of that amplitude changes in time. The amplitude of wave is equal to 1 located at $x=0$ for initial time, while that value decreases as the times goes on and it becomes about 0.2 located close by $x=10$.

Table 6. Numerical results for Example 3 with $h=0.2$ and $\Delta t=0.1$

$x \setminus t$	1	2	5	10
-30	-5.100E-13	-3.027E-12	-6.602E-11	-1.174E-09
-15	-4.952E-08	-4.754E-08	-1.567E-08	-1.462E-09
0	5.726E-01	2.871E-01	2.032E-02	-2.175E-03
15	2.199E-05	2.396E-04	8.813E-03	9.039E-02
30	1.653E-10	5.346E-09	2.186E-06	4.348E-04

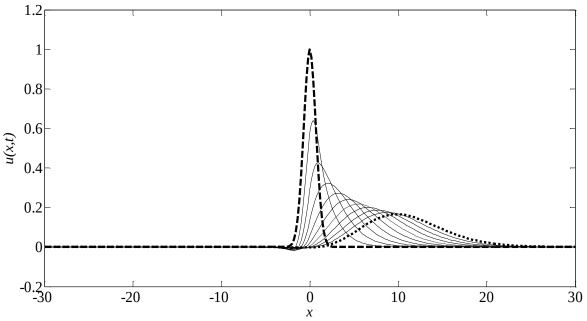


Figure 3. The behavior of the wave in Example 3 from $t=0$ to $t=10$ using $h=0.2$ and $\Delta t=0.1$.

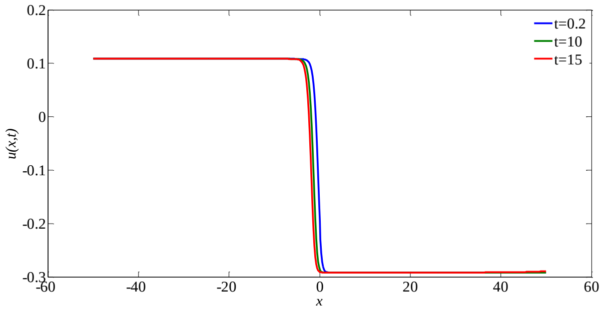


Figure 4. Numerical solutions for Example 4 at different times.

Example 4. As the fourth test problem, Eq. (1) under the consideration of parameters $\alpha=1$, $\beta=1$ and $g(u)=6u^2$ is studied using the following initial condition

$$u(x,0)=-\frac{23}{120}-\frac{1}{5}\tanh x+\frac{1}{10}\tanh^2 x$$

extracted from the exact solution

$$u(x,t)=-\frac{23}{120}-\frac{1}{5}\tanh\left(x+\frac{t}{10}\right)+\frac{1}{10}\tanh^2\left(x+\frac{t}{10}\right).$$

For comparison with an early work by Zhao et al. [2], L_∞ errors are calculated for various times over the domain $-50\le x\le 50$. The parameter h is taken to be 0.2 with both $\Delta t=0.01$ and $\Delta t=0.1$, and the results are presented up to time $t=10$ in Table 7. It is seen that the obtained L_∞ errors are less than the compared results. Furthermore, the presented errors still decrease when $\Delta t=0.01$. The behavior of the problem for three different times are given in Figure 4.

Table 7. L_∞ errors at various times for Example 4 using

t	L_∞		
	Present Method		[2]
	$h=0.2$, $\Delta t=0.01$	$h=0.2$, $\Delta t=0.1$	$h=0.2$, $\Delta t=0.1$
0.2	6.048E-06	6.553E-06	7.650E-05
0.3	9.310E-06	1.003E-05	6.954E-05
0.4	1.274E-05	1.365E-05	1.490E-04
0.5	1.632E-05	1.739E-05	1.334E-04
0.6	2.001E-05	2.123E-05	2.160E-04
0.7	2.377E-05	2.510E-05	1.918E-04
0.8	2.752E-05	2.895E-05	2.774E-04
0.9	3.119E-05	3.270E-05	2.474E-04
1	3.472E-05	3.629E-05	3.385E-04
2	5.823E-05	6.143E-05	-
3	9.216E-05	9.493E-05	-
5	1.481E-04	1.511E-04	-
10	2.871E-04	2.889E-04	-

Example 5. For the last problem, the nonlinear function $g(u)$ is chosen as $u^3/3$, $u^5/5$ and $u^9/9$, respectively, for Eq. (1) with $\alpha=1/2, \beta=1$.

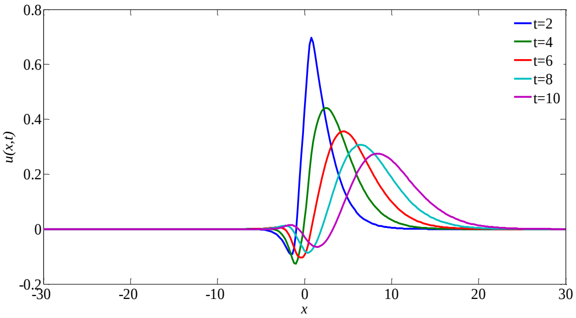
The initial condition is taken to be

$$u(x,0)=1/\left(1+x^4\right).$$

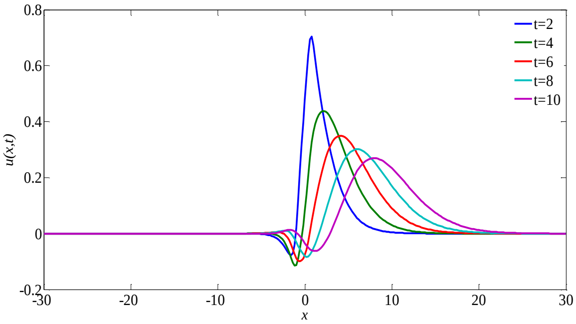
In this example, the solutions with considering various nonlinear function $g(u)$ are investigated up to time $t=10$. In Figure 5, the solutions are plotted at different times using $h=0.2$ and $\Delta t=0.1$. To show the effect of parameter h, the solutions at $t=10$ are also displayed in Figure 6 using $\Delta t=0.1$ and various h values. It is observed from the corresponding figures that due to the value of α parameter, a slight oscillation occurs at the beginning of the wave, and the amplitude of both wave and oscillation decreases as the time goes on. To see the effect of α parameter on the behavior of the wave, Figure 7 is presented for various values of α. In the calculations, the parameters are taken to be $h=0.2$, $\Delta t=0.1$, $g(u)=u^3/3$. Figure 7 shows that as the value of α decreases, the amplitude of the wave slightly increases. However, any oscillation does not appear in the wave motion if α value is taken larger. It can be also seen that as the α value decreases, some oscillations occur. The results in the above examples revealed that the proposed method has been seen to be usually more convergent and easier than its rival methods from the literature.

4. Conclusion

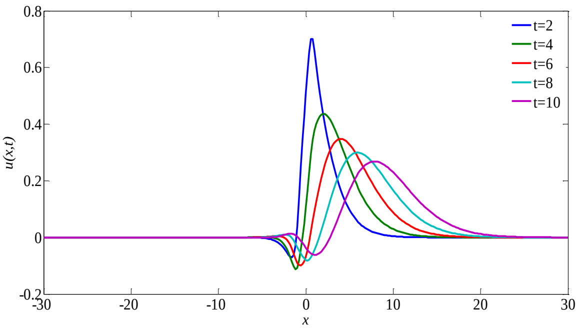
A hybrid approach based on two different types of finite difference scheme has been introduced and applied for the solutions of some physical problems constructed with the RLW-Burgers equation. To reveal the accuracy of the proposed scheme, five test problems are considered for various values of parameters taken part in the RLW-Burgers equation, and some error norms, such as absolute, L_2 and L_∞, are presented. The computed results revealed that the suggested method highly accurate, computationally powerful and user-friendly. The present approach is also believed to be easier in producing computer codes for applications. Therefore, it is seen to be a strongly advisable alternative to discover both qualitative and quantitative behaviors of similar processes for further studies.



(a) $g(u)=u^3/3$

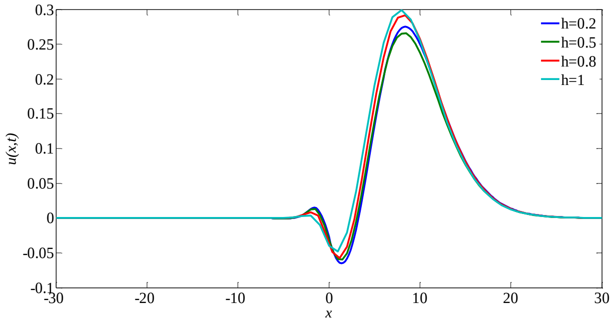


(b) $g(u)=u^5/5$

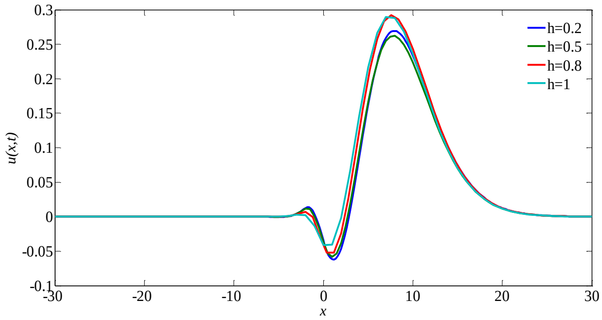


(c) $g(u)=u^9/9$

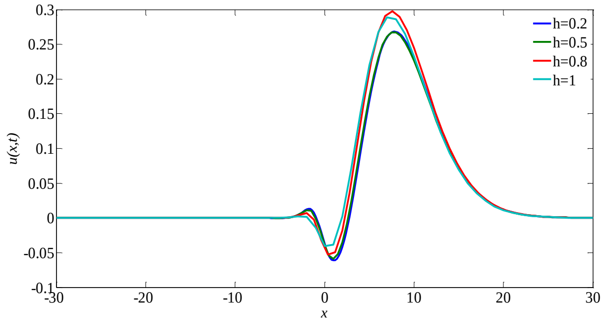
Figure 5. The behaviors of the wave in Example 5 using $h=0.2,\ \ \Delta t=0.1$ for various $g(u)$



(a) $g(u)=u^3/3$



(b) $g(u)=u^5/5$



(c) $g(u)=u^9/9$

Figure 6. Numerical solutions for Example 5 at $t=10$ using $\Delta t=0.1$ and various values of h with various $g(u)$

Acknowledgments

This research was supported by Suleyman Demirel University through Scientific Research Project Program (3539-D1-13).

References

[1] Bona, J.L., Pritchard, W.G. & Scott, L.R. (1981). An evaluation of a model equation for water waves. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 302(1471), 457-510.

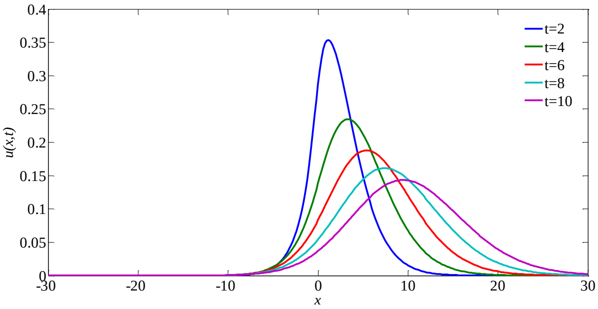
[2] Zhao, X., Li, D. & Shi, D. (2008). A finite difference scheme for RLW-Burgers equation. *Journal of Applied Mathematics & Informatics*, 26(3,4), 573-581.

[3] Al-Khaled, K., Momani, S. & Alawneh, A. (2005). Approximate wave solutions for generalized Benjamin-Bona-Mahony-Burgers equations. *Applied Mathematics and Computation*, 171(1), 281-292.

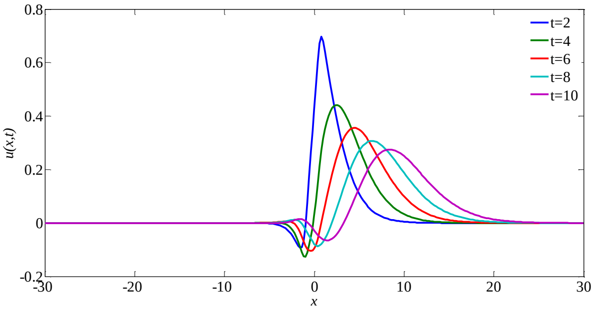
[4] Alquran, M. & Al-Khaled, K. (2011). Sinc and solitary wave solutions to the generalized Benjamin-Bona-Mahony-Burgers equations. *Physica Scripta*, 83(6), 6 pages.

[5] Arora, G., Mittal, R.C. & Singh, B.K. (2014). Numerical solution of BBM-Burgers equation with quartic B-spline collocation method. *Journal of Engineering Science and Technology*, 9, 104-116.

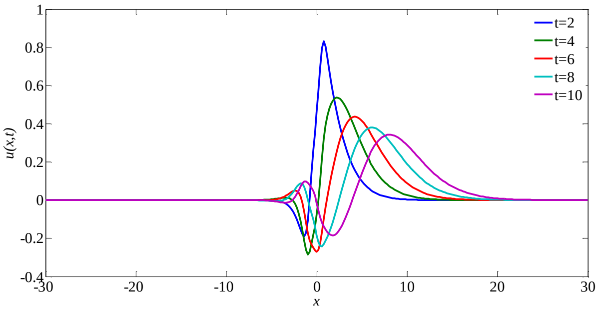
[6] Che, H., Pan, X., Zhang, L. & Wang, Y. (2012). Numerical analysis of a linear-implicit average scheme for generalized Benjamin-Bona-Mahony-Burgers equation. *Journal of Applied Mathematics*, 2012, 14 pages.



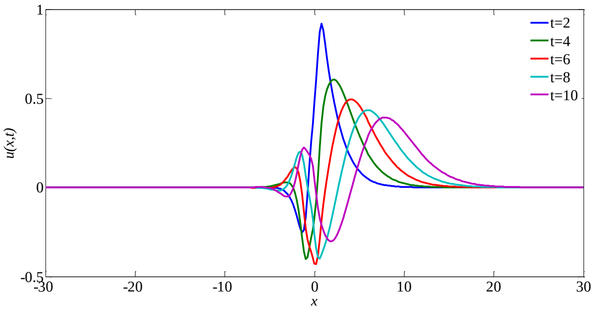
(a) $\alpha = 2$



(b) $\alpha = 0.5$



(c) $\alpha = 0.25$



(d) $\alpha = 0.125$

Figure 7. The behaviors of the wave in Example 5 using $h = 0.2$, $\Delta t = 0.1$ and $g(u) = u^3/3$ for various values of α.

[7] Omrani, K. & Ayadi, M. (2008). Finite difference discretization of the Benjamin-Bona-Mahony-Burgers equation. *Numerical Methods for Partial Differential Equations*, 24(1), 239-248.

[8] Zarebnia, M. & Parvaz, R. (2013). Cubic B-spline collocation method for numerical solution of the Benjamin-Bona-Mahony-Burgers equation. *WASET International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, 7(3), 540-543.

[9] Zarebnia, M. & Parvaz, R. (2016). On the numerical treatment and analysis of Benjamin-Bona-Mahony-Burgers equation. *Applied Mathematics and Computation*, 284, 79-88.

[10] Shen, Y. & Zha, G. (2008). A Robust seventh-order WENO scheme and its applications, 46th AIAA Aerospace Sciences Meeting and Exhibit, 2008 Jan 7-10, Reno, Nevada, AIAA 2008-075.

[11] Shen, Y. & Zha, G. (2010). Improved seventh order WENO scheme. 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, 2010 Jan 4-7, Orlando, Florida, AIAA 2010-1451 .

[12] Zahran, Y.H. & Babatin, M.M. (2013). Improved ninth order WENO scheme for hyperbolic conservation laws. *Applied Mathematics and Computation*, 219(15), 8198-8212.

[13] Jiang, G.S. & Shu, C.W. (1996). Efficient implementation of weighted ENO schemes. *Journal of Computational Physics*, 126(1), 202-228.

[14] Wang, Z.J. & Chen, R.F. (2001). Optimized weighted essentially non-oscillatory schemes for linear waves with discontinuity. *Journal of Computational Physics*, 174(1), 381-404.

[15] Ponziani, D., Prizzoli, S. & Grasso, F. (2003). Develpoment of optimized weighted-ENO schemes for multiscale compressible flows. *International Journal for Numerical Methods in Fluids*, 42(9), 953-977.

[16] Balsara, D.S. & Shu, C.W. (2000). Monotonicity preserving weighted essentially non-oscillatory schemes with increasingly high order of accuracy. *Journal of Computational Physics*, 160(2), 405-452.

[17] Pirozzoli, S. (2002). Conservative hybrid compact-WENO schemes for shock-turbulence interaction. *Journal of Computational Physics*, 178(1), 81-117.

[18] Kim, D. & Kwon, J.H. (2005). A high-order accurate hybrid scheme using a central flux scheme and a WENO scheme for compressible flowfield analysis. *Journal of Computational Physics*, 210(2), 554-583.

[19] Shen, Y.Q. & Yang, G.W. (2007). Hybrid finite compact WENO schemes for shock calculation. *International Journal for Numerical Methods in Fluids*, 53(4), 531-560.

[20] Sari, M., Gürarslan, G. & Zeytinoglu, A. (2010). High-order finite difference schemes for solving the advection diffusion equation. *Mathematical and Computational Applications*, 15(3), 449-460.

[21] Zeytinoglu, A. (2010). *Some approximate solutions of Burgers equations*. MSc Thesis. Suleyman Demirel University.

[22] Liu, X.D., Osher, S. & Chan, T. (1994). Weighted essentially non-oscillatory schemes. *Journal of Computational Physics*, 126, 200-212.

[23] Xie, P., (2007). *Uniform weighted compact/non-compact schemes for shock/boundary layer interaction.* PhD Thesis. The University of Texas.

[24] Pletcher, R.H., Tannehill, J.C. & Anderson, D.A. (2013). *Computational fluid mechanics and fluid transfer*. Taylor&Francis.

***Asuman Zeytinoglu** is a Research Assistant in the Department of Mathematics at Suleyman Demirel University, Isparta, Turkey. She received her BSc degree in Mathematics from Pamukkale University, Denizli, Turkey in 2007, and MSc degree in Mathematics from Suleyman Demirel University, Isparta, Turkey in 2010. She is currently continuing the PhD studies in Department of Mathematics at Suleyman Demirel University, Isparta, Turkey. Her research interests are numerical solutions of partial differential equations, computational methods and wave propagation.*

***Murat Sari** is an Associate Professor in the Department of Mathematics at Yildiz Technical University, Istanbul, Turkey. He received his BSc degree from Ondokuzmayis University, Turkey in 1991. He completed his graduate studies (MSc/PhD degrees) in Mathematics from University of Glamorgan, UK in 2000. His research interests are numerical methods, wave propagation, computational methods, mathematical modelling, biophysical modelling, medical modelling and economical modelling.*

***Bilender P. Allahverdiev** is a Professor in the Department of Mathematics at Suleyman Demirel University, Isparta, Turkey. He received his BSc degree in Mathematics from Baku State University, Baku, Azerbaijan in 1977. He received his PhD degree from Azerbaijan National Academy of Sciences, Azerbaijan in 1982. His research interests are operator theory, spectral theory of differential operators, Jacobian matrix and difference equations, Schrödinger equation and Sturm- Liouville boundary value problems.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.1, pp.17-25 (2018)
http://doi.org/10.11121/ijocta.01.2018.00532

RESEARCH ARTICLE

Analysis of rubella disease model with non-local and non-singular fractional derivatives

İlknur Koca*

Department of Mathematics, Faculty of Sciences, Mehmet Akif Ersoy University, 15100, Burdur, Turkey
ikoca@mehmetakif.edu.tr

ARTICLE INFO

Article History:
Received 14 August 2017
Accepted 02 October 2017
Available 18 October 2017

Keywords:
Rubella disease model
Special solution
Fixed point theorem
Numerical simulations

AMS Classification 2010:
34A34, 47H10, 65L07

ABSTRACT

In this paper we investigate a possible applicability of the newly established fractional differentiation in the field of epidemiology. To do this we extend the model describing the Rubella spread by changing the derivative with the time fractional derivative for the inclusion of memory. Detailed analysis of existence and uniqueness of exact solution is presented using the Banach fixed point theorem. Finally some numerical simulations are showed to underpin the effectiveness of the used derivative.

1. Introduction

The aim of mathematical biology is to develop mathematical equations and to describe some physical problems encountered in biology. Noting that, the establishment of such mathematical formula is achieved using the concept of differentiation or more practically the notion of derivatives. There exist two classes of differentiation in the literatures. The first one is based on the concept of rate of change [8-11,21]. The second one is based on the convolution of some functions including exponential decay law and the generalized Mittag-Leffler law. The derivatives based on exponential appear naturally in many problems in nature as being able to describe the effect of fading memory. This class of derivative has been applied in several research papers for instance [5,7,13,15,16,18-20,22]. However, it was noted by several experts in the field that, this new derivative does not have a non-local kernel as its corresponding integral is not fractional, thus a new kernel was suggested by Atangana and Baleanu [6] where after some manipulations, the exponential decay kernel was replaced by the generalized Mittag-Leffler kernel. This last derivative, therefore appears to be a very powerful mathematical tools form modeling real world problems as the generalized Mittag-Leffler function is combination of the power law and exponential decay law.

Several research papers have been published using this new concept of fractional differentiation with Mittag-Leffler. More importantly the results obtained in [1-4,6] revealed that, the new concept of more adequate for modeling real world problems to take into account the non-locality and also to have a memory effect. We shall note that the choice of a kernel is very important when modeling real world problems. When looking at experimental data obtained from real world observations, we can see that, many biological problems may not always follow the power law based on the function $x^{-\alpha}$ which is the kernel mostly used in the literature nowadays. For instance the case of Rubella, which is also known as the German measles or more precisely the three-day measles is enveloped and has a single-stranded RNA genome. The virus spreads via breathing

*Corresponding Author

route and photocopies in the nasopharynx and lymph nodes. This virus can only be detected in the stream blood after a period of between 5 to 7 days when the infection has taken place, later spreads throughout the body. With its properties of teratogenic and the ability of overpass the placenta and infecting the fetus where it stops cells developing or destroys them. Such a complex dynamic will be suitable to portray a more advance concept of power with of course a non-local concept which is the property inherited by the newly established derivative with fractional order called Atangana-Baleanu derivatives [6]. This paper is therefore devoted to the analysis of the dynamic of the spread of Rubella virus exploring the Atangana-Baleanu fractional derivative. The aim of the research in this field, requires the use of the new fractional derivative for Rubella disease virus. The exactness and uniqueness of the solution of the fractional model is proved by applying the fixed-point theorem.

The remainder part of this paper is broken into sections. In Section 2, we give the definitions of the new fractional derivative with non-singular and non-local kernel. Section 3 deals with the existence of solutions for the spread of rubella disease model via Picard-Lindelof method. In Section 4, we provide a special solution of the model which is considered using Atangala-Balenau derivative in Caputo sense. Finally in Section 5, some numerical results obtained at different instances of fractional order are presented to justify the suitability of the adopted derivative.

2. New fractional derivative with non-singular and non-local kernel

Let us remind the definitions of the new fractional derivative with non-singular and non-local kernel [6].

Definition 1. *Let $f \in H^1(a,b)$, $b > a$, $\alpha \in [0,1]$ then, the definition of the new fractional derivative (Atangana-Baleanu derivative in Caputo sense) is given as:*

$$ {}_{a}^{ABC}D_t^{\alpha}\left(f\left(t\right)\right) = \frac{B(\alpha)}{1-\alpha}\int_a^t f'(x)E_{\alpha}\left[-\alpha\frac{(t-x)^{\alpha}}{1-\alpha}\right]dx, \tag{1}$$

where ${}_{a}^{ABC}D_t^{\alpha}$ is fractional operator with Mittag-Leffler kernel in the Caputo sense with order α with respect to t and $B(\alpha) = B(0) = B(1) = 1$ is a normalization function [12].

Definition 2. *Let $f \in H^1(a,b)$, $b > a$, $\alpha \in [0,1]$ and not differentiable then, the definition of the new fractional derivative (Atangana-Baleanu fractional derivative in Riemann-Liouville sense) is given as:*

$$ {}_{a}^{ABR}D_t^{\alpha}\left(f\left(t\right)\right) = \frac{B(\alpha)}{1-\alpha}\frac{d}{dt}\int_a^t f(x)E_{\alpha}\left[-\alpha\frac{(t-x)^{\alpha}}{1-\alpha}\right]dx. \tag{2}$$

Definition 3. *The fractional integral of order α of a new fractional derivative is defined as:*

$$ {}_{a}^{AB}I_t^{\alpha}\left\{f(t)\right\} = \frac{1-\alpha}{B(\alpha)}f(t) + \frac{\alpha}{B(\alpha)\Gamma(\alpha)}\int_a^t f(y)(t-y)^{\alpha-1}dy. \tag{3}$$

When α is zero, initial function is obtained and when α is 1, the ordinary integral is obtained.

3. Existence of solutions for the spread of rubella disease model

Let us consider the following model employing the Atangana-Baleanu fractional derivative in Caputo sense :

$$\begin{aligned}
{}_{0}^{ABC}D_t^{\alpha}S\left(t\right) &= B(a) - \left[\lambda(a,t) + P(a) + \mu(a)\right]S\left(t\right),\\
{}_{0}^{ABC}D_t^{\alpha}E\left(t\right) &= \lambda(a,t)S\left(t\right) - \left(\sigma + \mu(a)\right)E\left(t\right),\\
{}_{0}^{ABC}D_t^{\alpha}I\left(t\right) &= \sigma E\left(t\right) - \left(\beta + \mu(a)\right)I\left(t\right),\\
{}_{0}^{ABC}D_t^{\alpha}R\left(t\right) &= \beta I\left(t\right) - \mu(a)R\left(t\right),\\
{}_{0}^{ABC}D_t^{\alpha}V\left(t\right) &= D(a)S\left(t\right) - \mu(a)V\left(t\right),
\end{aligned} \tag{4}$$

where $S\left(t\right)$, $E\left(t\right)$, $I\left(t\right)$, $R\left(t\right)$, $V\left(t\right)$ are susceptible, latent, infectious, recovered and vaccinated parameters respectively. $P(a)$ is a parameter for which immunized by vaccination and $\lambda(a,t)$ is the force of infection of age a at time t. Finally, σ is the latent rate and β is the infection rate [14]. The aim of this section is to find existence of solutions for rubella disease model with Atangana-Balenau fractional derivative. The system state is made up with S, E, I, R, V. The above system (4) can be converted to Volterra type integral equation with the Atangana-Baleanu fractional integral.

Theorem 1. *The following time fractional ordinary differential equation*

$$ {}_{0}^{ABC}D_t^{\alpha}(f\left(t\right)) = u(t), \tag{5}$$

has a unique solution with taking the inverse Laplace transform and using the convolution theorem below [4]:

$$f(t) = \frac{1-\alpha}{B(\alpha)} u(t) + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_a^t u(y)(t-y)^{\alpha-1} dy. \tag{6}$$

By the theorem above, the model can be written as (7):

$$\begin{cases} S(t) - g_1(t) = \frac{1-\alpha}{B(\alpha)} \{B(a) - [\lambda(a,t) + P(a) + \mu(a)] S(t)\} \\ + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \\ \times \{B(a) - [\lambda(a,y) + P(a) + \mu(a)] S(y)\} dy, \\ E(t) - g_2(t) = \frac{1-\alpha}{B(\alpha)} \{\lambda(a,t) S(t) - (\sigma + \mu(a)) E(t)\} \\ + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \\ \times \{\lambda(a,y) S(y) - (\sigma + \mu(a)) E(y)\} dy, \\ I(t) - g_3(t) = \frac{1-\alpha}{B(\alpha)} \{\sigma E(t) - (\beta + \mu(a)) I(t)\} \\ + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \{\sigma E(y) - (\beta + \mu(a)) I(y)\} dy, \\ R(t) - g_4(t) = \frac{1-\alpha}{B(\alpha)} \{\beta I(t) - \mu(a) R(t)\} \\ + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \{\beta I(y) - \mu(a) R(y)\} dy, \\ V(t) - g_5(t) = \frac{1-\alpha}{B(\alpha)} \{D(a) S(t) - \mu(a) V(t)\} \\ + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \{D(a) S(y) - \mu(a) V(y)\} dy, \end{cases} \tag{7}$$

The above system (7) of equations can be iteratively represented as:

$$\begin{cases} S_0(t) = g_1(t), \\ E_0(t) = g_2(t), \\ I_0(t) = g_3(t), \\ R_0(t) = g_4(t), \\ V_0(t) = g_5(t). \end{cases} \tag{8}$$

$$\begin{aligned} S_{n+1}(t) &= \frac{1-\alpha}{B(\alpha)} \\ &\quad \times \{B(a) - [\lambda(a,t) + P(a) + \mu(a)] S_n(t)\} \\ &\quad + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \\ &\quad \times \{B(a) - [\lambda(a,y) + P(a) + \mu(a)] S_n(y)\} dy, \end{aligned} \tag{9}$$

$$\begin{aligned} E_{n+1}(t) &= \frac{1-\alpha}{B(\alpha)} \{\lambda(a,t) S_n(t) - (\sigma + \mu(a)) E_n(t)\} \\ &\quad + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \\ &\quad \times \{\lambda(a,y) S_n(y) - (\sigma + \mu(a)) E_n(y)\} dy, \end{aligned}$$

$$\begin{aligned} I_{n+1}(t) &= \frac{1-\alpha}{B(\alpha)} \{\sigma E_n(t) - (\beta + \mu(a)) I_n(t)\} \\ &\quad + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \\ &\quad \times \{\sigma E_n(y) - (\beta + \mu(a)) I_n(y)\} dy, \end{aligned}$$

$$\begin{aligned} R_{n+1}(t) &= \frac{1-\alpha}{B(\alpha)} \{\beta I_n(t) - \mu(a) R_n(t)\} \\ &\quad + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \\ &\quad \times \{\beta I_n(y) - \mu(a) R_n(y)\} dy, \end{aligned}$$

$$\begin{aligned} V_{n+1}(t) &= \frac{1-\alpha}{B(\alpha)} \{D(a) S_n(t) - \mu(a) V_n(t)\} \\ &\quad + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-y)^{\alpha-1} \\ &\quad \times \{D(a) S_n(y) - \mu(a) V_n(y)\} dy. \end{aligned}$$

As the exact solution of the iterative formula of a Picard series used here converges toward the exact solution as the number of series terms tends to infinity. If we take the limit with greater than n, we expect to obtain the exact solution of equation as below:

$$\begin{cases} \lim_{n\to\infty} S_n(t) = S(t), \\ \lim_{n\to\infty} E_n(t) = E(t), \\ \lim_{n\to\infty} I_n(t) = I(t), \\ \lim_{n\to\infty} R_n(t) = R(t), \\ \lim_{n\to\infty} V_n(t) = V(t). \end{cases}$$

3.1. Existence of solution via Picard-Lindelof method

Let us define the following operator for showing the existence of solution:

$$\begin{aligned} f_1(a,t) &= B(a) - [\lambda(a,t) + P(a) + \mu(a)] S(t), \\ f_2(a,t) &= \lambda(a,t) S(t) - (\sigma + \mu(a)) E(t), \\ f_3(a,t) &= \sigma E(t) - (\beta + \mu(a)) I(t), \\ f_4(a,t) &= \beta I(t) - \mu(a) R(t), \\ f_5(a,t) &= D(a) S(t) - \mu(a) V(t). \end{aligned} \tag{10}$$

Let

$$\begin{aligned} N_1 &= \sup_{C[b,c_1]} \|f_1(a,t)\|, \ N_2 = \sup_{C[b,c_2]} \|f_2(a,y)\|, \\ N_3 &= \sup_{C[b,c_3]} \|f_3(a,z)\|, \ N_4 = \sup_{C[b,c_4]} \|f_4(a,p)\|, \\ N_5 &= \sup_{C[b,c_5]} \|f_5(a,r)\|, \end{aligned} \tag{11}$$

where

$$\begin{aligned} C[b,c_1] &= [t-b,t+b]\times[x-c_1,x+c_1] = B_1\times C_1,\\ C[b,c_2] &= [t-b,t+b]\times[x-c_2,x+c_2] = B_1\times C_2,\\ C[b,c_3] &= [t-b,t+b]\times[x-c_3,x+c_3] = B_1\times C_3,\\ C[b,c_4] &= [t-b,t+b]\times[x-c_4,x+c_4] = B_1\times C_4,\\ C[b,c_5] &= [t-b,t+b]\times[x-c_5,x+c_5] = B_1\times C_5. \end{aligned} \tag{12}$$

We will make use of Banach fixed-point theorem using the metric on $C[b,c_i]$, $(i=1,2,...,5)$ made by the uniform norm

$$\|X(t)\|_\infty = \sup_{t\in[t-b,t+b]}|f(t)|. \tag{13}$$

The next operator is defined between the two functional spaces of continuous functions, Picard's operator as follows:

$$\begin{aligned} O &: C(B_1,C_1,C_2,C_3,C_4,C_5)\\ &\to C(B_1,C_1,C_2,C_3,C_4,C_5). \end{aligned} \tag{14}$$

For simplicity, let us define $f_i(a,t) = X(t)$, $f_i(a,0) = X_0(t)$, $(i=1,2,...,5)$. Then the system is reduced the following:

$$\begin{aligned} OX(t) &= X_0(t) + F(t,X(t))\frac{1-\alpha}{B(\alpha)}\\ &+ \frac{\alpha}{B(\alpha)\Gamma(\alpha)}\int_0^t (t-y)^{\alpha-1}F(y,X(y))dy, \end{aligned} \tag{15}$$

where X is the matrice of given as

$$X(t) = \begin{cases} S(t)\\ E(t)\\ I(t)\\ R(t)\\ V(t)\end{cases}, X_0(t) = \begin{cases} S(0)\\ E(0)\\ I(0)\\ R(0)\\ V(0)\end{cases}, \tag{16}$$

$$F(a,X(t)) = \begin{cases} f_1(a,t)\\ f_2(a,t)\\ f_3(a,t)\\ f_4(a,t)\\ f_5(a,t)\end{cases}.$$

Let us assume that the physical problem under investigation satisfies followings:

$$\|X(t)\|_\infty \le \max\{c_1,c_2,c_3,c_4,c_5\}. \tag{17}$$

$$\begin{aligned} &\|OX(t) - X_0(t)\| \qquad (18)\\ &= \left\| F(t,X(t))\tfrac{1-\alpha}{B(\alpha)} + \tfrac{\alpha}{B(\alpha)\Gamma(\alpha)}\int_0^t (t-y)^{\alpha-1}F(y,X(y))dy \right\|\\ &\le \frac{1-\alpha}{B(\alpha)}\|F(t,X(t))\|\\ &+ \frac{\alpha}{B(\alpha)\Gamma(\alpha)}\int_0^t (t-y)^{\alpha-1}\|F(y,X(y))\|dy\\ &\le \frac{1-\alpha}{B(\alpha)}N = \max\{N_1,N_2,N_3,N_4,N_5\}\\ &+ \frac{\alpha}{B(\alpha)}Nb^\alpha < bN \le c = \max\{c_1,c_2,c_3,c_4,c_5\}, \end{aligned}$$

where we demand that

$$b < \frac{c}{N}.$$

Also we evaluate the following equality

$$\|OX_1 - OX_2\|_\infty = \sup_{t\in B}|X_1 - X_2|. \tag{19}$$

Nonetheless using the definition of our defined operator, we have

$$\|OX_1 - OX_2\|$$

$$= \left\| \begin{array}{c} \{F(t,X_1(t)) - F(t,X_2(t))\}\frac{1-\alpha}{B(\alpha)}\\ +\frac{\alpha}{B(\alpha)\Gamma(\alpha)}\int_0^t (t-l)^{\alpha-1}\left\{\begin{array}{c} F(l,X_1(l))\\ -F(l,X_2(l))\end{array}\right\}dl \end{array}\right\| \tag{20}$$

$$\begin{aligned} &\le \frac{1-\alpha}{B(\alpha)}\|F(t,X_1(t)) - F(t,X_2(t))\| \qquad (21)\\ &+ \frac{\alpha}{B(\alpha)\Gamma(\alpha)}\int_0^t (t-y)^{\alpha-1}\\ &\times \|F(l,X_1(y)) - F(l,X_2(y))\|dy\\ &\le \frac{1-\alpha}{B(\alpha)}q\|X_1(t) - X_2(t)\|\\ &+ \frac{\alpha q}{B(\alpha)\Gamma(\alpha)}\int_0^t (t-y)^{\alpha-1}\|X_1(y) - X_2(y)\|dy\\ &\le \left\{\frac{1-\alpha}{B(\alpha)}q + \frac{\alpha q b^\alpha}{B(\alpha)\Gamma(\alpha)}\right\}\|X_1(t) - X_2(t)\|\\ &\le bq\|X_1(t) - X_2(t)\|. \end{aligned}$$

with $q < 1$ since F is a contraction we have that $bq < 1$, thus the defined operator O is a contraction. So system has a unique set of solution.

4. Special solutions via iteration approach

The aim of this section is to provide a special solution of the model which is considered using

Atangala-Balenau derivative in Caputo sense. Let us apply the Sumudu transform on both sides of equation (4) together with an iterative method. We shall give the Sumudu transform of Atangana-Balenau fractional derivative in Caputo sense below:

Theorem 2. *Let $f \in H^1(a,b)$, $b > a$, $\alpha \in [0,1]$ then, the Sumudu transform of Atangana-Balenau fractional derivative in Caputo sense is given as:*

$$\begin{aligned}&ST\left\{{}_0^{ABC}D_t^{\alpha}(f(t))\right\} \\ &= \frac{B(\alpha)}{1-\alpha}\left(\alpha\Gamma(\alpha+1)E_\alpha(-\frac{1}{1-\alpha}p^\alpha)\right) \\ &\times\left(ST(f(t)) - f(0)\right).\end{aligned} \tag{22}$$

Proof. Proof of the theorem can be found in [4]. □

To solve Equation (4), we apply the Sumudu transform of the Atangana-Balenau fractional derivative of $f(t)$ on system with both sides. Then we obtain below:

$$\begin{aligned}&\tfrac{B(\alpha)}{1-\alpha}\left(\alpha\Gamma(\alpha+1)E_\alpha(-\tfrac{1}{1-\alpha}p^\alpha)\right)(ST(S(t)) - S(0)) \\ &= ST\left\{B(a) - [\lambda(a,t) + P(a) + \mu(a)]\,S(t)\right\}, \\ &\tfrac{B(\alpha)}{1-\alpha}\left(\alpha\Gamma(\alpha+1)E_\alpha(-\tfrac{1}{1-\alpha}p^\alpha)\right)(ST(E(t)) - E(0)) \\ &= ST\left\{\lambda(a,t)S(t) - (\sigma + \mu(a))\,E(t)\right\}, \\ &\tfrac{B(\alpha)}{1-\alpha}\left(\alpha\Gamma(\alpha+1)E_\alpha(-\tfrac{1}{1-\alpha}p^\alpha)\right)(ST(I(t)) - I(0)) \\ &= ST\left\{\sigma E(t) - (\beta + \mu(a))\,I(t)\right\}, \\ &\tfrac{B(\alpha)}{1-\alpha}\left(\alpha\Gamma(\alpha+1)E_\alpha(-\tfrac{1}{1-\alpha}p^\alpha)\right)(ST(R(t)) - R(0)) \\ &= ST\left\{\beta I(t) - \mu(a)R(t)\right\}, \\ &\tfrac{B(\alpha)}{1-\alpha}\left(\alpha\Gamma(\alpha+1)E_\alpha(-\tfrac{1}{1-\alpha}p^\alpha)\right)(ST(V(t)) - V(0)) \\ &= ST\left\{D(a)S(t) - \mu(a)V(t)\right\}.\end{aligned} \tag{23}$$

Rearranging, we obtain following inequalities where,

$$\begin{aligned}&ST(S(t)) = S(0) \\ &+ \theta * ST\left\{B(a) - [\lambda(a,t) + P(a) + \mu(a)]\,S(t)\right\}, \\ &ST(E(t)) = E(0) \\ &+ \theta * ST\left\{\lambda(a,t)S(t) - (\sigma + \mu(a))\,E(t)\right\}, \\ &ST(I(t)) = I(0) \\ &+ \theta * ST\left\{\sigma E(t) - (\beta + \mu(a))\,I(t)\right\}, \\ &ST(R(t)) = R(0) \\ &+ \theta * ST\left\{\beta I(t) - \mu(a)R(t)\right\}, \\ &ST(V(t)) = V(0) \\ &+ \theta * ST\left\{D(a)S(t) - \mu(a)V(t)\right\}.\end{aligned}$$

For simplicity, here

$$\theta = \frac{1-\alpha}{B(\alpha)\left(\alpha\Gamma(\alpha+1)E_\alpha(-\frac{1}{1-\alpha}p^\alpha)\right)},$$

is considered and "$*$" means multiplication sign . We next obtain the following recursive formula;

$$\begin{aligned}&S_{n+1}(t) = S_n(0) \\ &+ST^{-1}\left\{\theta * ST\left\{B(a) - [\lambda(a,t) + P(a) + \mu(a)]\,S_n(t)\right\}\right\}, \\ &E_{n+1}(t) = E_n(0) \\ &+ST^{-1}\left\{\theta * ST\left\{\lambda(a,t)S_n(t) - (\sigma + \mu(a))\,E_n(t)\right\}\right\}, \\ &I_{n+1}(t) = I_n(0) \\ &+ST^{-1}\left\{\theta * ST\left\{\sigma E_n(t) - (\beta + \mu(a))\,I_n(t)\right\}\right\}, \\ &R_{n+1}(t) = R_n(0) \\ &+ST^{-1}\left\{\theta * ST\left\{\beta I_n(t) - \mu(a)R_n(t)\right\}\right\}, \\ &V_{n+1}(t) = V_n(0) \\ &+ST^{-1}\left\{\theta * ST\left\{D(a)S_n(t) - \mu(a)V_n(t)\right\}\right\}.\end{aligned} \tag{24}$$

Therefore, the solution of equation (24) approximate to following

$$\begin{aligned}S(t) &= \lim_{n\to\infty} S_n(t), \\ E(t) &= \lim_{n\to\infty} E_n(t), \\ I(t) &= \lim_{n\to\infty} I_n(t), \\ R(t) &= \lim_{n\to\infty} R_n(t), \\ V(t) &= \lim_{n\to\infty} V_n(t).\end{aligned} \tag{25}$$

4.1. Application of fixed-point theorem for stability analysis of iteration method

Let $(X, \|.\|)$ be a Banach space and H a self-map of X. Let $y_{n+1} = g(H, y_n)$ be recurcive procedure. Suppose that, $F(H)$ the fixed-point set of

H has at least one element and that $y_n$converges to a point $p \in F(H)$. Let $\{x_n\} \subseteq X$ and define $e_n = \|x_{n+1} - g(H, x_n)\|$. If $\lim_{n\to\infty} e_n = 0$ implies that $\lim_{n\to\infty} x_n = p$, then the iteration method $y_{n+1} = g(H, y_n)$ is H-Stable. Then let we assume that, our sequence $\{x_n\}$ has an upper boundary . If all these conditions are satisfied for $y_{n+1} = Hy_n$ which is known as Picard's iteration, consequently the iteration is H-Stable. We shall then state the following theorem.

Theorem 3. *Let $(X, \|.\|)$ be a Banach space and H a self-map of X satisfying*

$$\|H_x - H_y\| \leq K\|x - H_x\| + k\|x - y\|,$$

for all x, y in X where $0 \leq K$, $0 \leq k < 1$. Suppose that H is Picard H-Stable [17].

Let us consider the following recursive formula equation (27) with (4) where

$$\theta = \frac{1-\alpha}{B(\alpha)\left(\alpha\Gamma(\alpha+1)E_\alpha(-\frac{1}{1-\alpha}p^\alpha)\right)}, \tag{26}$$

is the fractional Lagrange multiplier.

Theorem 4. *Let H be a self-map defined as (27) as below.*

$$H(S_n(t)) = S_{n+1}(t) = S_n(t) \tag{27}$$
$$+ST^{-1}\{\theta * ST\{B(a) - [\lambda(a,t) + P(a) + \mu(a)] S_n(t)\}\},$$
$$H(E_n(t)) = E_{n+1}(t) = E_n(t)$$
$$+ST^{-1}\{\theta * ST\{\lambda(a,t)S_n(t) - (\sigma + \mu(a)) E_n(t)\}\},$$
$$H(I_n(t)) = I_{n+1}(t) = I_n(t)$$
$$+ST^{-1}\{\theta * ST\{\sigma E_n(t) - (\beta + \mu(a)) I_n(t)\}\},$$
$$H(R_n(t)) = R_{n+1}(t) = R_n(t)$$
$$+ST^{-1}\{\theta * ST\{\beta I_n(t) - \mu(a) R_n(t)\}\},$$
$$H(V_n(t)) = V_{n+1}(t) = V_n(t)$$
$$+ST^{-1}\{\theta * ST\{D(a)S_n(t) - \mu(a)V_n(t)\}\}.$$

Then (27) is H-stable in $L^1(a,b)$ if following statement can be obtained.

$$\begin{aligned}
(1 - [\lambda(a,t) + P(a) + \mu(a)] A(\gamma)) &< 1, \\
(1 + \lambda(a,t)B(\gamma) - (\sigma + \mu(a))C(\gamma)) &< 1, \\
(1 + \sigma D(\gamma) - (\beta + \mu(a))E(\gamma)) &< 1, \\
(1 + \beta F(\gamma) - \mu(a)G(\gamma)) &< 1, \\
(1 + D(a)H(\gamma) - \mu(a)J(\gamma)) &< 1.
\end{aligned} \tag{28}$$

Proof. Let we start with showing that H has a fixed point. To achieve this, we evaluate the followings for all $(n, m) \in \mathbb{N} \times \mathbb{N}$.

$$H(S_n(t)) - H(S_m(t)) = S_n(t) - S_m(t) \tag{29}$$
$$+ST^{-1}\{\theta * ST\{B(a) - [\lambda(a,t) + P(a) + \mu(a)] S_n(t)\}\}$$
$$-ST^{-1}\{\theta * ST\{B(a) - [\lambda(a,t) + P(a) + \mu(a)] S_m(t)\}\}.$$

Let us consider (29) and apply norm on both sides and without loss of generality

$$\|H(S_n(t)) - H(S_m(t))\| \tag{30}$$
$$= \left\| \begin{array}{c} S_n(t) - S_m(t) \\ +ST^{-1}\left\{\theta * ST\left\{ \begin{array}{c} B(a) - [\lambda(a,t) + P(a) + \mu(a)] S_n(t) \\ -(B(a) - [\lambda(a,t) + P(a) + \mu(a)] S_m(t)) \end{array} \right\}\right\} \end{array} \right\|$$

$$\leq \|S_n(t) - S_m(t)\| \tag{31}$$
$$+\|ST^{-1}\{\theta * ST\{-[\lambda(a,t) + P(a) + \mu(a)](S_n(t) - S_m(t))\}\}\|.$$

Now we obtain :

$$\|H(S_n(t)) - H(S_m(t))\| \leq \|S_n(t) - S_m(t)\|$$
$$\times (1 - [\lambda(a,t) + P(a) + \mu(a)] A(\gamma)), \tag{32}$$

where $A(\gamma)$ is the $ST^{-1}\{\theta * ST\}$. Since all solutions have same role also we have following:

$$\begin{aligned}
&\|H(E_n(t)) - H(E_m(t))\| \leq \|E_n(t) - E_m(t)\| \\
&\times (1 + \lambda(a,t)B(\gamma) - (\sigma + \mu(a))C(\gamma)), \\
&\|H(I_n(t)) - H(I_m(t))\| \leq \|I_n(t) - I_m(t)\| \\
&\times (1 + \sigma D(\gamma) - (\beta + \mu(a))E(\gamma)), \\
&\|H(R_n(t)) - H(R_m(t))\| \leq \|R_n(t) - R_m(t)\| \\
&\times (1 + \beta F(\gamma) - \mu(a)G(\gamma)), \\
&\|H(V_n(t)) - H(V_m(t))\| \leq \|V_n(t) - V_m(t)\| \\
&\times (1 + D(a)H(\gamma) - \mu(a)J(\gamma)).
\end{aligned} \tag{33}$$

For

$$\begin{aligned}
(1 - [\lambda(a,t) + P(a) + \mu(a)] A(\gamma)) &< 1, \\
(1 + \lambda(a,t)B(\gamma) - (\sigma + \mu(a))C(\gamma)) &< 1, \\
(1 + \sigma D(\gamma) - (\beta + \mu(a))E(\gamma)) &< 1, \\
(1 + \beta F(\gamma) - \mu(a)G(\gamma)) &< 1, \\
(1 + D(a)H(\gamma) - \mu(a)J(\gamma)) &< 1,
\end{aligned} \tag{34}$$

then H-self mapping has a fixed point. Also nonlinear mapping H has to satisfy the conditions. So let we assume

$$K = (0,0,0,0,0)$$
$$k = \left\{ \begin{array}{c} (1 - [\lambda(a,t) + P(a) + \mu(a)] A(\gamma)) \\ (1 + \lambda(a,t)B(\gamma) - (\sigma + \mu(a))C(\gamma)) \\ (1 + \sigma D(\gamma) - (\beta + \mu(a))E(\gamma)) \\ (1 + \beta F(\gamma) - \mu(a)G(\gamma)) \\ (1 + D(a)H(\gamma) - \mu(a)J(\gamma)) \end{array} \right., \tag{35}$$

then all conditions of Theorem 3 hold. This completes the proof. □

5. Numerical Simulation

In this part, we present the numerical replication of the model for different values of fractional order using the proposed numerical scheme. The numerical simulations are shown in figure 1, 2, 3, and 4. Figures 1 is considered alpha to be 0.95, figure 2 is considered alpha to be 0.65, figure 3 is considered alpha to be 0.45 and finally in figure 4 is considered alpha to be 0.05. The paremeters used in this simulations are given below:

$$\begin{array}{lll} B = 100, & P = 0.3, & \lambda = 0.4, \\ \mu = 0.4, & \sigma = 0.3, & \beta = 0.4 \end{array}. \tag{36}$$

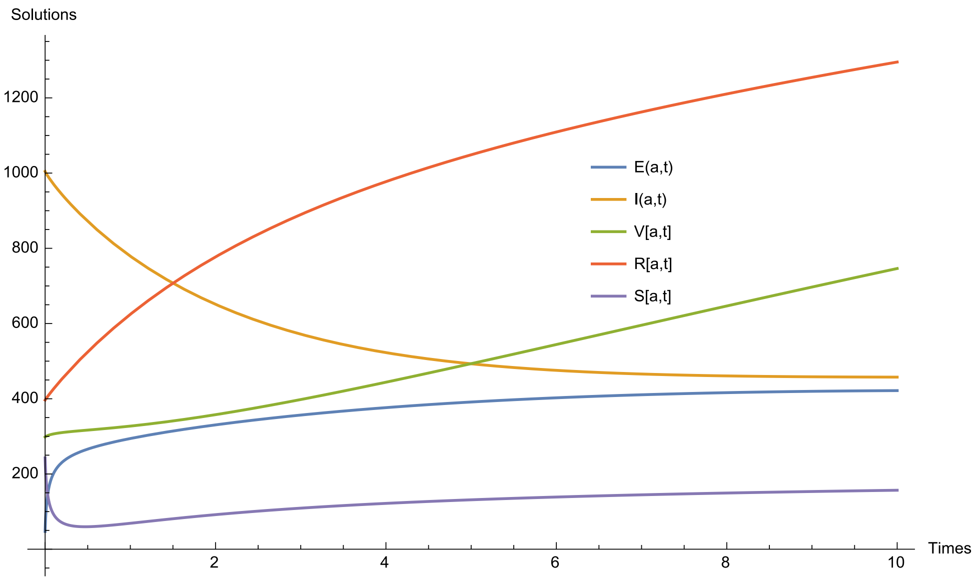


Figure 1 : Numerical simulation of solution for $\alpha = 0.95$

n

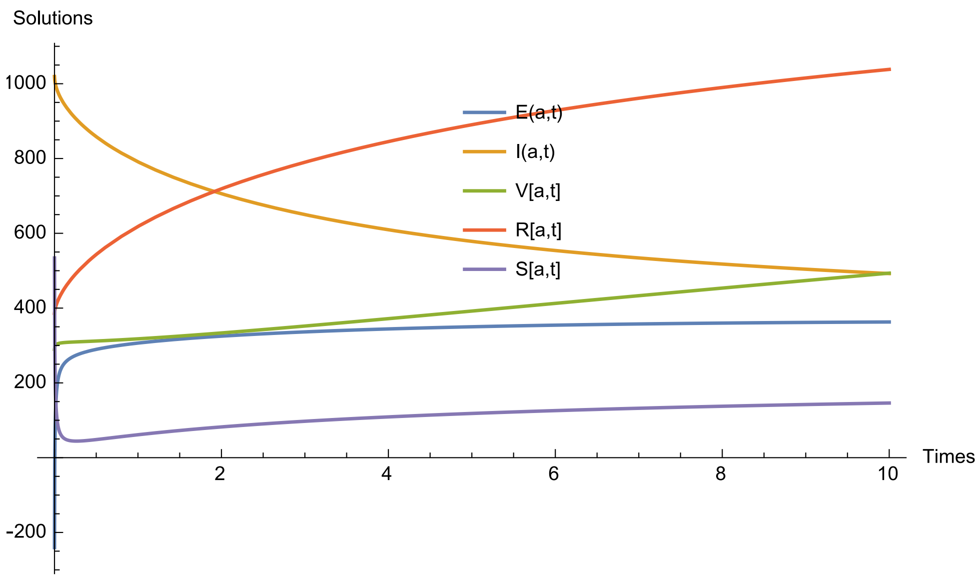


Figure 2 : Numerical simulation of solution for $\alpha = 0.65$

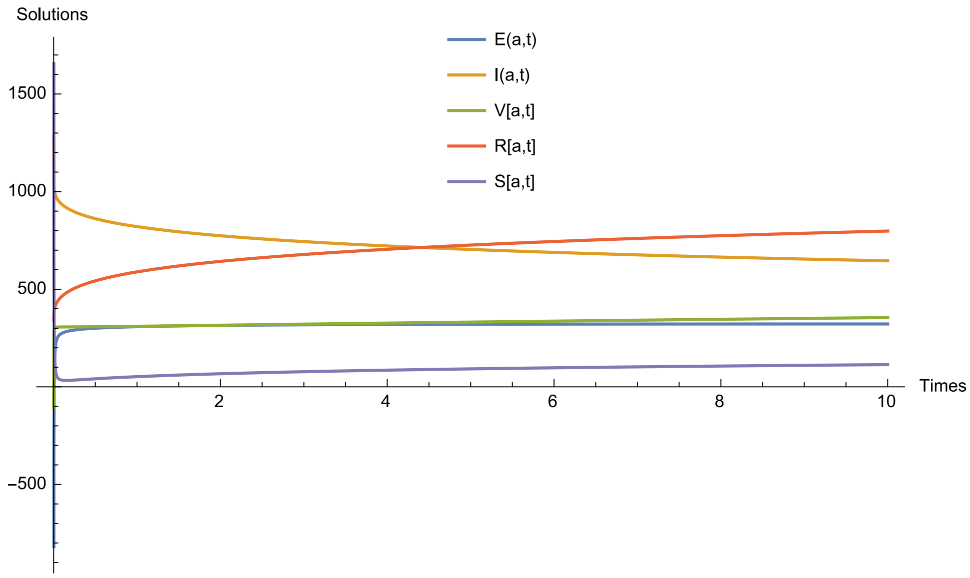


Figure 3 : Numerical simulation of solution for $\alpha = 0.45$

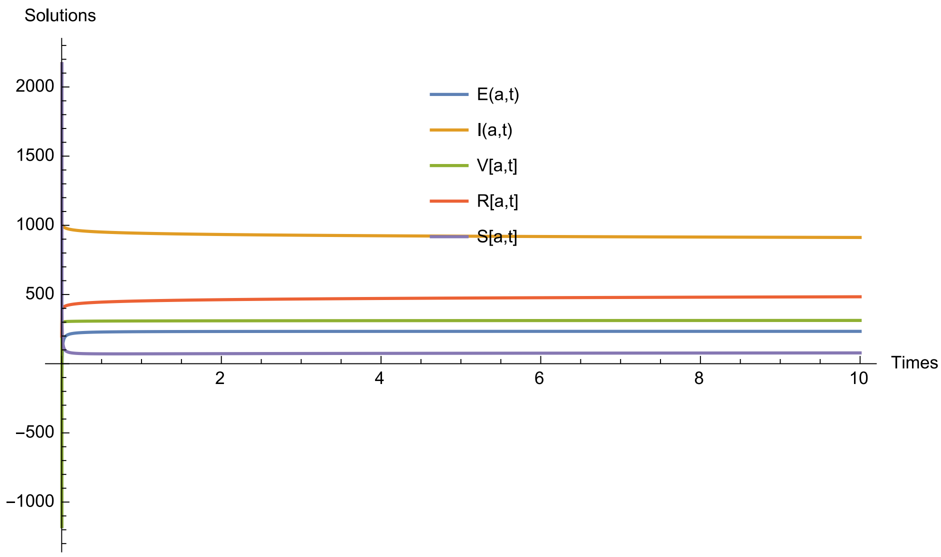


Figure 4 : Numerical simulation of solution for $\alpha = 0.05$

6. Conclusion

In this work, we have extended the model of rubella disease to the concept of fractional differential based on the Mittag-Leffler. We studied the existence of the generalized model using the fixed-point theorem. We presented the derivation of the solution using the Sumudu transform of Atanagana-Balenau derivative in Caputo sense. The stability analysis of the method is validated with the H-stable approach. Finally. Numerical simulations presented for different values of α.

References

[1] Abdeljawad, T. and Baleanu, D. (2016). Discrete fractional differences with nonsingular discrete Mittag-Leffler kernels. *Advances in Difference Equations*. (1), 232.

[2] Abdeljawad, T. and Baleanu, D. (2016). Integration by parts and its applications of a new nonlocal fractional derivative with Mittag-Leffler nonsingular kernel. arXiv preprint arXiv, 1607.00262.

[3] Alkahtani, B.S.T. (2016). Chua's circuit model with Atangana–Baleanu derivative with fractional order. *Chaos Solitons and Fractals*. 89, 547–551.

[4] Atangana, A. and Koca, I. (2016). Chaos in a simple nonlinear system with Atangana-Baleanu derivatives with fractional order. *Chaos, Solitons and Fractals*. 89, 447-454.

[5] Atangana, A. and Koca, I. (2016). On the new fractional derivative and application to Nonlinear Baggs and Freedman model. *Journal of Nonlinear Sciences and Applications*. (9), 2467-2480.

[6] Atangana, A. and Baleanu, D. (2016). New fractional derivatives with nonlocal and non-singular kernel: theory and application to heat transfer model. *Thermal Science*. 20 (2), 763-769.

[7] Atangana, A. and Owolabi, K.M. (2017). New numerical approach for fractional differential equations. preprint, arXiv, 1707.08177.

[8] Baskonus, H.M. and Bulut, H. (2015). On the numerical solutions of some fractional ordinary differential equations by fractional Adams-Bashforth-Moulton Method. *Open Mathematics*. 13 (1), 547–556.

[9] Baskonus, H.M. and Bulut, H. (2016). Regarding on the prototype solutions for the nonlinear fractional-order biological population model. *AIP Conference Proceedings*. 1738, 290004.

[10] Baskonus, H.M., Mekkaoui, T., Hammouch, Z. and Bulut, H. (2015). Active control of a chaotic fractional order economic system. *Entropy*. 17 (8), 5771-5783.

[11] Baskonus, H.M., Hammouch, Z., Mekkaoui, T. and Bulut, H. (2016). Chaos in the fractional order logistic delay system: circuit realization and synchronization. *AIP Conference Proceedings*, 1738, 290005.

[12] Caputo, M. & Fabrizio, M. (2015). A new definition of fractional derivative without singular kernel. *Progress in Fractional Differentiation and Applications*. 1, 73-85.

[13] Coronel-Escamilla, A., Gómez-Aguilar, J.F., Baleanu, D., Escobar-Jiménez, R. F., Olivares-Peregrino, V. H., and Abundez-Pliego, A. (2016). Formulation of Euler-Lagrange and Hamilton equations involving fractional operators with regular kernel. *Advances in Difference Equations*. (1), 283.

[14] Gay, N.J., Pelletier, L. and Duclos, P. (1998). Modelling the incidence of measles in Canada: An assessment of the options for vaccination policy. *Vaccine*. 16, 794–801.

[15] Gencoglu, M.T., Baskonus, H.M. and Bulut, H. (2017). Numerical simulations to the nonlinear model of interpersonal Relationships with time fractional derivative. *AIP Conference Proceedings*. 1798, 1-9, 020103.

[16] Gómez-Aguilar, J.F., Morales-Delgado, V.F., Taneco-Hernández, M.A., Baleanu, D., Escobar-Jiménez, R.F., and Al Qurashi, M.M. (2016). Analytical solutions of the electrical RLC circuit via Liouville–Caputo operators with local and non-Local kernels. *Entropy*. 18 (8), 402.

[17] Odibat, Z.M. and Momani, S. (2006). Application of variational iteration method to nonlinear differential equation of fractional order. *International Journal of Nonlinear Sciences and Numerical Simulations*. 7, 27-34.

[18] Owolabi, K.M. (2016). Numerical solution of diffusive HBV model in a fractional medium. *SpringerPlus*. 5, 19 pages.

[19] Owolabi, K.M. and Atangana, A. (2017). Numerical approximation of nonlinear fractional parabolic differential equations with Caputo–Fabrizio derivative in Riemann–Liouville sense. *Chaos, Solitons and Fractals*. 99, 171-179.

[20] Owolabi, K.M. (2017). Mathematical modelling and analysis of two-component system with Caputo fractional derivative order. *Chaos Solitons and Fractals*. 103, 544-554.

[21] Ozalp, N. and Koca, I. (2012). A fractional order nonlinear dynamical model of interpersonal relationships. *Advances in Difference Equations*. 189, 7 pages.

[22] Singh, J., Kumar, D., Al Qurashi, M. and Baleanu, D. (2017). A new fractional model for giving up smoking dynamics. *Advances in Difference Equations*, 88 (1), 16 pages.

İlknur Koca *received the B.Sc. and M.Sc. degrees in 2007 and 2009 from the Department of Mathematics from Ankara University, Turkey respectively. She received her Ph.D. degree from the same university in 2013. She is working as an Assistant Professor at Department of Mathematics of Mehmet Akif Ersoy University. Her research interests are methods and applications of partial and ordinary differential equations, fractional differential equations, iterative methods.*

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.1, pp.26-42 (2018)
http://doi.org/10.11121/ijocta.01.2018.00388

RESEARCH ARTICLE

Discretization based heuristics for the capacitated multi-facility Weber problem with convex polyhedral barriers

M. Hakan Akyüz*

Department of Industrial Engineering, Faculty of Engineering and Technology, Galatasaray University, Turkey
mhakyuz@gsu.edu.tr

ARTICLE INFO

Article History:
Received 04 September 2016
Accepted 13 September 2017
Available 23 October 2017

Keywords:
Polyhedral barriers
Location-allocation
Mixed-integer linear programming
Discretization heuristics
Lagrangean relaxation

AMS Classification 2010:
49M25, 90C11, 90C59

ABSTRACT

The Capacitated Multi-facility Weber Problem (CMWP) tries to determine the location of I capacitated facilities in the plane and to satisfy demand of J customers so as to minimize the total transportation cost. The CMWP assumes that the facilities can be located anywhere on the plane and customers are directly connected to them. This study considers an extension of the CMWP where there exist convex polyhedral barriers blocking passage and locating facilities inside. As a result, the distances between facilities and customers have to be measured by taking into account the polyhedral barriers. The CMWP with convex polyhedral barriers (CMWP-B) is a non-convex problem that is difficult to solve. We propose specially tailored discretization based heuristic procedures. Since CMWP-B is novel to the literature, a new set of test problems is randomly generated. Then, the performance of the suggested methods are tested on the test instances. Our results imply that the suggested heuristics yield quite accurate and efficient solutions for the CMWP-B.

1. Introduction

The Capacitated Multi-facility Weber Problem (CMWP) tries to locate I facilities with capacity restrictions in the plane and to meet the demand of J customers while minimizing the total cost of transportation. The objective function of the CMWP is shown to be neither convex nor concave [1]. When capacity restrictions of the facilities are removed, the CMWP becomes the so called Multi-facility Weber Problem (MWP). Both of the MWP and CMWP are NP-hard as shown by Sherali and Nordai [2], and Meggido and Supowit [3], respectively. Moreover, they both generalize the Weber Problem (WP) that aims to find the optimal location of a single facility. The objective function of the WP is convex, and thus, easy to solve. However, this is not true for its generalizations of the MWP and CMWP as additional allocation decisions have to be made for them. One can partition the customer set into non-intersecting subsets each of which is served by an uncapacitated facility for the MWP. This indicates that each customer is served from exactly one facility at the optimal solution of the MWP. On the other hand, customer demands may need to be met from multiple facilities for the CMWP when capacity restrictions are imposed. Nevertheless, all of the WP, MWP and CMWP assume that each facility can be freely located in the plane. Such an assumption can be misleading in practice since there may exist physical areas that obstruct to travel and to open facilities inside. For example, lakes, mountains, glaciers and forests are natural barrier areas which can prohibit both travelling and facility location inside. Indeed, traveling (or passage) is frequently blocked by existing work-shop (office room) areas in a manufacturing plant (in an office building). This work focuses on an extension of the CMWP where the existence of convex polyhedral barriers

*Corresponding Author

are considered as obstacles to pass inside and to locate facilities. Namely, the CMWP with convex polyhedral barriers (CMWP-B) is addressed. We suggest several discretization based heuristic approaches to efficiently solve the CMWP-B.

Mostly, the transportation cost between a facility and customer is a function of the distance and the amount of shipment between them. In particular, the distance is usually modelled with the Euclidean, squared Euclidean, rectilinear and ℓ_r (with $1 \leq r < \infty$) norms in location-allocation type problems such as the CMWP. Alpaydın et al. [4] and Brimberg et al. [5] provide excellent surveys that include techniques and functions to model distances. In this work, the distances between facilities and customers are measured by the so called "barrier distance" which uses the Euclidean distance, i.e. ℓ_2, as its underlying distance measure. Observe that the Euclidean distance may not be feasible when polyhedral barriers exist between facilities and customers. Therefore, "barrier distance" is determined as the shortest path, which does not pass through the barrier regions where the edge costs are calculated with the Euclidean distance, between a facility and a customer. Clearly, the "barrier distances" require extra efforts to calculate the distances which makes the CMWP-B a more realistic and difficult problem to solve than the unrestricted CMWP. Nonetheless, the "barrier distance" still preserves the metric properties [6].

This work has the following contributions. For all we know, the CMWP-B is novel to the literature and we suggest a mathematical programming formulation for it. Next, we offer three discretization based (DB) heuristics which reduce the continuous location space into a discrete space using a discretization strategy. The first DB heuristic employs the solution of an approximating mixed-integer linear programming (MILP) formulation suggested for the CMWP-B. The second DB heuristic applies a Lagrangean Relaxation (LR) scheme on the suggested MILP formulation. The third DB heuristic performs a Tabu Search (TS) using neighborhoods defined over the discretized location space. All upper bounds are calculated with alternate location-allocation type heuristics for the DB heuristics. We generate new test instances for the CMWP-B where they inherit the data of the standard CMWP instances. For that purpose, convex polyhedral barriers are randomly constructed such that the feasibility of the CMWP-B instances is maintained. Lastly, we perform extensive computational experiments on randomly generated test instances derived from the standard CMWP instances.

The remainder of this study is organized as follows. In Section 2 a brief review of the relevant literature is presented. This is followed by the mathematical programming formulation of the CMWP-B in Section 3. Section 4 presents details of the suggested DB heuristics. Section 5 is where we discuss our computational findings. Lastly, the paper is concluded and future research directions are given in Section 6.

2. Literature Review

In this section, we present a short review of the relevant literature. Both the MWP and CMWP have attracted the attention of researchers starting with the seminal works by Cooper [1,7]. There exist several exact solution procedures suggested for the MWP ([8–11]) as well as heuristic approaches for the MWP ([7,12–18]). Furthermore, we can cite the works by Cooper [1], Sherali et al. [19] and Akyüz et al. [20] as examples of exact solution methods developed for the CMWP. On the other hand, we can mention Cooper's [7] Alternate Location-Allocation (ALA) type heuristics [21,22], Discrete Approximation (DA) heuristics [23,24] and metaheuristics [25] developed for the CMWP. Akyüz et al. [26,27] also offer heuristic procedures for a multi-commodity extension of the CMWP. A survey on the location-allocation type problems can be found in the work by Brimberg et al. [28].

As an extension of the single facility WP, the WP-B arises when there exist barriers which prohibit to travel inside and to locate a facility. Katz and Cooper [29] introduce the WP-B having a single circular barrier region. The WP-B with rectilinear distances is considered in the works by Larson and Sadiq [30] and Batta et al. [31]. Aneja and Parlar [32] design an algorithm for the WP-B as well as for a variant of the WP-B where travelling is permitted within the barriers. Butt and Cavalier [33] address the WP-B where barriers are convex polygons and propose an algorithm which yields local optimum solution. Klamroth [34] derives a reduction result for the WP-B having convex polyhedral barriers. Then, an exact and a heuristic procedure is developed for the WP-B. McGarvey and Cavalier [35] develop a branch and bound approach that partitions the continuous location space to optimally solve the WP-B in the presence of polyhedral barriers. Bischoff and Klamroth [36] deal with the WP-B where the barriers are convex polyhedral sets. Their method benefits from the reduction result by Klamroth [34] and decomposes the WP-Bs into multiple subproblems (i.e. WPs) each can be solved over a convex and bounded region. The authors [36] apply a genetic algorithm to decrease the number of subproblems to solve.

The MWP with convex polyhedral barriers (MWP-B) is first considered by Krau [10] which applies a column generation approach based on partitioning of the customer set. Bischoff et al. [37] suggest two ALA type heuristics that have similar allocation phases for the MWP-B. Their first ALA type heuristic alternately solves multiple WP-Bs and set partitioning problems resulting in an inefficient algorithm. The second ALA type heuristic suggested by Bischoff et al. [37] yields better results. In the location phase, the reduction result of [34] is extended to the multi-facility case. Unlike their first ALA type heuristic, this results in solving multiple WPs over convex restricted regions and increase the efficiency of the location phase for the second ALA type heuristic. In what follows, we give a formal definition of the CMWP-B.

3. Capacitated Multi-Facility Weber Problem with Convex Polyhedral Barriers

Let I, J and P denote the number of facilities, the number of customers and the number of polyhedral barriers, respectively. The coordinates of customer j is shown by $\mathbf{a}_j = (a_{j1}, a_{j2})^T$ and its demand is denoted by q_j. The parameter s_i stands for the capacity of facility i. There are two decisions to be made for the CMWP-B: location decisions and allocation decisions. Then, the unknown location of facility i is represented as $\mathbf{x}_i = (x_{i1}, x_{i2})^T$. The unknown amount of flow between facility i and customer j is denoted by f_{ij}. c_{ij} is the cost of transportation per unit flow between customer j and facility i per unit distance. A polyhedral barrier p is indicated with set $\mathcal{B}_p$ and the union of barriers are stated with the set $\mathcal{B}$, i.e. $\mathcal{B} = \bigcup_{p=1}^{P} \mathcal{B}_p$. Now, the feasible region to locate facilities is given with the set $\mathcal{X}$ that can be defined as $\mathcal{X} = \mathbb{E}^2 \setminus \mathcal{B}$ where $\mathbb{E}^2$ is the two dimensional Euclidean space. The Euclidean norm, i.e. $||\mathbf{x}_i - \mathbf{a}_j||_2 = \left[(x_{i1} - a_{j1})^2 + (x_{i2} - a_{j2})^2\right]^{1/2}$, is used as the underlying distance measure to calculate the "barrier distance" between facility i and customer j. Note that barrier distance is represented with $d_{\mathcal{B}}(\mathbf{x}_i, \mathbf{a}_j)$ between facility i and customer j. The details about its calculation will be discussed later on. Now a mathematical programming formulation of the CMWP-B is given as follows.

CMWP-B:

$$\min \quad Z = \sum_{i=1}^{I}\sum_{j=1}^{J} c_{ij} f_{ij} d_{\mathcal{B}}(\mathbf{x}_i, \mathbf{a}_j) \tag{1}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} f_{ij} = s_i \qquad i = 1, \ldots, I, \tag{2}$$

$$\sum_{i=1}^{I} f_{ij} = q_j \qquad j = 1, \ldots, J, \tag{3}$$

$$f_{ij} \geq 0 \quad i = 1, \ldots, I; j = 1, \ldots, J, \tag{4}$$

$$\mathbf{x}_i \in \mathcal{X} \qquad i = 1, \ldots, I. \tag{5}$$

The objective function (1) is to minimize the sum of total transportation cost between facilities and customers. Notice that constraints (2)-(4) are the constraints of the well-known Transportation Problem (TP). Constraints (2) state that the total flow sent from a facility i is equal to its capacity s_i. Constraints (3) imply that the demand q_j of customer j is exactly met. Constraints (4) ensure the nonnegativity of flows. Constraints (5) guarantee that facilities are placed within the feasible region $\mathcal{X}$. It is possible to further restrict the set of feasible region $\mathcal{X}$. Wendell and Hurter [38] show that an optimal solution of the WP can be found within the convex hull of customers. The result by Wendell and Hurter [38] can be directly generalized to the WP-B as well as to the CMWP-B in the case where barrier regions stay within the convex hull of customer locations. In this case, it is enough to consider the region within convex hull of customers that is outside of the polyhedral barriers. However, unlike the CMWP, the resulting feasible region to locate facilities is non-convex for the CMWP-B. Moreover, distances between facilities and customers are measured with barrier distances $d_{\mathcal{B}}(\mathbf{x}_i, \mathbf{a}_j)$ for the CMWP-B. Observe that, the CMWP employs the Euclidean distance between each facility and customer. However, barrier distance is calculated as the shortest path distance between any two points when there exist barriers. The CMWP-B can be reduced to the CMWP when all barriers are removed. The optimal objective value of the CMWP is a lower bound on the optimal objective value of the CMWP-B since barrier distance between any two points is larger than or equal to their Euclidean distances. That is to say, the CMWP-B is more difficult to solve than the CMWP which is NP-hard. This motivates the implementation of efficient heuristic methods for the CMWP-B.

The calculation of the barrier distance $d_{\mathcal{B}}(.)$ requires additional efforts. Notice that convex polyhedral barriers can be represented with their extreme points. Let $\mathbf{e}_{pb} = (e_{pb1}, e_{pb2})^T$ be the coordinates of their extreme points for each polyhedral barrier $p = 1, \ldots, P$ with corresponding extreme points $b = 1, \ldots, B_p$. Here, B_p gives the number of extreme points defining the polyhedral barrier p represented with the set $\mathcal{B}_p$. It should be emphasized that considering only convex polyhedral barriers does not cause a loss of generality. McGarvey and Cavalier [35] have shown that optimal solution of the WP-B can be found outside of the convex hull of barriers which does not contain a customer location. As the convex hull of any barrier (convex or non-convex) can be approximated with a convex polyhedron, barriers are assumed to be convex polyhedrons. The barrier distance between two points $\mathbf{a}$ and $\mathbf{e}$ is defined as the distance of the shortest feasible path from $\mathbf{a}$ to $\mathbf{e}$. To ensure that a path is feasible, it should not coincide with the interior of polyhedral barriers. The concept of visibility graph, e.g. Ghosh [39], is used to ensure that a path is feasible. Two points are said to be visible to each other if and only if they can be connected with a line which does not pass through the interior of polyhedral barriers. Nodes of the visibility graph are customer locations and the extreme points of polyhedral barriers. An edge is defined in the visibility graph if and only if two points, say customer $\mathbf{a}$ and extreme point $\mathbf{e}$ are visible to each other. Edge costs are calculated with the Euclidean distance $d(\mathbf{a}, \mathbf{e}) = \left[(a_1 - e_1)^2 + (a_2 - e_2)^2\right]^{1/2}$. Clearly, when two nodes are not visible to each other, there is no corresponding edge in the visibility graph. Finally, the barrier distance between two points (nodes) $d_{\mathcal{B}}(\mathbf{a}, \mathbf{e})$ is equal to the cost of the shortest path from $\mathbf{a}$ to $\mathbf{e}$ on the visibility graph. Shortest paths can be obtained by a shortest path algorithm such as the Dijkstra's algorithm [40].

The discussion above only considers the barrier distances among the nodes existing in the visibility graph. For an arbitrary point $\mathbf{x}$ within the set of feasible locations, i.e. $\mathbf{x} \in \mathcal{X}$, the calculation of barrier distances is slightly different. One may consider re-constructing the visibility graph from scratch such that $\mathbf{x}$ is also added into the node set. Fortunately, it is sufficient to determine only the set of nodes visible from $\mathbf{x}$, say $\mathcal{V}_{\mathbf{x}}$, in the visibility graph. Then, the barrier distances from the point $\mathbf{x}$ to any customer location $\mathbf{a}$ is measured using the formula:

$$d_{\mathcal{B}}(\mathbf{x}, \mathbf{a}) = \min_{\mathbf{h} \in \mathcal{V}_{\mathbf{x}}} \{d(\mathbf{x}, \mathbf{h}) + SPD(\mathbf{h}, \mathbf{a})\}, \qquad (6)$$

where $\mathbf{h}$ stands for the customer locations and/or extreme points of barriers that are visible from the point $\mathbf{x}$. $SPD(\mathbf{h}, \mathbf{a})$ represents the shortest path distance between points $\mathbf{h} \in \mathcal{V}_{\mathbf{x}}$ and $\mathbf{a}$ in the visibility graph. The barrier distance function $d_{\mathcal{B}}(\mathbf{x}, \mathbf{a})$ defines a metric on $\mathcal{X}$ satisfying the following properties of a metric: positivity, definiteness, symmetry and triangle inequality. Klamroth [6] is an excellent reference to resort for more details on the single facility location problems with barriers and their properties. In the left-hand side of Figure 1, an example with four customers and single tetragon barrier is given. In the right-hand side of Figure 1, the corresponding visibility graph is illustrated.

4. Discretization Based (DB) Heuristics

The CMWP-B reduces to solving I WP-B's when the amount of shipments are known. This results in a two dimensional minimum location problem in continuous space. When all barriers are removed, the resulting problem is a classical WP and using the results by Wendell and Hurter [38] and Hansen et al. [41], an optimal solution can be found within the convex hull of customer locations. Clearly, the single WP is easy to solve because it is a convex programming problem. That is to say, it can be solved by the Weiszfeld's [42] algorithm or one of its generalizations ([43, 44]). On the other hand, the objective function of the CMWP-B is non-convex and the problem itself is not easy to solve on continuous space. This sparks a discretization strategy to solve the CMWP-B using a discretized location space. It is conceivable that the CMWP-B can be transformed into a MILP problem formulation when facility locations are chosen from a set of candidate locations. Indeed, the resulting MILP problem formulation may yield the optimal solution of the CMWP-B when the set of candidate locations to place facilities includes the optimal facility locations. Obviously, it is not possible to determine the optimal facility locations in advance when constructing the set of candidate facility locations. However, solving a MILP problem considering a set of candidate facility locations produces an approximate solution for the CMWP-B. For that purpose, a systematic way of choosing candidate facility locations is of high importance and may provide good solutions for the CMWP-B. In fact, such a strategy is formerly offered by Hansen et al. [14] and Aras et al. [23] for the MWP and CMWP, respectively. They use customer locations as the set of candidate facility locations and

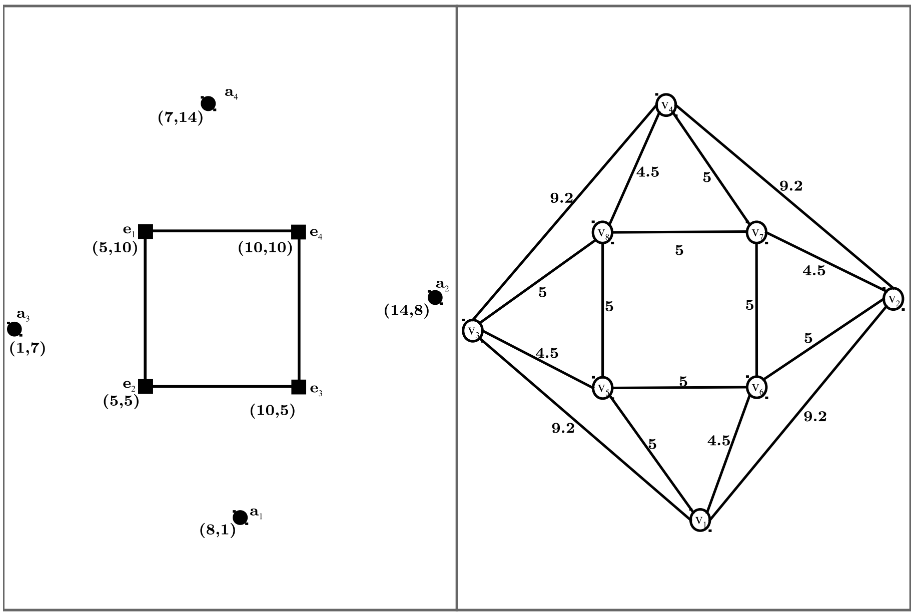


Figure 1. An illustrative example with 4 customers, a single tetragon barrier and the corresponding visibility graph.

solve approximating MILPs. Hansen et al. [14] solve an approximating p-median problem and obtain highly accurate solutions for the MWP. Similar results also hold for the CMWP from the work by Aras et al. [23]. The discretization strategy that we pursue in this study is analogous to the previous ones [14,23]. In what follows, we first present the approximating MILP problem formulation and the first DB (DB-I) heuristic. This part also elaborates how we tackle with the difficulties imposed by the barrier distances for the single WP-B. Then, the efficiency of DB-I heuristic is improved by a LR scheme. This heuristic is denominated as DB-II. Lastly, a Tabu Search (TS) algorithm is employed to further improve the efficiency of the DA-I heuristic. This heuristic is called as DB-III heuristic.

4.1. Approximating MILP Problem Formulation: DB-I Heuristic

Let $k = 1, \dots, K$ denote candidate facility locations with known coordinates given as $\mathbf{v}_k = (v_{k1}, v_{k2})^T$. The decision variables w_{ijk} represent the amount of flow between facility i located at candidate point k and customer j. Binary variables u_{ik} take a value of 1 if and only if facility i is opened at candidate point k and 0 otherwise. c_{ijk} is the corresponding transportation cost for flow w_{ijk}. Specifically, it is calculated as $c_{ijk} = c_{ij} d_{\mathcal{B}}(\mathbf{v}_k, \mathbf{a}_j)$. Now, an approximating MILP problem formulation of the CMWP-B can be stated as follows.

DA:

$$\min \quad Z_{DA} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk} w_{ijk} \tag{7}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} w_{ijk} = s_i u_{ik} \quad i = 1, \dots, I; \; k = 1, \dots, K, \tag{8}$$

$$\sum_{i=1}^{I}\sum_{k=1}^{K} w_{ijk} = q_j \quad j = 1, \dots, J, \tag{9}$$

$$\sum_{k=1}^{K} u_{ik} = 1 \quad i = 1, \dots, I, \tag{10}$$

$$w_{ijk} \geq 0 \quad i = 1, \dots, I; j = 1, \dots, J; \; k = 1, \dots, K, \tag{11}$$

$$u_{ik} \in \{0, 1\} \quad i = 1, \dots, I; \; k = 1, \dots, K. \tag{12}$$

Here, constraints (8), (9) and (11) are analogous to the TP constraints (2)-(4) for the approximating MILP problem formulation of the CMWP-B. Constraints (10) ensure that every facility is opened at exactly one candidate facility location.

Before giving the details of the DB-I heuristic, we first introduce an improvement heuristic that is employed within the DB heuristics in the following. This improvement heuristic is an adaptation of the famous Alternate Location-Allocation (ALA) heuristic that is offered for the MWP by Cooper [7]. It consists of alternately solving two subproblems: allocation and location subproblems. These subproblems respectively arise when the facility locations and allocations between facility and customers are fixed. ALA heuristic repeats alternating until the objective function value does not change from one iteration to another. This also indicates that facility locations and allocation values do not significantly change and the objective value becomes stable. It can be observed that initialized from given facility locations, the CMWP-B reduces to solving a TP, which is the "allocation" subproblem, in order to determine allocations, i.e. flow between facilities and customers. TP can be solved straightforwardly using linear programming solvers. On the other hand, with a given feasible shipment plan, the CMWP-B can be decomposed into I single facility WP-Bs so called the "location" subproblems. Solving the resulting location subproblems, i.e. the WP-Bs, is not trivial and our approach for solving the location subproblems is summarized in the following. Recall that the WP-B is a non-convex problem and several solution approaches can be used to solve ([32, 33, 35]). Instead of an exact solution procedure, which can be prohibitive to use within an efficient heuristic approach, we prefer to make use of the Weiszfeld's [42] procedure that solves the unrestricted WP optimally. Weiszfeld's algorithm employs the gradient direction to update facility locations from one iteration to the other and eventually converges to the optimal facility locations. This procedure works well on a convex problem like the WP. However, there is no guarantee of optimality and it cannot be directly used for the non-convex WP-B. Nevertheless, observe that, when Weiszfeld's algorithm ends up with a solution $\mathbf{x}^*$ that is within the feasible location space outside of the polyhedral barriers, i.e. $\mathbf{x}^* \in \mathcal{X}$, then it is also optimal for the WP-B. Otherwise, when Weiszfeld's algorithm terminates with a solution $\mathbf{x}^*$ that is within a polyhedral barrier p, then, the facility location can be approximated by choosing the closest point $\hat{\mathbf{x}}$ on the border of the barrier p. In other words, our solution approach finds a solution for the location subproblems that yields approximate (or hopefully optimal) facility locations. This procedure is illustrated with Figure 2. This improvement heuristic solves the location and allocation subproblems as described until the objective value remains the same from one iteration to another. This heuristic is named as ALA with barriers (ALAB) heuristic. ALAB heuristic is frequently resorted within our DB heuristics.

The DB-I heuristic works as follows. First, the DA formulation is solved using the set of candidate facility locations consisting of customer locations. The DA formulation yields the facility locations which are used to initialize the improvement heuristic. Second, the ALAB heuristic is used to enhance the solution initialized with the facility locations obtained in the first phase. Then, the best solution found is reported as the outcome of the DB-I heuristic. Note that applying the ALAB heuristic in the second phase of the DB-I heuristic usually improves the initial solution obtained from the DA formulation.

4.2. A Lagrangean Relaxation (LR) Scheme: DB-II Heuristic

The DA formulation can be intractable for large problems and thus computationally very expensive to solve exactly. Therefore, it may be better to use an approximate solution approach. To this end, a LR scheme and subgradient optimization is employed to compute good feasible solutions for the DA. The demand constraints (9) are relaxed with Lagrangean multipliers μ_j to obtain the Lagrangean subproblem of DA, namely the LDA.

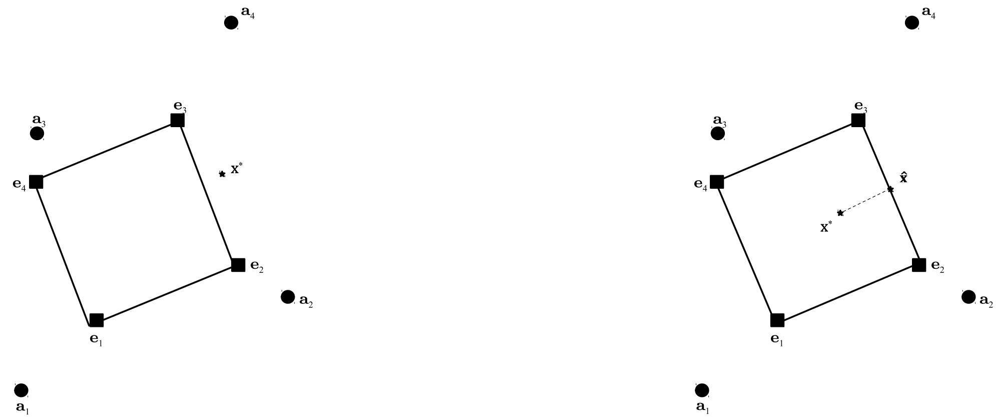
LDA($\boldsymbol{\mu}$):

$$\min Z_{LDA}(\boldsymbol{\mu}) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(c_{ijk} - \mu_j)w_{ijk} + \sum_{j=1}^{J}\mu_j q_j \tag{13}$$

$$\text{s.t.} \quad (8), (10), (11), (12),$$

$$w_{ijk} \leq \min\{s_i, q_j\} \quad i = 1, \dots, I; \; j = 1, \dots, J; k = 1, \dots, K. \tag{14}$$

Observe that constraints (14) are introduced as basic upper bounds on the flow quantities to the LDA formulation. Clearly, these constraints are redundant for the DA formulation. However, they may significantly improve the optimal objective value, say $Z^*_{LDA}(\boldsymbol{\mu})$, of the Lagrangean subproblem LDA. The last term of the objective function (13) is constant and LDA($\boldsymbol{\mu}$) can be further decomposed over the facilities. That is to say, solving the following subproblems for each facility $i = 1, \dots, I$ is equivalent to solving the Lagrangean subproblem LDA($\boldsymbol{\mu}$)



(a) Case-I: x^* is outside the barrier region **(b)** Case-II: x^* is inside the barrier region

Figure 2. Two possible cases for the output of a Weiszfeld-like procedure.

$\text{LDA}_i(\boldsymbol{\mu})$:

$$\min Z_{LDA_i}(\boldsymbol{\mu}) = \sum_{j=1}^{J}\sum_{k=1}^{K}\bar{c}_{ijk}w_{ijk} \tag{15}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} w_{ijk} = s_i u_{ik} \quad k = 1,\ldots,K, \tag{16}$$

$$\sum_{k=1}^{K} u_{ik} = 1, \tag{17}$$

$$w_{ijk} \le \min\{s_i, q_j\} \quad j = 1,\ldots,J; \quad k = 1,\ldots,K, \tag{18}$$

$$w_{ijk} \ge 0 \quad j = 1,\ldots,J; \quad k = 1,\ldots,K, \tag{19}$$

$$u_{ik} \in \{0,1\} \quad k = 1,\ldots,K. \tag{20}$$

where the unit costs $\bar{c}_{ijk}$ are determined as $\bar{c}_{ijk} = (c_{ijk} - \mu_j)$ with a given Lagrange multiplier vector $\boldsymbol{\mu}$. The resulting subproblem $\text{LDA}_i(\boldsymbol{\mu})$ can be solved by an inspection procedure for each candidate facility location k. Note that, when facility i is placed at a candidate location k, then $\text{LDA}_i(\boldsymbol{\mu})$ further reduces to the following problem $\text{LDA}_{ik}(\boldsymbol{\mu})$:

$$\min Z_{LDA_{ik}}(\boldsymbol{\mu}) = \sum_{j=1}^{J}\bar{c}_{ijk}w_{ijk} \tag{21}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} w_{ijk} = s_i \tag{22}$$

$$w_{ijk} \le \min\{s_i, q_j\} \quad j = 1,\ldots,J, \tag{23}$$

$$w_{ijk} \ge 0 \quad j = 1,\ldots,J. \tag{24}$$

$\text{LDA}_{ik}(\boldsymbol{\mu})$ is a continuous bounded knapsack problem that can be optimally solved in polynomial time [45]. This approach requires sorting of the cost coefficients for each candidate location k. Hence, we need to run a sorting procedure K times. The least cost candidate point k^* is chosen as the optimal location of facility i, and thus for the subproblem $\text{LDA}_i(\boldsymbol{\mu})$. Now, the optimal value of $\text{LDA}_i(\boldsymbol{\mu})$ is calculated as $Z^*_{LDA_i}(\boldsymbol{\mu}) = \min_k \left\{Z^*_{LDA_{ik}}(\boldsymbol{\mu})\right\}$. It is conceivable that the values of binary variables are set as $u_{ik^*} = 1$ for the least cost candidate location k^* and $u_{ik} = 0$ for $k = 1,\ldots,K$ and $k \neq k^*$. Once all subproblems $\text{LDA}_i(\boldsymbol{\mu})$ are solved, for a given Lagrange multiplier vector $\boldsymbol{\mu}$, the optimal value of the Lagrangean subproblem $\text{LDA}(\boldsymbol{\mu})$ is determined as $Z^*_{LDA}(\boldsymbol{\mu}) = \sum_{i=1}^{I} Z^*_{LDA_i}(\boldsymbol{\mu}) + \sum_{j=1}^{J}\mu_j q_j$. Notice that $Z^*_{LDA}(\boldsymbol{\mu})$ constitutes a lower bound on the optimal value of the DA for any Lagrange multiplier vector $\boldsymbol{\mu}$. The best lower bound can be obtained by solving the Lagrangean dual problem $\max_{\boldsymbol{\mu}}\{Z^*_{LDA}(\boldsymbol{\mu})\}$ that is accomplished using subgradient algorithm by Held et al. [46]. For the sake of brevity, we do not give details of the subgradient algorithm. The subgradient algorithm calculates upper bounds during its run, and hence, feasible solutions for the CMWP-B. For that purpose, the ALAB heuristic is used. The values of binary variables u_{ik} are used to determine the locations of each facility for each solution of the resulting Lagrangean subproblem $\text{LDA}(\boldsymbol{\mu})$ with given multiplier values $\boldsymbol{\mu}$. Lastly, the DB-II heuristic reports the best feasible solution obtained from the subgradient algorithm.

4.3. A Tabu Search Algorithm: DB-III Heuristic

Inspired with the promising results obtained by using the customer locations as the set of candidate facility location, the DB-III heuristic consists of applying a TS algorithm on the set of candidate facility locations. A simple neighborhood search structure can be defined by exchanging the location of a facility i' at a candidate location k' with another one in order to move from one feasible solution to another one. This yields a feasible solution for the CMWP-B. Further, feasible solutions can be improved by applying a ALAB heuristic.

The suggested TS algorithm basically exchanges the location of a facility with another candidate location at each iteration. The selected candidate location is declared as tabu for that facility and its status can not be revoked during the tabu tenure. A greedy strategy of selecting the closest candidate location is applied to move from the current feasible solution to a neighbor solution. To increase the intensity of the neighbor search the closest N neighbor feasible solutions can be checked and the best feasible solution is picked as the new feasible solution. The TS algorithm searches over the candidate location set and reports the best solution found. For further details on TS we refer to the work by Glover and Laguna [47].

The TS algorithm uses a tabu list $\mathcal{T}$ that keeps record of candidate facility locations which are declared as tabu for a duration (tabu tenure) say α iterations for each facility i. Each facility i is associated with a set of candidate locations which are declared as tabu. A tabu declared candidate location k for a facility i can not be in the current solution until the tabu tenure record, represented as $T(i,k)$, decreases to zero. Let t denote the current iteration number and θ stand for the maximum number of iterations completed by the TS, namely, the iteration limit of the TS. At each iteration, a facility i^* is chosen and a non-tabu candidate location k^* with $T(i^*,k^*) = 0$ is declared as tabu for i^*. The facility i^* is determined with respect to their facility index in the order from smallest facility index to largest facility index. In other words, if facility i is selected at iteration t then facility $i+1$ is the next facility to be selected at iteration $t+1$. Note that, when $i^* = I$ for the selected facility, the facility to be declared as tabu is the one with index number 1 at the next tabu iteration. Once tabu facility i^* is set at iteration t, the candidate point k^* is searched over the closest neighbors of the current candidate location, say k', of facility i^*. Clearly, all facilities other than i^* maintain their facility locations determined at previous iteration $t-1$ and the new feasible solution only changes the location of facility i^* at iteration t. Let $k'_{(1)}$ be the closest candidate location to k', $k'_{(2)}$ denote the second closest candidate location to k', and so forth. The upper bound obtained at tabu iteration t and the best upper bound found so far are respectively represented with Z^t_{UB} and Z^{best}_{Tabu}. The upper bound Z^t_{UB} is determined as the lowest upper bound value obtained by exchanging k' with the candidate locations of the set $\{k'_{(1)}, k'_{(2)}, \ldots, k'_{(N)}\}$ one by one from the closest $(k'_{(1)})$ to farthest $(k'_{(N)})$ neighbor candidate location. Here, N is the number of neighbor solutions checked (or the width of the neighborhood) and $k'_{(N)}$ is the N^{th} closest candidate location to k'. This strategy is followed as long as Z^t_{UB} does not improve the best upper bound Z^{best}_{Tabu}. When an improvement is obtained, the neighborhood search is stopped. The best upper bound Z^{best}_{Tabu} is updated whenever the improvement occurs and the TS proceeds to the next iteration $t+1$ with the next facility. For example, when the feasible solution obtained by exchanging k' with the closest candidate location $k'_{(1)}$ yields a better upper bound than the best upper bound Z^{best}_{Tabu}, then Z^t_{UB} is calculated with candidate location $k'_{(1)}$ for facility i^* and Z^{best}_{Tabu} is updated. In addition, the current tabu solution is updated so that facility i^* is located on the candidate location $k'_{(1)}$. Similarly, such an approach is followed until N^{th} closest candidate location whenever an improvement is achieved. If there is no improvement for the best upper bound Z^{best}_{Tabu} after the exchange of N^{th} closest candidate location $k'_{(N)}$, then facility i^* is located at the candidate location which gives the lowest objective value among these N neighbor candidate locations.

Selection of the candidate location k^* of facility i^* to move to another neighbor solution requires further attention. The tabu status of candidate locations terminates after a while and it is likely that the same candidate location is selected repetitively. To avoid such a case, a tabu frequency list (FL) keeps record of the total number of times a candidate point k is declared tabu for a facility i within the TS algorithm. $FL(i,k)$ stands for the FL, and initially, $FL(i,k) = 0$ for all $i = 1, \ldots, I$; $k = 1, \ldots, K$. When a candidate location k for facility i^* is tested within the neighborhood search described at an iteration, $FL(i^*,k)$ is increased by one. The candidate location k which has the highest $FL(i^*,k)$ value is excluded from the neighborhood search in the next iteration. Then the set of neighbor candidate locations does not contain a tabu declared candidate location or the most frequent candidate location for i^*. Therefore, this strategy favors diversification of the solutions during the TS. The upper bound Z^t_{UB} obtained at iteration t is associated with a solution vector $(k^1, k^2, \ldots, k^I)$ which represents the candidate locations of each facility in the solution. Here, k^1

Algorithm 1: Tabu Search Algorithm

Step 1. *(Initialization):* Find initial upper bound Z^0_{UB} and its associated solution vector of candidate locations $(k^1, k^2, \dots, k^I)^0$. Set $Z^{best}_{Tabu} = Z^0_{UB}$ and iteration counter $t = 1$. Set $FL(i,k) = 0$ for $i = 1, \dots, I; k = 1, \dots, K$. Set $T(i, k^i) = \alpha$ and $FL(i, k^i) = 1$ for each facility $i = 1, \dots, I$.

Step 2. For facility $i = 1, \dots, I$,

(i) determine the neighborhood set $\mathcal{K}^i$ as $\mathcal{K}^i = \{k^i_{(1)}, k^i_{(2)}, \dots k^i_{(N)}\}$ so that $T(i, k^i_{(n)}) = 0$ for $n = 1, \dots, N$ and $FL(i, k^i_{(n)}) < \max\{FL(i,k)\}$ for $n = 1, \dots, N; k = 1, \dots, K$.

(ii) for neighbor $n = 1, \dots, N$, set $(k^1, k^2, \dots, k^i, \dots, k^I)^t = (k^1, k^2, \dots, k^i = k^i_{(n)}, \dots, k^I)^{(t-1)}$ and find its associated objective value as $Z^{k^i_{(n)}}_{UB}$.

(iii) if $Z^{k^i_{(n)}}_{UB} < Z^{best}_{Tabu}$ then update $Z^{best}_{Tabu} = Z^{k^i_{(n)}}_{UB}$, $Z^t_{UB} = Z^{k^i_{(n)}}_{UB}$. Set $FL(i, k^i_{(n)}) = FL(i, k^i_{(n)}) + 1$ and $T(i, k^i_{(n)}) = \alpha$. Go to Step 4.

Step 3. If Z^{best}_{Tabu} does not improve, find k^* such that $k^* = \underset{k \in \mathcal{K}^i}{\operatorname{argmin}}\{Z^{k^i_{(n)}}_{UB}\}$, $(k^1, k^2, \dots, k^i, \dots, k^I)^t = (k^1, k^2, \dots, k^i = k^*, \dots, k^I)^{(t-1)}$. Set $FL(i, k^i_{(n)}) = FL(i, k^i_{(n)}) + 1$ for $n = 1, \dots, N$ and $T(i, k^*) = \alpha$.

Step 4. Set $t = t + 1$ and decrease each tabu tenure value $T(i, k^i) > 0$ by one. If $t = \theta$ or Z^{best}_{Tabu} does not improve for 30 consecutive iterations STOP and report Z^{best}_{Tabu}, otherwise go to Step 2.

is the candidate location of the first facility, k^2 is the candidate location of the second facility and so forth, in the solution at iteration t. A formal outline of the suggested TS algorithm is given in Algorithm 1.

5. Computational Experiments

In this section, first the test bed used in this work is given. Second, the results obtained with our DB heuristics are presented for the CMWP-B. A Dell Precision T5810 workstation with Intel(R) Xeon(R) E5-1650v3 processor of 3.50 GHz and 64 GB RAM operating within Microsoft Windows 7 Pro 64-bit environment is employed as our computing platform. The callable library of Gurobi 5.6.3 with default settings is used to solve MILP and LP formulations presented and all codes are written in C^{++} programming language.

5.1. Test Bed

We have performed our computational experiments on randomly generated test instances that are produced using standard CMWP test instances from the literature. Standard test instances are taken from the works by Sherali et al. [19] and Boyacı [48]. Recall that the CMWP instances do not have barriers forbidding location and travel. However, the CMWP-B can inherit facility capacities, customers' demand and location information from the CMWP instances. We have generated convex polyhedral barriers for each CMWP instance by considering the following two issues. First, we ensure that barrier regions do not contain any customer location inside. Second, interior areas of barriers should not be coinciding with others. Then, these barriers are integrated with the data of the CMWP instances resulting in CMWP-B test instances.

The optimal objective value of the CMWP constitutes a lower bound on the CMWP-B. We have employed the best known objective values of standard CMWP test instances as benchmark values to make a comparison among the performance of the suggested DB heuristics. These benchmark values representing the best known objective values of the CMWP instances are reported in Table 1. The first column indicates the names of the instances. Original instance numbers are used as stated in the work by Sherali et al. [19]. We add a prefix "S" before the corresponding instance number. For example, the CMWP instance number 10 in the work by Sherali et al. [19] is shown as S10. 11 CMWP test instances from Sherali et al. [19] are considered. Analogously, a total of 40 CMWP test instances from Boyacı [48], which have non-unique cost values, are represented with a prefix "B" followed by an instance number. The second column stands for the size of the instance by giving the number of facilities and the number of customers in parenthesis, respectively. The instances which have the same number of facilities and customers are distinguished by adding a suffix letter starting from "a" to "e" in accordance with the original denomination presented in the reference works. The last column includes the best known objective values that are taken from the reference works ([19, 27, 48]. As a remark, these benchmark values are not necessarily optimal for the CMWP as the latter two studies present heuristic outcomes.

To construct barriers for each standard CMWP test instance given, we have randomly generated different number of extreme points and convex polygons. P polygons are chosen from the set $P \in \{1, 3, 5, 10\}$ for the CMWP-B test instances. When $P = 1$, that is a single polygon, the total number of extreme points, denoted as B, of the polygon is determined within the set $\mathcal{H} = \{2, 3, 4, 5, 6, 7, 8, 9\}$. For $P > 1$, when there are more than one polygon, B is calculated by multiplying the elements of $\mathcal{H}$ with the number of polygons P. For example, for $P = 3$, B is chosen such that $B \in \mathcal{H} = \{6, 9, 12, 15, 18, 21, 24, 27\}$. This makes a total of $4 \times 8 = 32$ combinations of CMWP-B test instances which are generated for each source instance of the CMWP. Further, two different strategies are followed to generate instances that are grouped as regular and random instances. Regular instances contains polygons with the same number of extreme points. For example, an instance is called regular when all barriers are triangles. On the other hand, random instances contains polygons each of which do not necessarily have the same number of extreme points as long as their total does not exceed B. Shortly, each of regular and random instance groups for the CMWP-B contain $4 \times 8 = 32$ test instances. This makes $11 \times 32 \times 2 = 704$ and $40 \times 32 \times 2 = 1920$ CMWP-B test instances that are constructed using existing CMWP test instances Sherali et al. [19] and Boyacı [48] instances, respectively.

Table 1. Benchmark values for the standard CMWP test instances from the literature.

Instance Name	Size (I, J)	Benchmark Value
S6	(3,9a)	221.40
S7	(3,9b)	871.62
S8	(4,8)	609.23
S9	(5,15)	8169.79
S10	(5,20a)	12846.87
S11	(5,20b)	1107.18
S12	(5,30)	23990.04
S15	(5,10)	2595.47
S16	(6,10)	7797.21
S18	(8,10)	1564.46
S20	(10,10)	7719.00
B1	(5,20)	22146.54[b]
B2	(5,25)	38236.43[b]
B3	(5,30)	42899.92[b]
B4	(5,40)	104983.19[b]
B5	(5,50)	71433.35[b]
B6	(5,100a)	119564.85[b]
B7	(5,250)	891589.16[b]
B8	(6,20)	14693.26[b]
B9	(6,25)	19036.28[b]
B10	(6,30)	59665.36[b]
B11	(7,20)	15075.29[b]
B12	(7,25)	16086.30[b]
B13	(7,30)	33627.21[b]
B14	(8,20)	6499.80[b]
B15	(8,25)	9486.23[b]
B16	(8,30)	18081.73[b]
B17	(9,20)	6551.37[b]
B18	(9,25)	5983.27[b]
B19	(9,30)	10643.15[b]
B20	(10,20)	4702.16[b]
B21	(10,25)	3950.70[b]
B22	(10,30)	6863.79[b]
B23	(10,40)	37834.75[b]
B24	(10,50)	47125.8[b]
B25	(10,100)	171126.36[b]
B26	(20,40)	8240.14[a]
B27	(20,50)	13535.76[b]
B28	(20,250)	220033.88[b]
B29	(25,250a)	208587.12[a]
B30	(25,250b)	278454.58[a]
B31	(25,250c)	259745.76[a]
B32	(25,250d)	195125.60[a]
B33	(25,250e)	289012.15[a]
B34	(25,500a)	646653.93[a]
B35	(25,500b)	512483.15[a]
B36	(25,500c)	484237.33[a]
B37	(25,500d)	643832.22[a]
B38	(25,500e)	491602.33[a]
B39	(30,100)	26185.03[b]
B40	(50,250)	81055.87[b]

[a] Results from [27]
[b] Results from [48]

5.2. Computational Results

We have employed the best known values of the CMWP as our benchmark values as given in Table 1 to compare the performance of our DB heuristics. The percent deviations of the heuristics are determined using the following formula.

$$100 \times \frac{Z_{UB} - Z_{CMWP}^{best}}{Z_{CMWP}^{best}} \tag{25}$$

where Z_{UB} is the upper bound value found by the suggested heuristic for the CMWP-B. Z^{best}_{CMWP} is the best known objective value for standard CMWP test instances as shown in Table 1. Clearly, the formula (25) yields an upper bound on the performance of the heuristics. The instances are divided into two groups as small and large instances depending on their number of facilities and customers. For example, large instances contain $I > 10$ facilities to locate. In addition, instances with $J \geq 100$ customers are also considered as large instances.

The upper bounds are determined with the ALAB heuristic as described in Section 4.1 for our DB heuristics. The DB-I heuristic employs the ALAB heuristic as an improvement step by using the facility locations obtained after the proposed MILP formulation is solved. For the DB-II heuristic, an initial upper bound is needed to increase the efficiency of the subgradient algorithm. Initially, facilities are assumed to have unlimited capacity and they can meet all customer demand. Then, facilities are assigned to the candidate locations from which they serve the customers with least total cost. Once the corresponding candidate locations are fixed as the starting facility locations, the ALAB heuristic is run initialized from those facility locations. Lastly, the resulting upper bound is used as the initial upper bound of the subgradient algorithm for the DB-II heuristic. Besides, the ALAB heuristic is systematically run (i.e. once in ten iterations) to update best upper bound found within the subgradient algorithm. Finally, the ALAB heuristic is applied on the best solution found by the TS algorithm within the DB-III heuristic.

Table 2 gives the outcomes of the DB heuristics on both random and regular CMWP-B test instances generated from the instances by Sherali et al. [19]. The first three columns explain instance properties. The first column gives the name of the standard CMWP instance as mentioned. (I, J) stands for the number of facilities and customers in the second column. The third column states the total number of barrier regions P in the instances. "% Dev." and "CPU" denote the percent deviation calculated by the formula (25) and the CPU time of the corresponding DB heuristic in seconds, respectively. To be precise, the cells under these columns are the average of 8 test instances having different number of total extreme points from the set $\mathcal{H}$ depending on the number of barriers P. We should emphasize that all instances in Table 2 are small instances. The last row indicates the average of each column. Starting with column 4, six columns are consecutively dedicated to the outcomes of the DB heuristics for random and regular test instances. Best percent deviations at each row are marked in bold characters. On these instances, we observed that, DB-I heuristic outperforms both DB-II and DB-III heuristics in terms of accuracy. Indeed, the DB-I heuristic yields outstanding accuracy for both random and regular test instances. Using LR scheme as in the DB-II heuristic increases the efficiency and yields outcomes in almost half time of the DB-I heuristic. Additionally, the DB-II heuristic produces similar accuracy to the DB-I heuristic in 12 out of 44 cases of random instances. This value is 11 out of 44 cases when regular instances are considered.

The tabu tenure parameter α for the DB-III heuristic is set to $\alpha = \max\{I, 20\}$ in the light of our preliminary experiments. A maximum number of tabu iterations θ is set to 300. Another stopping condition of 30 non-improving tabu iterations is also imposed to avoid from unnecessary computations for the DB-III heuristic. A neighborhood search width denoted with N is empirically determined as $N = \min\{\lceil J/3 \rceil, 30\}$. Here, $\lceil a \rceil$ is the smallest integer value that is larger than or equal to a. These TS settings are employed in all of our computational experiments. Unfortunately, the performance of the DB-III heuristic is not promising on these small instances. Nevertheless, running time of the DB-III heuristic is shorter than the DB-I heuristic on the average for small instances.

Although the instances based on Sherali et al. [19] are small instances, their sizes are of limited variety. Therefore, Table 3 shows additional results obtained for random and regular CMWP-B instances (small instances) based on Boyacı [48] instances. Diversity of these instances is higher than the ones by Sherali et al. [19]. Table 3 can be read similar to Table 2 since they share the same structure. The results of Boyacı [48] instances are similar to Sherali et al. [19] instances and strengthen the success of the DB-I heuristic for small instances. The DB-I heuristic is the winner for both random and regular CMWP-B instances over all test instances in terms of accuracy. In particular, the DB-I heuristic yields percent deviations with a difference of 10% more than that of the closest approach, the DB-II heuristic, on the average for small instances. For these small instances, it becomes evident that the efficiency of the DB-I heuristic decreases drastically, i.e. 110.6 and 108.96 seconds on the average for random and regular instances, respectively. The performance

Table 2. The performance of the DB heuristics over random and regular CMWP-B instances based on Sherali et al. [19] instances.

Instance			Random CMWP-B Instances						Regular CMWP-B Instances					
Name	Size		DB-I		DB-II		DB-III		DB-I		DB-II		DB-III	
	$(I,\,J)$	P	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU
S6	(3,9)	1	**0.27**	0.05	**0.27**	0.24	6.64	0.16	**0.51**	0.33	0.51	0.24	5.54	0.13
		3	**0.48**	0.16	**0.48**	0.29	24.44	0.16	**0.36**	0.39	0.37	0.24	17.07	0.13
		5	**0.27**	0.33	0.45	0.29	6.34	0.17	**0.72**	0.56	**0.72**	0.23	13.51	0.16
		10	**0.27**	0.23	0.28	0.35	1.96	0.20	**0.20**	0.22	0.26	0.30	7.03	0.19
S7	(3,9)	1	**0.57**	0.02	0.57	0.05	4.35	0.18	**0.53**	0.02	0.53	0.05	4.14	0.17
		3	**1.89**	0.03	**1.89**	0.06	6.95	0.18	**1.29**	0.02	**1.29**	0.04	6.03	0.20
		5	**0.24**	0.02	**0.24**	0.06	5.06	0.21	**1.26**	0.02	**1.26**	0.05	10.27	0.20
		10	**0.34**	0.03	**0.34**	0.05	4.42	0.20	**1.04**	0.04	1.04	0.05	7.47	0.23
S8	(4,8)	1	**0.34**	0.06	**0.34**	0.30	32.06	0.19	**0.22**	0.41	**0.22**	0.29	32.44	0.18
		3	**0.69**	0.47	**0.69**	0.29	33.77	0.17	**0.09**	0.52	**0.09**	0.28	35.85	0.18
		5	**0.42**	0.08	**0.42**	0.29	30.89	0.20	**0.34**	0.62	**0.34**	0.27	32.42	0.19
		10	**0.33**	0.32	**0.33**	0.32	30.78	0.25	**0.72**	0.42	**0.72**	0.28	29.74	0.23
S9	(5,15)	1	**0.31**	0.25	1.30	0.10	33.97	0.52	**0.25**	0.65	1.21	0.10	42.65	0.44
		3	**0.46**	0.65	1.25	0.10	35.38	0.54	**0.66**	1.01	2.63	0.15	29.94	0.48
		5	**0.20**	0.56	1.24	0.11	36.96	0.46	**0.19**	0.95	1.25	0.11	40.13	0.50
		10	**0.47**	0.74	1.46	0.10	42.03	0.49	**0.54**	1.25	1.47	0.10	44.45	0.48
S10	(5,20)	1	**0.00**	0.37	1.81	0.31	4.16	0.39	**0.05**	0.99	2.34	0.31	3.22	0.44
		3	**0.26**	0.75	1.36	0.31	4.60	0.39	**0.10**	1.29	2.43	0.32	4.32	0.40
		5	**0.69**	0.91	2.16	0.35	4.72	0.39	**0.45**	0.68	1.89	0.31	4.47	0.41
		10	**1.11**	0.69	3.23	0.34	5.30	0.47	**0.86**	0.63	2.96	0.31	5.54	0.45
S11	(5,20)	1	**0.04**	0.51	**0.04**	0.43	51.89	0.68	**0.05**	0.89	**0.05**	0.50	60.21	0.79
		3	**0.05**	0.95	0.83	0.47	56.74	0.70	**0.43**	0.90	**0.43**	0.50	63.36	0.80
		5	**2.00**	1.18	**2.00**	0.47	52.39	0.75	**1.34**	0.77	**1.34**	0.52	67.39	0.71
		10	**0.15**	0.72	**0.15**	0.49	53.14	1.08	**0.40**	0.90	**0.40**	0.52	75.35	0.77
S12	(5,30)	1	**0.09**	1.84	5.46	0.48	19.49	0.88	**0.20**	2.49	4.83	0.60	20.60	0.90
		3	**0.16**	2.09	5.90	0.50	17.59	0.96	**0.07**	2.43	5.16	0.56	18.85	0.84
		5	**0.02**	3.68	5.50	0.54	20.28	0.95	**0.08**	2.52	5.11	0.55	17.63	0.90
		10	**0.11**	2.13	6.72	0.55	14.22	1.04	**0.28**	2.59	4.87	0.58	20.82	1.00
S15	(5,10)	1	**0.00**	0.10	2.32	0.36	123.50	0.29	**0.00**	0.25	2.33	0.32	125.95	0.34
		3	**0.00**	0.09	2.33	0.30	124.69	0.37	**0.00**	0.29	2.38	0.32	127.35	0.42
		5	**0.21**	0.45	2.54	0.35	129.09	0.32	**0.02**	0.28	2.35	0.25	130.11	0.36
		10	**0.02**	0.27	2.35	0.37	123.87	0.42	**0.08**	0.24	2.41	0.36	123.51	0.42
S16	(6,10)	1	**0.00**	0.51	1.65	0.31	54.28	0.22	**0.01**	0.53	1.66	0.28	56.84	0.23
		3	**0.02**	0.12	1.67	0.28	50.54	0.24	**0.22**	0.90	1.87	0.32	58.82	0.28
		5	**0.02**	0.61	1.68	0.32	59.27	0.24	**0.05**	0.45	1.71	0.37	73.10	0.28
		10	**0.21**	0.47	1.70	0.38	45.93	0.26	**0.55**	0.47	2.20	0.40	54.90	0.28
S18	(8,10)	1	**0.24**	0.36	82.75	0.27	101.38	0.26	**0.22**	0.56	86.94	0.30	108.18	0.28
		3	**0.14**	0.29	67.41	0.31	100.41	0.27	**0.22**	0.78	80.52	0.29	117.53	0.30
		5	**0.37**	0.82	69.61	0.30	91.55	0.30	**0.53**	0.58	75.48	0.34	96.01	0.29
		10	**0.11**	0.57	48.99	0.32	102.53	0.34	**0.08**	0.67	46.85	0.30	115.58	0.34
S20	(10,10)	1	**0.43**	0.40	18.82	0.22	61.25	0.45	**0.77**	0.57	19.16	0.22	50.91	0.48
		3	**0.32**	0.22	18.62	0.22	50.76	0.58	**0.62**	1.09	19.04	0.24	56.04	0.61
		5	**0.60**	0.54	18.92	0.22	53.93	0.55	**1.03**	0.76	19.60	0.23	51.51	0.49
		10	**0.33**	0.47	18.73	0.26	46.57	0.57	**0.07**	0.67	18.57	0.24	37.40	0.56
Average			**0.35**	0.59	9.25	0.30	44.55	0.42	**0.40**	0.76	9.74	0.30	47.82	0.42

of the DB-II heuristic using LR scheme outperforms the DB-I heuristic in CPU times. Moreover, the DB-II heuristics yields the best results in 18 out of 88 cases for the random instances. Analogously, the DB-II heuristic finds the best results in 17 out of 88 cases for the regular instances. The DB-II heuristic is slightly more efficient than the DB-III heuristic with CPU times of 0.75 (0.74) and 0.98 (0.98) seconds on the average, respectively, for random (regular) instances. The accuracy of the DB-III heuristic is poor for small instances. The efficiency of the DB-I heuristic deteriorates especially for the instances with more than 30 customers (or similarly 30 candidate facility locations). The DB-II and DB-III heuristics show different behavior than the DB-I heuristic does since they are not significantly affected from number of customers. Observe that increasing the number of facilities deteriorates the CPU time of all DB heuristics.

Table 4 presents the results of the large CMWP-B instances based on Boyacı [48] instances. The

Table 3. The performance of the DB heuristics over small sized random and regular CMWP-B instances based on Boyacı [48] instances.

Instance			Random CMWP-B Instances						Regular CMWP-B Instances					
Name	Size		DB-I		DB-II		DB-III		DB-I		DB-II		DB-III	
	(I, J)	P	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU
B1	(5,20)	1	**0.05**	0.34	**0.05**	0.39	32.18	0.33	**0.17**	0.34	**0.17**	0.41	38.30	0.33
		3	**0.54**	0.37	**0.54**	0.34	37.42	0.36	**0.46**	0.33	**0.46**	0.35	38.98	0.36
		5	**0.40**	0.31	**0.40**	0.39	38.94	0.38	**0.78**	0.51	**0.78**	0.47	34.55	0.38
		10	**0.78**	0.38	**0.78**	0.47	36.49	0.42	**0.39**	0.47	**0.39**	0.51	36.85	0.42
B2	(5,25)	1	**0.00**	0.70	3.76	0.60	51.09	0.52	**0.00**	0.83	3.37	0.56	58.36	0.52
		3	**0.67**	0.75	8.27	0.50	63.57	0.54	**0.78**	0.87	7.08	0.51	56.15	0.54
		5	**0.44**	0.81	8.49	0.57	57.21	0.60	**0.68**	0.88	9.85	0.53	53.32	0.60
		10	**0.74**	0.77	9.19	0.63	57.23	0.65	**0.66**	1.12	11.14	0.59	61.46	0.65
B3	(5,30)	1	**0.20**	2.76	**0.20**	0.58	47.59	0.62	**0.20**	2.50	**0.20**	0.52	47.79	0.62
		3	**0.42**	2.21	**0.42**	0.57	46.91	0.74	**0.74**	5.28	**0.74**	0.53	49.04	0.74
		5	**0.92**	4.37	**0.92**	0.56	54.74	0.77	**0.44**	4.98	**0.44**	0.60	44.22	0.77
		10	**0.43**	5.12	**0.43**	0.60	53.33	1.01	**0.86**	6.21	**0.86**	0.70	48.66	1.01
B4	(5,40)	1	**0.02**	3.02	0.05	0.32	21.15	1.04	**0.07**	2.99	0.10	0.28	21.45	1.04
		3	**0.19**	3.10	0.22	0.29	22.98	1.23	**0.58**	3.31	0.61	0.27	26.89	1.23
		5	**0.59**	2.80	0.62	0.28	27.10	1.22	**0.23**	3.23	0.26	0.31	21.80	1.22
		10	**0.32**	2.85	0.35	0.34	27.77	1.34	**0.36**	3.01	0.39	0.34	33.58	1.34
B5	(5,50)	1	**0.03**	101.95	5.32	1.04	15.12	1.26	**0.03**	116.17	4.76	0.87	14.56	1.26
		3	**0.16**	76.65	5.12	0.94	12.92	1.34	**0.23**	82.50	6.33	0.95	13.04	1.34
		5	**0.16**	63.15	5.14	1.02	13.31	1.38	**0.20**	49.95	6.69	0.98	12.62	1.38
		10	**0.34**	63.81	4.32	1.06	13.54	1.51	**0.32**	44.87	2.80	1.12	12.81	1.51
B8	(6,20)	1	**0.31**	0.59	**0.31**	0.53	68.57	0.50	**0.19**	0.48	**0.19**	0.48	58.26	0.50
		3	**0.32**	0.56	**0.32**	0.65	64.36	0.54	**0.00**	0.57	**0.00**	0.57	70.93	0.54
		5	**0.42**	0.56	**0.42**	0.54	52.49	0.63	**0.53**	0.54	**0.53**	0.57	61.18	0.63
		10	**0.23**	0.56	**0.23**	0.54	73.76	0.58	**0.35**	0.61	**0.35**	0.77	68.48	0.58
B9	(6,25)	1	**0.03**	1.05	4.10	0.47	124.74	0.53	**0.21**	1.06	5.37	0.51	112.04	0.53
		3	**0.54**	0.99	5.82	0.52	111.94	0.67	**0.92**	1.02	8.54	0.44	105.25	0.67
		5	**0.36**	1.11	3.33	0.51	105.92	0.59	**0.73**	1.06	7.58	0.51	81.73	0.59
		10	**0.38**	1.17	5.16	0.54	94.46	0.75	**0.91**	1.10	4.31	0.61	70.42	0.75
B10	(6,30)	1	**0.00**	8.24	7.50	0.62	28.19	0.80	**0.05**	8.21	5.41	0.68	35.84	0.80
		3	**0.57**	19.68	18.18	0.61	44.18	0.87	**0.30**	16.25	12.86	0.62	35.18	0.87
		5	**0.66**	10.77	15.38	0.66	36.87	0.84	**0.57**	19.02	10.88	0.70	36.86	0.84
		10	**0.24**	13.75	15.67	0.77	29.06	1.03	**0.59**	17.23	12.53	0.69	36.15	1.03
B11	(7,20)	1	**0.04**	0.76	0.09	0.54	108.15	0.46	**0.06**	0.75	0.21	0.67	87.65	0.46
		3	**0.02**	0.82	**0.02**	1.44	82.21	0.51	**0.01**	0.82	0.06	1.18	93.72	0.51
		5	**0.05**	0.74	7.28	0.74	92.82	0.64	**0.02**	0.78	0.07	1.10	90.46	0.64
		10	**0.16**	0.83	**0.16**	1.43	87.46	0.74	**0.02**	0.84	**0.02**	1.13	89.41	0.74
B12	(7,25)	1	**0.00**	1.39	0.17	0.25	105.95	0.63	**0.00**	1.20	0.17	0.24	95.98	0.63
		3	**0.07**	1.20	0.30	0.30	93.29	0.68	**0.13**	1.26	0.40	0.29	92.26	0.68
		5	**0.17**	1.16	0.40	0.33	89.13	0.73	**0.13**	1.22	0.30	0.38	77.31	0.73
		10	**0.19**	1.36	0.55	0.38	83.14	0.83	**0.18**	1.49	0.38	0.28	85.83	0.83
B13	(7,30)	1	**0.04**	16.59	**0.04**	1.00	76.97	0.77	**0.02**	13.80	**0.02**	0.99	83.92	0.77
		3	**0.18**	8.93	**0.18**	1.02	71.56	0.89	**0.33**	13.29	**0.33**	1.03	78.50	0.89
		5	**0.29**	13.54	**0.29**	0.91	88.40	0.81	**0.27**	7.47	**0.27**	0.89	87.11	0.81
		10	**0.36**	11.81	**0.36**	0.96	86.16	0.95	**0.72**	18.39	**0.72**	0.96	81.93	0.95
B14	(8,20)	1	**0.30**	5.39	4.50	0.48	133.92	0.63	**0.10**	7.61	4.30	0.44	115.08	0.63
		3	**0.46**	5.31	4.13	0.47	117.73	0.63	**0.81**	3.42	6.07	0.48	97.48	0.63
		5	**0.50**	4.43	4.59	0.50	105.46	0.75	**0.64**	5.58	4.21	0.49	107.66	0.75
		10	**0.43**	4.26	4.77	0.49	120.21	0.80	**0.13**	7.06	4.06	0.51	103.52	0.80
B15	(8,25)	1	**0.03**	2.01	3.55	0.87	65.21	0.71	**0.14**	6.83	3.67	0.87	68.04	0.71
		3	**0.21**	5.41	3.77	0.80	67.40	0.77	**0.36**	2.23	3.79	0.84	74.58	0.77
		5	**0.17**	1.95	3.76	0.84	65.58	0.79	**0.28**	3.35	3.94	0.80	67.69	0.79
		10	**0.36**	2.04	3.75	1.05	67.55	0.89	**0.15**	1.91	4.03	0.92	68.80	0.89
B16	(8,30)	1	**0.08**	51.08	10.58	0.75	45.20	1.00	**0.04**	39.32	10.36	0.73	34.97	1.00
		3	**0.52**	31.69	14.88	0.76	49.86	0.86	**0.76**	38.38	13.66	0.83	53.35	0.86
		5	**0.37**	43.56	14.77	0.80	30.17	0.93	**0.42**	36.88	11.79	0.83	40.73	0.93
		10	**0.27**	42.10	8.35	0.88	39.20	1.28	**0.49**	43.64	11.31	0.86	52.64	1.28
B17	(9,20)	1	**0.04**	9.82	42.06	0.55	220.47	0.84	**0.11**	8.80	44.80	0.55	228.22	0.84
		3	**0.13**	7.01	39.32	0.56	204.10	0.67	**0.20**	16.08	44.82	0.56	188.51	0.67
		5	**0.57**	6.32	33.91	0.62	186.02	0.68	**0.29**	13.87	45.44	0.57	169.32	0.68
		10	**0.25**	15.25	33.06	0.66	214.30	0.96	**0.28**	4.69	45.34	0.59	186.90	0.96
B18	(9,25)	1	**0.00**	45.28	12.98	0.78	158.12	0.78	**0.00**	44.64	14.78	0.72	200.06	0.78
		3	**0.24**	34.22	13.33	0.71	178.78	0.80	**0.43**	35.12	12.70	0.80	228.81	0.80
		5	**0.03**	33.45	16.84	0.68	167.18	0.92	**0.11**	45.52	12.41	0.69	244.27	0.92
		10	**0.15**	45.24	15.98	0.75	157.24	1.08	**0.06**	40.49	28.64	0.69	175.83	1.08
B19	(9,30)	1	**0.03**	42.40	0.77	0.77	134.87	1.18	**0.04**	47.03	0.59	0.78	142.30	1.18
		3	**0.44**	43.16	0.89	0.74	143.19	1.06	**0.25**	45.49	1.11	0.79	123.20	1.06
		5	**0.11**	42.03	1.22	0.74	115.09	1.45	**0.10**	39.37	0.64	0.75	122.89	1.45
		10	**0.35**	23.60	2.79	0.86	123.19	1.57	**0.11**	38.27	2.10	0.77	109.46	1.57
B20	(10,20)	1	**0.01**	24.28	2.19	0.58	135.01	0.65	**0.04**	24.88	2.01	0.57	150.63	0.65
		3	**0.23**	23.92	9.23	0.54	156.90	0.69	**0.19**	24.32	2.95	0.55	165.59	0.69
		5	**0.35**	23.60	9.64	0.59	134.63	0.70	**0.28**	24.63	7.31	0.60	144.17	0.70
		10	**0.19**	22.88	2.24	0.63	118.57	0.89	**0.29**	25.04	2.62	0.65	130.06	0.89
B21	(10,25)	1	**0.00**	47.80	115.53	0.61	246.99	1.02	**0.00**	35.73	131.34	0.62	205.23	1.02
		3	**0.02**	46.10	127.92	0.61	215.16	1.12	**0.04**	45.53	151.11	0.60	237.05	1.12
		5	**0.04**	42.90	107.31	0.68	252.52	1.05	**0.09**	45.24	153.95	0.60	177.75	1.05
		10	**0.26**	36.21	109.13	0.75	178.67	1.44	**0.24**	42.55	183.06	0.72	243.54	1.44
B22	(10,30)	1	**0.00**	62.66	13.94	1.12	205.34	1.09	**0.00**	81.07	15.82	0.95	158.65	1.09
		3	**0.03**	122.41	13.92	1.02	137.33	1.06	**0.33**	79.10	31.12	0.99	150.42	1.06
		5	**0.09**	70.16	11.64	1.13	147.51	1.23	**0.04**	86.41	13.43	1.07	156.84	1.23
		10	**0.31**	56.70	19.67	1.09	183.74	1.36	**0.38**	67.46	10.02	1.06	177.48	1.36
B23	(10,40)	1	**0.10**	456.79	3.54	1.10	38.53	1.40	**0.06**	321.95	2.08	1.16	34.11	1.40
		3	**0.18**	296.80	2.08	1.20	38.00	1.37	**0.21**	418.63	2.79	1.15	36.06	1.37
		5	**0.15**	432.46	2.59	1.24	29.16	1.49	**0.14**	547.25	2.80	1.19	35.61	1.49
		10	**0.33**	369.49	2.21	1.34	31.96	1.81	**0.49**	319.49	3.15	1.23	37.22	1.81
B24	(10,50)	1	**0.17**	1964.71	10.03	1.68	53.50	2.54	**0.03**	1357.01	8.91	1.63	53.16	2.54
		3	**0.16**	1676.72	8.91	1.68	63.20	2.62	**0.15**	1626.71	9.73	1.64	54.25	2.62
		5	**0.22**	1537.81	10.12	1.74	58.06	2.81	**0.27**	1691.68	7.83	1.60	61.16	2.81
		10	**0.43**	1417.33	9.82	1.87	60.08	3.60	**0.17**	1689.67	9.43	1.86	67.56	3.60
Average			**0.25**	110.60	11.67	0.75	91.20	0.98	**0.29**	108.96	13.81	0.74	91.16	0.98

structure of Table 4 is similar to Table 2 and Table 3. We have imposed an additional CPU time limit of 3600 seconds for large instances. The DB-I heuristic outperforms the others in 26 (28) out of 72 cases for random (regular) test instances. However, the efficiency of the DB-I heuristic significantly diminishes on large instances. Only instances B6 with 5 facilities and 100 customers can be solved within CPU time limits. Moreover, the DB-I heuristic can not produce a solution for 20 out of 72 cases that correspond to the instances with 25 facilities and 500 customers for both random and regular instances. On the other hand, our findings are outstanding for the DB-II heuristic on large instances. The DB-II heuristic yields the best outcome on 46 (42) out of 72 cases for large random (regular) test instances. Clearly, the computational requirements of the DB-II heuristic is quite reasonable for large instances when compared to the DB-I heuristic. Lastly, we have observed that the DB-III heuristic yields superior results in 2 out of 72 cases on regular instances. The DB-III performs even faster than the DB-II heuristic. Therefore, it can be used as a compromise for larger instances where both DB-I and DB-II heuristics require prohibitive solution times. Finally, we should point on the DB-III heuristic performs almost twice faster than the DB-II heuristic does on large instances in average.

6. Conclusion

In this work, we have focused on the capacitated multi-facility Weber problem with polyhedral barriers (CMWP-B). We have suggested a mathematical formulation for the CMWP-B and its discrete equivalent that is a MILP problem. Discretization based heuristic procedures for the CMWP-B have been proposed. We have carried out extensive computational experiments on randomly generated test instances that are based on standard CMWP instances. The proposed methods compute upper bounds for the optimal objective value of the CMWP-B.

The first discretization based heuristic solves the mixed integer linear programming formulation using the customer locations as the set of candidate facility locations. Its efficiency is improved by using a Lagrangean relaxation scheme and subgradient algorithm that resulted in the second discretization based heuristic. Lastly, the third discretization based heuristic employs a tabu search algorithm using a neighborhood search over customer locations. Among the discretization based heuristics, the first one yields the highest accuracy. However, its performance is limited with relatively small instances. As a remedy, we have applied a Lagrangean relaxation scheme within the second discretization based heuristic which constitutes a compromise between accuracy and efficiency. The performance of the third discretization based heuristic is usually poor, however, its performance is promising for large instances. Implementing exact solution procedures can be a good direction of research for the CMWP-B in the future. Efforts to provide an effective branch and bound algorithm is of high importance for this type of location-allocation type problems. Last but not least, a probabilistic extension of the CMWP-B, where the customer locations and/or their demand quantities are stochastic, is a worthwhile open research area.

Acknowledgments

The work described in this paper is supported by a grant from Galatasaray University Scientific Research Project Grant No: 16.402.013 and the Turkish Scientific and Technological Research Council (TÜBİTAK) Research Grant No: 214M222.

References

[1] Cooper, L. (1972). The transportation-location problem. *Oper. Res.*, 20, 94-108.

[2] Sherali, H.D. and Nordai, F.L. (1988). NP-hard, capacitated, balanced p-median problems on a chain graph with a continuum of link demands. *Math. Oper. Res.*, 13, 32-49.

[3] Meggido, N. and Supowit, K.J. (1984). On the complexity of some common geometric location problems. *SIAM J. Comput.*, 13, 182-196.

[4] Alpaydın, E., Altınel, İ.K. and Aras, N. (1996). Parametric distance functions vs. nonparametric neural networks for estimating road travel distances. *Eur. J. Oper. Res.*, 93, 230-243.

[5] Brimberg, J., Walker, J.H. and Love, R.F. (2007). Estimation of travel distances with the weighted l_p norm: some empirical results. *J. Transp. Geogr.*, 15, 62-72.

[6] Klamroth, K. (2002). *Single-facility location problems with barriers.* Springer, Berlin.

[7] Cooper, L. (1964). Heuristic methods for location-allocation problems. *SIAM Rev.*, 6, 37-53.

[8] Kuenne, R.E. and Soland, R.M. (1972). Exact and approximate solutions to the multisource Weber problem. *Math. Program.*, 3, 193-209.

[9] Rosing, K.E. (1992). An optimal method for solving the (generalized) multi-Weber problem. *Eur. J. Oper. Res.*, 58, 414-426.

[10] Krau, S. (1997). *Extensions du problème de Weber.* Thesis (PhD). University of Montréal.

[11] Righini, G. and Zaniboni, L. (2007). A branch-and-price algorithm for the multi-source Weber problem. *Int. J. Oper. Res.*, 2, 188-207.

[12] Love, R.F. and Juel, H. (1982). Properties and solution methods for large location-allocation problems. *J. Oper. Res. Soc.*, 33, 443-452.

[13] Bongartz, I., Calamai, P.H. and Conn, A.R. (1994). A projection method for the ℓ_p norm location-allocation problems. *Math. Program.*, 66, 283-312.

Table 4. The performance of the DB heuristics over large sized random and regular CMWP-B instances based on Boyacı [48] instances.

Instance			Random CMWP-B Instances						Regular CMWP-B Instances					
Name	Size		DB-I		DB-II		DB-III		DB-I		DB-II		DB-III	
	(I, J)	P	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU	% Dev.	CPU
B6	(5,100a)	1	**0.03**	1828.94	1.03	2.28	7.22	3.74	**0.07**	1871.62	0.66	2.17	5.85	3.74
		3	**0.20**	1886.21	1.55	2.46	7.08	3.80	**0.08**	2014.49	1.27	2.19	9.48	3.80
		5	**0.13**	1547.83	1.25	2.41	7.19	3.32	**0.15**	2198.07	0.83	2.63	7.13	3.32
		10	**0.29**	1639.21	1.35	2.72	6.67	3.93	**0.31**	2301.38	0.97	2.57	8.93	3.93
B7	(5,250)	1	**1.75**	3630.16	6.74	13.46	8.05	5.48	**1.21**	3626.48	6.01	14.43	7.72	5.48
		3	**1.01**	3630.67	6.08	14.20	7.37	6.20	**1.19**	3625.06	5.72	14.10	7.36	6.20
		5	**1.39**	3628.52	5.06	13.38	7.23	6.14	**1.06**	3638.15	7.14	14.19	6.95	6.14
		10	**2.01**	3626.56	7.27	14.89	7.07	6.72	**0.99**	3624.78	6.36	13.86	7.29	6.72
B25	(10,100)	1	**1.82**	3699.10	3.88	5.46	24.44	4.56	**1.89**	3602.56	4.57	4.91	22.57	4.56
		3	**1.52**	3615.53	3.67	4.81	21.93	4.12	**1.89**	3604.93	2.92	5.18	24.27	4.12
		5	**1.12**	3657.24	4.34	5.22	23.86	4.33	**1.59**	3649.64	2.35	5.41	24.04	4.33
		10	**2.89**	3521.21	4.40	5.48	24.32	5.11	**2.50**	3605.03	4.17	4.97	23.01	5.11
B26	(20,40)	1	**9.72**	3639.06	56.09	1.86	239.90	3.45	**8.50**	3625.53	48.80	1.85	236.09	3.45
		3	**7.01**	3632.89	64.83	1.76	281.70	3.36	**10.09**	3633.84	63.08	1.95	242.18	3.36
		5	**11.61**	3631.01	58.21	1.96	231.54	3.49	**10.98**	3624.91	56.91	2.00	254.30	3.49
		10	**10.59**	3622.99	62.21	2.23	236.63	4.27	**8.53**	3643.25	52.23	2.28	246.29	4.27
B27	(20,50)	1	**8.43**	3619.96	77.15	2.45	144.96	4.36	**7.56**	3622.60	80.12	2.44	157.92	4.36
		3	**6.70**	3625.67	86.18	2.39	142.29	4.27	**8.36**	3621.60	89.58	2.32	153.58	4.27
		5	**8.87**	3633.69	68.75	2.51	153.49	5.09	**12.48**	3641.95	90.35	2.54	150.77	5.09
		10	**11.43**	3624.86	96.97	2.79	150.03	5.42	**9.01**	3630.88	83.07	2.95	167.62	5.42
B28	(20,250)	1	**11.10**	3670.45	17.38	55.72	26.30	35.27	**10.41**	3669.79	15.64	57.16	25.29	35.27
		3	18.03	3722.43	**16.26**	55.57	25.51	37.97	**12.32**	3636.56	20.10	52.61	26.26	37.97
		5	**11.89**	3698.75	18.52	58.54	25.84	39.17	**11.90**	3664.88	17.01	55.20	25.24	39.17
		10	17.40	3784.30	**16.69**	57.62	27.21	43.92	**10.96**	3640.60	20.02	54.30	26.56	43.92
B29	(25,250a)	1	23.68	3792.63	**22.23**	72.10	39.07	37.96	22.35	3835.17	**20.47**	67.38	37.49	37.96
		3	24.00	3803.01	**21.17**	66.36	34.41	44.34	27.05	3786.01	**20.52**	70.57	34.44	44.34
		5	25.23	3879.52	**22.64**	68.93	34.98	50.14	24.81	3800.52	**22.33**	69.41	35.26	50.14
		10	23.71	3999.63	**20.30**	73.25	35.46	48.43	23.48	3784.19	**22.37**	73.74	36.97	48.43
B30	(25,250b)	1	9.63	3812.81	**9.28**	71.35	21.36	48.12	13.71	3796.26	**8.84**	73.82	20.61	48.12
		3	11.03	3804.33	**9.37**	74.53	19.70	51.78	11.60	3790.43	**9.17**	65.78	18.27	51.78
		5	13.04	3805.01	**10.92**	72.72	21.19	55.35	12.35	3802.08	**9.00**	66.38	17.56	55.35
		10	12.04	3799.99	**10.39**	71.40	18.94	60.52	10.97	3800.07	**10.52**	72.73	19.11	60.52
B31	(25,250c)	1	10.84	3769.46	**9.10**	68.75	14.92	65.95	10.17	3770.65	**9.13**	68.55	18.49	65.95
		3	11.17	3769.92	**8.33**	66.51	17.49	61.66	8.15	3767.76	**8.02**	65.22	15.87	61.66
		5	12.66	3777.18	**9.00**	69.42	15.22	72.64	11.44	3787.90	**8.70**	68.70	17.15	72.64
		10	9.70	3778.33	**5.97**	71.59	17.97	65.99	8.98	3771.89	**7.36**	70.61	14.40	65.99
B32	(25,250d)	1	22.66	3758.52	**14.04**	68.08	21.65	60.69	25.28	3761.29	**12.91**	72.25	23.98	60.69
		3	20.44	3764.54	**13.32**	69.92	24.45	54.80	22.35	3771.12	**11.92**	75.27	25.50	54.80
		5	21.90	3765.79	**12.00**	67.80	21.19	64.30	24.40	3748.71	**14.67**	74.33	23.63	64.30
		10	24.61	3761.55	**13.47**	73.74	23.38	70.62	23.64	3768.80	**13.36**	72.50	24.48	70.62
B33	(25,250e)	1	13.72	3931.19	**6.89**	77.40	17.43	59.38	16.66	3847.32	**5.55**	68.52	15.59	59.38
		3	13.80	3788.08	**5.00**	71.64	15.62	66.10	12.74	3862.72	**5.58**	72.49	17.48	66.10
		5	13.32	3782.14	**5.68**	72.93	16.13	64.92	12.21	3938.20	**5.38**	73.00	17.25	64.92
		10	12.46	3796.71	**5.95**	72.69	19.59	68.20	15.89	4070.35	**4.81**	75.19	15.81	68.20
B34	(25,500a)	1	N/A	N/A	**5.78**	323.61	9.26	155.23	N/A	N/A	**5.24**	328.38	8.71	155.23
		3	N/A	N/A	**5.16**	337.02	6.55	162.01	N/A	N/A	5.70	309.09	**5.61**	162.01
		5	N/A	N/A	**4.85**	324.65	5.48	167.10	N/A	N/A	**4.87**	309.49	5.14	167.10
		10	N/A	N/A	**5.07**	339.89	5.70	182.96	N/A	N/A	5.30	323.41	**4.86**	182.96
B35	(25,500b)	1	N/A	N/A	**2.20**	334.21	12.95	92.05	N/A	N/A	**2.76**	318.95	12.32	92.05
		3	N/A	N/A	**1.64**	329.23	13.67	104.66	N/A	N/A	**3.61**	336.81	13.52	104.66
		5	N/A	N/A	**3.04**	322.77	13.66	96.81	N/A	N/A	**2.84**	346.44	13.63	96.81
		10	N/A	N/A	**3.82**	340.62	13.14	119.04	N/A	N/A	**2.48**	330.88	12.51	119.04
B36	(25,500c)	1	N/A	N/A	**8.97**	303.48	10.47	123.60	N/A	N/A	**8.44**	318.49	11.78	123.60
		3	N/A	N/A	**7.29**	310.64	12.65	114.78	N/A	N/A	**8.55**	314.52	10.74	114.78
		5	N/A	N/A	**7.19**	318.15	10.96	126.00	N/A	N/A	**7.16**	324.89	13.20	126.00
		10	N/A	N/A	**7.31**	315.51	13.52	129.35	N/A	N/A	**8.13**	311.83	10.78	129.35
B37	(25,500d)	1	N/A	N/A	**2.78**	319.46	7.40	161.42	N/A	N/A	**3.23**	332.95	7.74	161.42
		3	N/A	N/A	**3.12**	329.97	8.07	168.51	N/A	N/A	**3.24**	329.30	8.16	168.51
		5	N/A	N/A	**2.69**	335.53	7.33	171.57	N/A	N/A	**3.00**	322.69	8.03	171.57
		10	N/A	N/A	**3.47**	308.40	7.00	191.51	N/A	N/A	**3.31**	323.25	7.83	191.51
B38	(25,500e)	1	N/A	N/A	**4.96**	313.73	10.22	110.78	N/A	N/A	**4.44**	317.81	9.72	110.78
		3	N/A	N/A	**4.43**	323.08	10.01	113.92	N/A	N/A	**3.61**	324.16	9.37	113.92
		5	N/A	N/A	**5.15**	308.20	9.43	120.62	N/A	N/A	**5.40**	328.98	9.95	120.62
		10	N/A	N/A	**5.38**	326.54	9.52	128.10	N/A	N/A	**4.54**	349.60	11.24	128.10
B39	(30,100)	1	**20.79**	3629.47	46.12	13.34	146.56	20.78	**20.18**	3628.04	49.56	13.56	158.42	20.78
		3	**17.25**	3629.07	43.44	13.59	154.82	19.32	**21.62**	3652.17	49.94	13.58	148.84	19.32
		5	**20.75**	3630.10	43.83	13.66	167.71	21.94	**20.62**	3635.66	44.21	13.99	162.94	21.94
		10	**21.72**	3627.36	48.53	14.83	162.79	25.38	**16.13**	3625.33	46.85	14.51	160.60	25.38
B40	(50,250)	1	30.87	5704.95	**27.45**	132.80	76.57	69.14	32.62	5989.58	**25.41**	132.79	79.60	69.14
		3	32.10	5363.78	**29.91**	142.86	78.11	83.04	36.27	5633.25	**30.00**	137.68	76.49	83.04
		5	32.11	4638.97	**27.47**	133.93	75.84	73.29	33.33	5049.48	**29.48**	146.09	77.99	73.29
		10	32.43	3728.11	**30.00**	148.43	73.74	95.08	35.92	4329.07	**28.14**	140.66	75.60	95.08
Average			**13.17**[a]	3652.10[a]	**18.22**	122.66	47.52	63.44	**13.25**[a]	3708.13[a]	**18.19**	122.99	47.94	63.44

[a]These average values are calculated excluding unavailable values shown as N/A

[14] Hansen, P., Mladenović, N. and Taillard, É. (1998). Heuristic solution of the multisource Weber problem as a p-median problem. *Oper. Res. Lett.*, 22, 55-62.
[15] Gamal, M.D.H. and Salhi, S. (2003). A cellular heuristic for the multi-source Weber problem. *Comput. Oper. Res.*, 30, 1609-1624.
[16] Brimberg, J., Hansen, P. and Mladenović, N. (2006). Decomposition strategies for large-scale continuous location-allocation problems. *IMA J. Manag Math.*, 17, 307-316.
[17] Brimberg, J., Drezner, Z., Mladenović, N. and Salhi, S. (2014). A new local search for continuous location problems. *Eur. J. Oper. Res.*, 232, 256-265.
[18] Drezner, Z., Brimberg, J., Mladenović, N. and Salhi, S. (2016). New local searches for solving the multi-source Weber problem. *Ann. Oper. Res.*, 246, 181-203.
[19] Sherali, H.D., Al-Loughani, I. and Subramanian, S. (2002). Global optimization procedures for the capacitated Euclidean and L_p distance multifacility location-allocation problem. *Oper. Res.,* 50, 433-448.
[20] Akyüz, M.H., Altınel, İ.K. and Öncan, T. (2014). Location and allocation based branch and bound algorithms for the capacitated multi-facility Weber problem. *Ann. Oper. Res.*, 222, 45-71.
[21] Zainuddin, Z.M. and Salhi, S. (2007). A perturbation-based heuristic for the capacitated multisource Weber problem. *Eur. J. Oper. Res.*, 179, 1194-1207.
[22] Luis, M., Salhi, S. and Nagy, G. (2009). Region-rejection based heuristics for the capacitated multi-source Weber problem. *Comput. Oper. Res.*, 36, 2007-2017.
[23] Aras, N., Altınel, İ.K. and Orbay, M. (2007). New heuristic methods for the capacitated multi-facility Weber problem. *Nav. Res. Log.*, 54, 21-32.
[24] Boyacı, B., Altınel, İ.K. and Aras, N. (2013). Approximate solution methods for the capacitated multi-facility Weber problem. *IIE Trans.*, 45, 97-120.
[25] Luis, M., Salhi, S. and Nagy, G. (2011). A guided reactive GRASP for the capacitated multi-source Weber problem. *Comput. Oper. Res.*, 38, 1014-1024.
[26] Akyüz, M.H., Öncan, T. and Altınel, İ.K. (2012). Efficient approximate solution methods for the multi-commodity capacitated multi-facility Weber problem. *Comput. Oper. Res.*, 39, 225-237.
[27] Akyüz, M.H., Öncan, T. and Altınel, İ.K. (2013). Beam search heuristics for the single and multi-commodity capacitated Multi-facility Weber Problems. Comput. Oper. Res., 40, 3056-3068.
[28] Brimberg, J., Hansen, P., Mladenović, N. and Salhi, S. (2008). A survey of solution methods for the continuous location-allocation problem. *Int. J. Oper. Res.*, 5, 1-12.
[29] Katz, I.N. and Cooper, L. (1981). Facility location in the presence of forbidden regions, I: formulation and the case of Euclidean distance with one forbidden circle. *Eur. J. Oper. Res.*, 6, 166-173.
[30] Larson, R.C. and Sadiq, G. (1983). Facility Locations with the Manhattan Metric in the Presence of Barriers to Travel. *Oper. Res.*, 31, 652-669.
[31] Batta, R., Ghose, A. and Palekar, U.S., (1989). Locating Facilities on the Manhattan Metric with Arbitrarily Shaped Barriers and Convex Forbidden Regions. *Transport. Sci.*, 23, 26-36.
[32] Aneja, Y.P. and Parlar, M. (1994). Technical Note–Algorithms for Weber Facility Location in the Presence of Forbidden Regions and/or Barriers to Travel. *Transport. Sci.*, 28, 70-76.
[33] Butt, S.E. and Cavalier, T.M. (1996). An efficient algorithm for facility location in the presence of forbidden regions. *Eur. J. Oper. Res.*, 90, 56-70.
[34] Klamroth, K. (2001). A reduction result for location problems with polyhedral barriers. *Eur. J. Oper. Res.*, 130, 486-497.
[35] McGarvey, R.G. and Cavalier, T.M. (2003). A global optimal approach to facility location in the presence of forbidden regions. *Comput. Ind. Eng.*, 45, 1-15.
[36] Bischoff, M. and Klamroth, K. (2007). An efficient solution method for Weber problems with barriers based on genetic algorithm. *Eur. J. Oper. Res.*, 177, 22-41.
[37] Bischoff, M., Fleischmann, T. and Klamroth, K. (2009). The multi-facility location-allocation problem with polyhedral barriers. *Comput. Oper. Res.*, 36, 1376-1392.
[38] Wendell, R.E. and Hurter, A.P. (1973). Location theory, dominance and convexity. *Oper. Res.* 21, 314-320.
[39] Ghosh, S.K. (2007). *Visibility algorithms in the plane.* Cambridge University Press, New York.
[40] Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. Numer. Math., 1, 269-271.
[41] Hansen, P., Perreur, J. and Thisse, F. (1980). Location theory, dominance and convexity: some further results. *Oper. Res.*, 28, 1241-1250.
[42] Weiszfeld, E. (1937). Sur le point lequel la somme des distances de n points donné est minimum. *Tôhoku Math. J.*, 43, 355-386.
[43] Brimberg, J. and Love, R.F. (1993). Global convergence of a generalized iterative procedure for the minisum location problem with L_p distances. *Oper. Res.*, 41, 1153-1163.
[44] Brimberg, J., Chen, R. and Chen, D. (1998). Accelerating convergence in the Fermat-Weber location problem. *Oper. Res.*, 22, 151-157
[45] Martello, S. and Toth P. (1990). *Knapsack problems: algorithms and computer implementations.* Wiley, New York.
[46] Held, M., Wolfe, P. and , Crowder H.P. (1974). Validation of subgradient optimization. *Math. Program.*, 6, 62-88.
[47] Glover, F.W. and Laguna, M. (1997). *Tabu search.* Kluwer academic publishers, Boston
[48] Boyacı, B. (2009). *Solving the capacitated multifacility Weber problem approximately.* Thesis (MSc). Boğaziçi University.

M. Hakan Akyüz *received a PhD degree from the Department of Industrial Engineering at Boğaziçi University in 2011. He worked as a post-doctoral fellow at the Department of Industrial Engineering and Logistics Management at the Hong Kong University of Science and Technology afterwards. His research area includes applications of operations research in facility location, container terminal operations and maritime logistics. He is working as an Assistant Professor at the Department of Industrial Engineering at Galatasaray University since 2014.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.1, pp.43-62 (2018)
http://doi.org/10.11121/ijocta.01.2018.00325

RESEARCH ARTICLE

A decoupled Crank-Nicolson time-stepping scheme for thermally coupled magneto-hydrodynamic system

S.S. Ravindran*

Department of Mathematical Sciences, University of Alabama, Huntsville, Alabama, USA
ravinds@uah.edu

ARTICLE INFO

Article History:
Received 20 March 2016
Accepted 11 August 2017
Available 23 October 2017

Keywords:
Thermally coupled MHD
Crank-Nicolson
Mixed finite element
Error estimates
Non-homogeneous boundary condition
Partitioned method

AMS Classification 2010:
76W05, 65M12, 65M60, 76D05,35Q30

ABSTRACT

Thermally coupled magneto-hydrodynamics (MHD) studies the dynamics of electro-magnetically and thermally driven flows, involving MHD equations coupled with heat equation. We introduce a partitioned method that allows one to decouple the MHD equations from the heat equation at each time step and solve them separately. The extrapolated Crank-Nicolson time-stepping scheme is used for time discretization while mixed finite element method is used for spatial discretization. We derive optimal order error estimates in suitable norms without assuming any stability condition or restrictions on the time step size. We prove the unconditional stability of the scheme. Numerical experiments are used to illustrate the theoretical results.

1. Introduction

Thermally coupled magneto-hydrodynamics has many applications including in electromagnetic pumping design [35], electromagnetic filtration [4], contact-less electromagnetic stirring [32] and damping convective flow in metal-like melt [34]. Magnetohydrodynamics in general has broad applications including fusion [19], underwater propulsion [18], nuclear reactors [13], metallurgy [1, 2, 11, 31] and astrophysics [30]. In all of these applications, qualitative and quantitative understanding of the dynamics is important to achieve optimal operating conditions. This has led to considerable research efforts over the past three decades into the development of theoretical, see e.g [16, 24, 26, 27, 29] and efficient and accurate computational techniques, see e.g. [8, 9, 20, 21] for MHD equations. Majority of the numerical analysis work done on the equations has been for steady state equations. In [17, 23, 25, 33], time stepping schemes for unsteady MHD equations have been analyzed. However, these work consider MHD equations where thermal effects are negligible. Thermally coupled MHD equations model a complex flow phenomena which is in general three dimensional, highly nonlinear and represents multi-physics.

In this work, we propose and analyze a decoupled time stepping scheme for the thermally coupled MHD equations. It uses a semi-implicit Crank-Nicolson scheme, which combines an implicit treatment of the second derivative terms, a semi-implicit second order extrapolation of the nonlinear convective terms and an explicit treatment of the temperature coupling term in the Navier-Stokes equations. The proposed scheme solves the MHD equations and the heat equation separately in each time step (without iteration) allowing the possibility of optimizing the subproblem's respective physics. We show unconditional stability of the scheme and provide a complete error analysis for fully discrete scheme using finite element spatial discretization.

*Corresponding Author

The remaining of the paper is organized as follows: The continuum problem and some preliminaries are presented in Section 2. In Section 3, we present the decoupled time-stepping scheme and analyze its stability, accuracy and convergence. Finally, we present a numerical example that illustrates our theoretical results.

2. Continuum problem and preliminaries

To begin with,we present some notations and basic results that will be used throughout the article.

2.1. Continuum problem

The non-dimensional Boussinesq equations describing thermally coupled MHD equations are (see for e.g. [15])

$$\left\{\begin{array}{rcl} \dfrac{\partial \mathbf{u}}{\partial t} & - & Pr_\theta \Delta \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + Pr_\theta \nabla p \\ & - & S(\nabla\times\mathbf{B})\times\mathbf{B} = Pr_\theta Ra \theta \mathbf{i}_3 + \mathbf{f}_1 \,, \\ \frac{\partial \mathbf{B}}{\partial t} & + & Pr_B \nabla\times(\nabla\times\mathbf{B}) \\ & - & \nabla\times(\mathbf{u}\times\mathbf{B}) = \mathbf{0}\,, \\ \frac{\partial \theta}{\partial t} & - & \Delta\theta + \mathbf{u}\cdot\nabla\theta = f_2\,, \\ \nabla\cdot\mathbf{u} & = & 0\,, \\ \nabla\cdot\mathbf{B} & = & 0\,, \end{array}\right. \tag{1}$$

in $(0,T]$, where T denotes time and $\Omega \subset \mathbb{R}^d (d = 2,3)$ a bounded region with Lipschitz-continuous boundary Γ. Moreover the different fields appearing in the equations are $\mathbf{u}(\mathbf{x},t)$ the fluid velocity, $\mathbf{B}(\mathbf{x},t)$ the magnetic field, θ the temperature, $p(\mathbf{x},t)$ the pressure, $\mathbf{f}$ the source and $\mathbf{i}_3$ the unit basis vector. The non-dimensional numbers that appear in the MHD equations are $S := Pr_B Pr_\theta H^2$, the Hartman number H, the Rayleigh number Ra, the thermal Prandtl number Pr_θ and the magnetic Prandtl number Pr_B. The MHD system we consider is supplemented with the initial conditions

$$\begin{array}{rcl} \mathbf{u}(\mathbf{x},0) & = & \mathbf{u}_0(\mathbf{x})\,, \, \theta(\mathbf{x},0) = \theta_0(\mathbf{x}) \text{ and} \\ \mathbf{B}(\mathbf{x},0) & = & \mathbf{B}_0(\mathbf{x}) \quad \text{in}\, \Omega\,, \end{array} \tag{2}$$

along with the boundary conditions

$$\left\{\begin{array}{rcl} \mathbf{u}|_\Gamma & = & \mathbf{g} \text{ with } \int_\Gamma \mathbf{g}\cdot\mathbf{n}\, ds = 0\,, \\ \theta|_\Gamma & = & \widetilde{q}\,, \\ \mathbf{B}\cdot\mathbf{n}|_\Gamma & = & q \text{ with } \int_\Gamma q\, ds = 0\,, \\ Pr_B & & (\nabla\times\mathbf{B})\times\mathbf{n}|_\Gamma \\ & - & (\mathbf{u}\times\mathbf{B})\times\mathbf{n}|_\Gamma = \mathbf{k} \\ & & \text{with } \mathbf{k}\cdot\mathbf{n} = 0\,, \quad \int_\Gamma \mathbf{k}\, ds = 0\,. \end{array}\right. \tag{3}$$

2.2. Function spaces

For a Banach space X, we denote by $L^p(0,T;X)$ the time-space function space endowed with the norm $\|w\|_{L^p(0,T;X)} := \left(\int_0^T \|w\|_X^p\, dt\right)^{1/p}$ if $1 \leq p < \infty$ and $\operatorname{ess\,sup}_{t\in[0,T]} \|w\|_X$ if $p = \infty$.

We will often use the abbreviated notation $L^p(X) := L^p(0,T;X)$ for convenience. The symbol $C([0,T];X)$ denotes the set of continuous functions $u : [0,T] \to X$ endowed with the norm $\|u\|_{C(0,T;X)} := \max_{0\leq t\leq T} \|u(t)\|_X$. For any integer $k \geq 1$, let $W^{k,p}(\Omega)$ be the Sobolev space of functions in $L^p(\Omega)$ with derivatives up-to the k^{th} order endowed with the norm $\|\phi\|_{m,p} := \left[\sum_{|\alpha|\leq m} \int_\Omega |\partial_x^\alpha \phi(\mathbf{x})|^p dx\right]^{\frac{1}{p}}$ where $\partial_x^\alpha \phi(\mathbf{x}) := \frac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1}\cdots\partial_{x_d}^{\alpha_d}} \phi(\mathbf{x})$, $\alpha := (\alpha_1, \cdots, \alpha_d)$, $\alpha_i \geq 0$, $|\alpha| := \sum_{i=1}^d \alpha_i$.

We denote by $H^k(\Omega)$ the space $W^{k,2}(\Omega)$, when $p = 2$, and drop the subscripts $p(=2)$ in referring to the norm in $H^k(\Omega)$. Moreover, we will use the following simplified norm notations:

$$\|u\| := \|u\|_{L^2(\Omega)} \quad \text{and} \quad \|u\|_\infty := \|u\|_{L^\infty(\Omega)}\,.$$

For $\mathbf{g} \in \mathbf{H}^{\frac{1}{2}}(\Gamma)$ satisfying $\int_\Gamma \mathbf{g}\cdot\mathbf{n}\; ds = 0$ and $q \in H^{\frac{1}{2}}(\Gamma)$ satisfying $\int_\Gamma q\; ds = 0$, define $\mathbf{H}^1_{n,q}(\Omega) := \{\mathbf{v} \in \mathbf{H}^1(\Omega) \;:\; \mathbf{v}\cdot\mathbf{n}|_\Gamma = q\}$, $\mathbf{V}_g := \{\mathbf{v} \in \mathbf{H}^1(\Omega) \;:\; \mathbf{v}|_\Gamma = \mathbf{g}, \quad \nabla\cdot\mathbf{v} = 0\}$ and $H^1_{\widetilde{q}}(\Omega) := \{\theta \in H^1(\Omega) \;:\; \theta|_\Gamma = \widetilde{q}\}$.

We write $\mathbf{V} = \mathbf{V}_0$, $\mathbf{H}^1_n(\Omega) = \mathbf{H}^1_{n,0}(\Omega)$ and $\mathbb{V} := \{\mathbf{v} \in \mathbf{H}^1(\Omega) \;:\; \nabla\cdot\mathbf{v} = 0 \text{ in } \Omega\}$. We introduce the time discrete space $l^p(Z)$ associated with $L^p(0,T;Z)$; $l^p(Z)$ is the space of Z-valued sequences $w := \{w_n; n = 1,\ldots,N\}$ with norm $\|\cdot\|_{l^p(Z)}$ defined by

$$\|w\|_{l^p(Z)} := \begin{cases} (\Delta t \sum_{n=1}^{N} \|w_n\|_Z^p)^{1/p} & \text{if } 1 \le p < \infty \\ \max_{1\le n\le N} \|w_n\|_Z & \text{if } p = \infty\,. \end{cases}$$

For later purposes, we recall the inequality

$$\lambda_m \|\mathbf{B}\|_1^2 \le \|\nabla\cdot\mathbf{B}\|^2 + \|\nabla\times\mathbf{B}\|^2 \;\forall \mathbf{B} \in \mathbf{H}^1_n(\Omega)\,, \quad (4)$$

the Poincaré inequality

$$\|\mathbf{v}\|^2 \le \lambda_p \|\nabla \mathbf{v}\|^2 \quad \forall \mathbf{v} \in \mathbf{H}^1_0(\Omega)\,,$$

the Gagliardo-Nirenberg interpolation inequality [3]

$$\|\mathbf{u}\|_q \le C \|\nabla \mathbf{u}\|_p^{\lambda} \|\mathbf{u}\|_r^{1-\lambda} \;\forall \mathbf{u} \in \mathbf{W}^{1,p}(\Omega) \cap \mathbf{L}^r(\Omega)$$

for $0 \le \lambda \le 1$ and $\frac{1}{q} = \lambda(\frac{1}{p} - \frac{1}{d}) + (1-\lambda)\frac{1}{r}$ and the Agmon's inequality

$$\|\mathbf{u}\|_\infty \le C \|\mathbf{u}\|_1^{\frac{1}{2}} \|\mathbf{u}\|_2^{\frac{1}{2}} \quad \forall \mathbf{u} \in \mathbf{H}^2(\Omega) \cap \mathbf{H}^1_0(\Omega)\,.$$

We define the explicitly skew-symmetrized trilinear forms

$$\begin{aligned} c_1(\mathbf{u},\mathbf{v},\mathbf{w}) &:= \tfrac{1}{2}\textstyle\int_\Omega [(\mathbf{u}\cdot\nabla)\mathbf{v}\cdot\mathbf{w} - (\mathbf{u}\cdot\nabla)\mathbf{w}\cdot\mathbf{v}]\, d\Omega\,, \\ &= \textstyle\int_\Omega \left[(\mathbf{u}\cdot\nabla)\mathbf{v}\cdot\mathbf{w} + \tfrac{1}{2}(\nabla\cdot\mathbf{u})\mathbf{v}\cdot\mathbf{w}\right]\, d\Omega\,, \end{aligned}$$

for all $\mathbf{u},\mathbf{v},\mathbf{w} \in \mathbf{H}^1(\Omega)$ with $(\mathbf{u}\cdot\mathbf{n})\mathbf{v}\cdot\mathbf{w} = 0$ on Γ and

$$\begin{aligned} c_2(\mathbf{u},\theta,\psi) &:= \tfrac{1}{2}\textstyle\int_\Omega [(\mathbf{u}\cdot\nabla)\theta\,\psi - (\mathbf{u}\cdot\nabla)\psi\,\theta]\, d\Omega\,, \\ &= \textstyle\int_\Omega \left[(\mathbf{u}\cdot\nabla)\theta\,\psi + \tfrac{1}{2}(\nabla\cdot\mathbf{u})\psi\,\theta\right]\, d\Omega\,, \end{aligned}$$

for all $\mathbf{u} \in \mathbf{H}^1(\Omega)$, $\theta, \psi \in H^1(\Omega)$ with $(\mathbf{u}\cdot\mathbf{n})\theta\,\psi = 0$ on Γ.
Moreover, we define the bilinear forms

$$b(\mathbf{v}, r) := -\int_\Omega Pr_\theta\, r\, \nabla\cdot\mathbf{v}\, d\Omega\,,$$

$$e(\theta,\mathbf{v}) := Pr_\theta Ra \int_\Omega \theta \mathbf{i}_3 \cdot \mathbf{v}\, d\Omega\,,$$

and the trilinear form

$$d(\mathbf{B},\mathbf{C},\mathbf{v}) := \int_\Omega \mathbf{B} \times (\nabla\times\mathbf{C}) \cdot \mathbf{v}\, d\Omega\,.$$

Notice that the trilinear form $d(\cdot,\cdot,\cdot)$ is skew-symmetric with respect to the first and last arguments, i.e., $d(\mathbf{B},\mathbf{C},\mathbf{v}) = -d(\mathbf{v},\mathbf{C},\mathbf{B})$.
We end this section with a result regarding the existence and uniqueness of solutions to the initial-boundary value problem (1)-(3) whose proof can be furnished by using Galerkin approximations, a-priori estimates and compactness methods.

Proposition 1. *Assume that the given functions* $\mathbf{f}$, $\mathbf{g}$, $\mathbf{k}$, q, $\widetilde{q}$, $\mathbf{u}_0$ *and* $\mathbf{B}_0$ *satisfy* $\mathbf{f}_1 \in L^2(0,T;\mathbf{H}^{-1}(\Omega))$, $f_2 \in L^2(0,T;H^{-1}(\Omega))$, $\mathbf{g} \in H^1(0,T;\mathbf{H}^{\frac{1}{2}}(\Gamma))$, $\mathbf{k} \in L^2(0,T;\mathbf{H}^{-\frac{1}{2}}(\Gamma))$, $q \in H^1(0,T;H^{\frac{1}{2}}(\Gamma))$, $\widetilde{q} \in H^1(0,T;H^{\frac{1}{2}}(\Gamma))$, $\int_\Gamma \mathbf{g}\cdot\mathbf{n}\,ds = 0$, $\int_\Gamma q\,ds = 0$, $\mathbf{k}\cdot\mathbf{n}|_\Gamma = 0$, $\mathbf{u}_0 \in \mathbf{V}_{\mathbf{g}(\cdot,0)}$, $\mathbf{B}_0 \in \mathbf{H}^1_{n,q(\cdot,0)}(\Omega)$ *and* $\theta_0 \in H^1\hat{q}(\cdot,0)(\Omega)$. *Then, the problem (1)-(3) has at least one solution* $(\mathbf{u},p,\theta,\mathbf{B})$ *such that* $\mathbf{u} \in L^\infty(0,T;\mathbf{L}^2(\Omega)) \cap L^2(0,T;\mathbf{V}_\mathbf{g})$, $\theta \in L^2(0,T;H^1_{\widetilde{q}}(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$, $\mathbf{B} \in L^\infty(0,T;\mathbf{L}^2(\Omega)) \cap L^2(0,T;\mathbf{H}^1_{n,q}(\Omega))$ *and* $p \in L^2(0,T;L^2_0(\Omega))$. *In two-spatial dimension* ($d = 2$), *these solutions are unique.*

2.3. Properties of finite element spaces and projections

In order to keep the exposition simple, we restrict our attention to convex polyhedral domains. Let $\mathcal{T}_h$ be a family of subdivisions (e.g. triangulation) of $\overline{\Omega} \subset \mathbb{R}^d$ satisfying $\overline{\Omega} = \cup_{K\in\mathcal{T}_h} K$ so that $diameter(K) \le h$ and any two closed elements K_1 and $K_2 \in \mathcal{T}_h$ are either disjoint or share exactly one face, side or vertex. Suppose further that $\mathcal{T}_h$ is a shape regular and quasi-uniform triangulation. That is, there exists a constant $C > 0$ such that the ratio between the diameter h_K of an element $K \in \mathcal{T}_h$ and the diameter of the largest ball contained in K is bounded uniformly by C, and h_K is comparable with the mesh size $h = \max_{K\in\mathcal{T}_h} h_K$ for all $K \in \mathcal{T}_h$. For example, $\mathcal{T}_h$ consists of triangles for $d = 2$ or tetrahedra for $d = 3$ that are non-degenerate as $h \to 0$. We choose families of finite dimensional spaces $\mathbb{X}_h \subset H^1(\Omega)$, $\mathbb{Y}_h \subset H^1_n(\Omega)$, $\mathbb{Z}_h \subset H^1(\Omega)$ and $\mathbb{Q}_h \subset L^2(\Omega)$, parameterized by a parameter h such that $0 < h < 1$. Let $\mathbf{g}_h$, q_h and $\widetilde{q}_h$ be approximations of $\mathbf{g}$, q and $\widetilde{q}$, respectively, such that there exists $\mathbf{v}_h \in \mathbb{X}_h$, $\mathbf{C}_h \in \mathbb{Y}_h$ and satisfying $\mathbf{v}_h|_\Gamma = \mathbf{g}_h$, $\mathbf{C}_h\cdot\mathbf{n}|_\Gamma = q_h$ and $\theta_h|_\Gamma = \widetilde{q}_h$. We then define $\mathbf{X}_{h,g_h} := \mathbb{X}_h \cap \mathbf{H}^1{}_{g_h}$, $\mathbf{Y}_{h,q_h} := \{\mathbf{C}_h \in \mathbb{Y}_h(\Omega) : \mathbf{C}_h\cdot\mathbf{n}|_\Gamma = q_h\}$, $Z_{h,\widetilde{q}_h} := \mathbb{Z}_h \cap \mathbf{H}^1{}_{\widetilde{q}_h}$ and $Q_h := \mathbb{Q}_h \cap L^2_0(\Omega)$. We also define the discretely divergence free space is given by

$$\mathbf{V}_{h,g_h} := \{\mathbf{v}_h \in X_{h,g_h} : (\nabla\cdot\mathbf{v}_h, r_h) = 0 \forall r_h \in \mathbb{Q}_h\}.$$

We set $\mathbf{V}_h := \mathbf{V}_{h,0}$, $\mathbf{Y}_h := \mathbf{Y}_{h,0}$, $Z_h := Z_{h,0}$ and $X_h = X_{h,0}$.

We make the following assumptions on the finite dimensional subspaces $\mathbb{X}_h, \mathbb{Y}_h, \mathbb{Z}_h$ and $\mathbb{Q}_h$:

Assumption A1.

We have the approximation properties: there exists an integer k and a constant C, independent of h, $\mathbf{v}$, $\mathbf{B}$, θ and r, such that

$$\inf_{\mathbf{v}_h \in \mathbb{X}_h} [\|\mathbf{v} - \mathbf{v}_h\| + h\|\nabla(\mathbf{v} - \mathbf{v}_h)\|] \le Ch^{\ell+1}\|\mathbf{v}\|_{\ell+1}$$

$$\inf_{\mathbf{B}_h \in \mathbb{Y}_h} [\|\mathbf{B} - \mathbf{B}_h\| + h\|\nabla(\mathbf{B} - \mathbf{B}_h)\|] \le Ch^{\ell+1}\|\mathbf{B}\|_{\ell+1}$$

$$\inf_{\theta_h \in \mathbb{Z}_h} [\|\theta - \theta_h\| + h\|\nabla(\theta - \theta_h)\|] \le Ch^{\ell+1}\|\theta\|_{\ell+1}$$

and

$$\inf_{r_h \in Q_h} \|r - r_h\| \le Ch^{\ell}\|r\|_{\ell}$$

for all $\mathbf{v} \in \mathbf{H}^{\ell+1}(\Omega)$, $\mathbf{B} \in \mathbf{H}^{\ell+1}(\Omega)$, $\theta \in H^{\ell+1}(\Omega)$, and $r \in H^{\ell}(\Omega) \quad 1 \le \ell \le k$.

Assumption A2. (Discrete inf-sup condition) For every $r_h \in Q_h$, there exists a nonzero function $\mathbf{v}_h \in \mathbb{X}_h$ and $\beta > 0$ such that

$$|(r_h, \nabla \cdot \mathbf{v}_h)| \ge \beta \|\nabla \mathbf{v}_h\| \|r_h\|,$$

with an inf-sup constant $\beta > 0$ that is independent of the mesh size h.

Assumption A3. For any integers l and m $(0 \le l \le m \le 1)$ and any real numbers p and q $(1 \le p \le q \le \infty)$ it holds that

$$\|\psi_h\|_{m,q} \le ch^{l-m+d(1/q-1/p)}\|\psi_h\|_{l,p} \quad \forall \psi_h \in \mathbb{X}_h.$$

There are many conforming finite element spaces satisfying the assumptions $(A1)$-$(A3)$. One may choose, for example, the Taylor-Hood element pair for the velocity and pressure (i.e, piecewise quadratic polynomial for velocity and piecewise linear polynomial for pressure), and piecewise quadratic polynomials for the magnetic field and temperature. Then, hypothesis (A1)-(A3) hold with $k = 2$.

We define Stokes, Maxwell and Ritz projections as follows: Given $(\mathbf{u}, p) \in \mathbf{H}^1(\Omega) \times L_0^2(\Omega)$, $\theta \in H^1(\Omega)$ and $\mathbf{B} \in \mathbf{H}^1(\Omega)$, we define the Stokes projection $(P_h^s\mathbf{u}, P_h^s p) \in \mathbf{X}_{h,g_h} \times Q_h$ as the solution of the problem

$$\begin{aligned} Pr_\theta(\nabla(\mathbf{u} - P_h^s\mathbf{u}), \nabla \mathbf{v}_h) &+ b(\mathbf{v}_h, (p - P_h^s p)) \\ &= 0 \quad \forall \mathbf{v}_h \in \mathbf{X}_h, \\ b(\mathbf{u} - P_h^s\mathbf{u}, r_h) &= 0 \quad \forall r_h \in Q_h, \end{aligned} \tag{5}$$

the Maxwell projection $P_h^m\mathbf{B} \in \mathbf{Y}_{h,q_h}$ as the solution of the problem

$$\begin{aligned} &(\nabla \times (\mathbf{B} - P_h^m\mathbf{B}), \nabla \times \boldsymbol{\phi}_h) \\ &\quad + (\nabla \cdot (\mathbf{B} - P_h^m\mathbf{B}), \nabla \cdot \boldsymbol{\phi}_h) \\ &\quad = 0 \quad \forall \boldsymbol{\phi}_h \in \mathbf{Y}_h, \end{aligned} \tag{6}$$

and the Ritz projection $P_h^r\theta \in \mathbf{Z}_{h,\widetilde{q}_h}$ as the solution of the problem

$$(\nabla(\theta - P_h^r\theta), \nabla \psi_h) = 0 \quad \forall \psi_h \in \mathbf{Z}_h, \tag{7}$$

We have the following convergence and boundedness results for these projections.

Lemma 1. *Suppose that assumptions (A1)-(A2) hold with a positive integer k, and that $(\mathbf{u}, p) \in \mathbf{H}^{k+1} \times (L_0^2(\Omega) \cap H^k(\Omega))$, $\theta \in H^{k+1}(\Omega)$ and $\mathbf{B} \in \mathbf{H}^{k+1}(\Omega)$. Then, for any $h \in (0, h_0]$ the Stokes projection $(P_h^s\mathbf{u}, P_h^s p)$ of $(\mathbf{u}, p)$ satisfies*

$$\|\mathbf{u} - P_h^s\mathbf{u}\|_1 + \|p - P_h^s p\| \le ch^k(\|\mathbf{u}\|_{k+1} + \|p\|_k), \tag{8}$$

the Maxwell projection $P_h^m\mathbf{B}$ of $\mathbf{B}$ satisfies

$$\|\mathbf{B} - P_h^m\mathbf{B}\|_1 \le ch^k\|\mathbf{B}\|_{k+1}, \tag{9}$$

and the Ritz projection $P_h^r\theta$ of θ satisfies

$$\|\theta - P_h^r\theta\|_1 \le ch^k\|\theta\|_{k+1}. \tag{10}$$

Moreover, suppose that assumption (A3) holds. Then, $P_h^s\mathbf{u}$, $P_h^m\mathbf{B}$ and $P_h^r\theta$ satisfy

$$\|P_h^s\mathbf{u}\|_\infty + \|P_h^s\mathbf{u}\|_{1,3} \le c(\|\mathbf{u}\|_2 + \|p\|_1), \tag{11}$$

$$\|P_h^m\mathbf{B}\|_\infty + \|P_h^m\mathbf{B}\|_{1,3} \le c\|\mathbf{B}\|_2, \tag{12}$$

and

$$\|P_h^r\theta\|_\infty + \|P_h^r\theta\|_{1,3} \le c\|\theta\|_2. \tag{13}$$

Proof. The proof of (8)-(10) follows by the regularity properties of the Stokes, Maxwell and Ritz projections and by duality argument. In order to prove (11)-(13), we first notice that Gagliardo-Nirenberg's inequality yields

$$\|\phi\|_{0,\infty} + \|\phi\|_{1,3} \le C\|\phi\|_1^{1/2}\|\phi\|_2^{1/2} .$$

Therefore the approximation properties (8)-(1) together with Agmon's inequality yield the desired result. □

Let Δt denote the step size for t so that $t_n = n\Delta t,\ n = 0, 1, 2, \ldots, N$. For notational convenience, we denote $\phi^n := \phi(t_n)$, $\mathcal{D}(\phi^n) := \frac{\phi^{n+1}-\phi^n}{\Delta t}$, $\phi^{n+1/2} := \phi^{n+1} + \phi^n$ and $\mathcal{I}(\phi^{n+1/2}) := \phi^n + \frac{1}{2}\phi^{n-1} - \frac{1}{2}\phi^{n-2}$, [5, 14].

Lemma 2. *If $\phi(t)$ is smooth enough, then*

$$\begin{aligned}
(i)\|\phi^{n+1/2} &- \phi(t_{n+1/2})\|_k^2 \\
&\le \tfrac{(\Delta t)^3}{48}\int_{t_n}^{t_{n+1}} \|\partial_t^2\phi\|_k^2\,dt ,\\
(ii)\|\partial_t\phi(t_{n+1/2}) &- \mathcal{D}(\phi(t_n))\|^2 \\
&\le \tfrac{(\Delta t)^3}{1280}\int_{t_n}^{t_{n+1}} \|\partial_t^3\phi(t)\|^2\,dt ,\\
(iii)\|\mathcal{I}(\phi(t_{n+1/2})) &- \phi(t_{n+1/2})\|_{H^k}^2 \\
&\le c(\Delta t)^{3/2}\int_{t_n}^{t_{n+1}} \|\partial_t^2\phi(t)\|_k^2\,dt.
\end{aligned}$$

Moreover, let $P_h^s\mathbf{u}$ be the Stokes projection of $\mathbf{u}$, $P_h^m\mathbf{B}$ the Maxwell projection of $\mathbf{B}$ and $P_h^r\theta$ the Ritz projection of θ. If assumptions (A1)-(A2) hold with a positive integer k, then

$$\begin{aligned}
(iv)\|\mathcal{D}(\mathbf{u}(t_{n+1}) &- P_h^s\mathbf{u}(t_{n+1}))\| \\
&\le \tfrac{ch^k}{\sqrt{\Delta t}}\|(\partial_t\mathbf{u}, \partial_t p)\|_{L^2(t_n,t_{n+1};H^{k+1}\times H^k)} ,\\
(v)\|\mathcal{D}(\mathbf{B}(t_{n+1}) &- P_h^m\mathbf{B}(t_{n+1}))\| \\
&\le \tfrac{ch^k}{\sqrt{\Delta t}}\|\partial_t\mathbf{B}\|_{L^2(t_n,t_{n+1};H^{k+1})} ,\\
(vi)\|\mathcal{D}(\theta(t_{n+1}) &- P_h^r\theta(t_{n+1}))\| \\
&\le \tfrac{ch^k}{\sqrt{\Delta t}}\|\partial_t\theta\|_{L^2(t_n,t_{n+1};H^{k+1})} .
\end{aligned}$$

Proof. The proof of (i)-(iii) follows by Taylor expansion with integral remainder whereas the proof of (iv)-(vi) follows as a consequence of Lemma 1. □

We will need the following well known discrete Grönwall lemma.

Lemma 3. ***(Discrete Grönwall lemma)*** *Let d, Δt, $\{a_n\}_{n\ge0}$, $\{b_n\}_{n\ge0}$, $\{c_n\}_{n\ge0}$, and $\{d_n\}_{n\ge0}$ be nonnegative numbers such that*

$$a_m + \Delta t\sum_{n=1}^{m} b_n \le \Delta t\sum_{n=0}^{m-1} a_n d_n + \Delta t\sum_{n=0}^{m-1} c_n + d ,$$

for $m \ge 1$. Then we have

$$a_m + \Delta t\sum_{n=1}^{m} b_n \le \exp(\Delta t\sum_{n=0}^{m-1} d_n)(\Delta t\sum_{n=0}^{m-1} c_n + d)$$

for $m \ge 1$.

A proof of this result can be found, for e.g, in [12].

3. Decoupled Crank-Nicolson time-stepping scheme

We discretize the system (1) by Crank-Nicholson scheme in time and Galerkin finite element in space. The time discretization combines an implicit treatment of the second derivative terms, a semi-implicit second-order extrapolation for the nonlinear convective terms and explicit treatment of the temperature coupling term in the Navier-Stokes equations.

Algorithm 1. *Given $(\mathbf{u}_h^i, \mathbf{B}_h^i, p_h^i, \theta_h^i) \in \mathbf{X}_{h,g_h^i} \times \mathbf{Y}_{h,q_h^i} \times Q_h \times \mathbf{Z}_{h,\widetilde{q}_h^i}$, $i = 0, 1$, find $\{(\mathbf{u}_h^n, \mathbf{B}_h^n, p_h^n, \theta_h^n) \in \mathbf{X}_{h,g_h^n} \times \mathbf{Y}_{h,q_h^n} \times Q_h \times \mathbf{Z}_{h,\widetilde{q}_h^i}$ such that*

$$\left\{\begin{aligned}
(\mathcal{D}\mathbf{u}_h^n, \mathbf{v}_h) &+ Pr_\theta(\nabla\mathbf{u}_h^{n+1/2}, \nabla\mathbf{v}_h) \\
&+ c_1(\mathcal{I}(\mathbf{u}_h^{n+1/2}), \mathbf{u}_h^{n+1/2}, \mathbf{v}_h) \\
&+ b(\mathbf{v}_h, p_h^{n+1/2}) \\
&+ Sd(\mathcal{I}(\mathbf{B}_h^{n+1/2}), \mathbf{B}_h^{n+1/2}, \mathbf{v}_h) \\
&= e(\mathcal{I}(\theta_h^{n+1/2}), \mathbf{v}_h) \\
&+ (\mathbf{f}_1^{n+1/2}, \mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{X}_h ,\\
b(\mathbf{u}_h^{n+1/2}, r_h) &= 0 \quad \forall r_h \in Q_h ,\\
(\mathcal{D}\mathbf{B}_h^n, \boldsymbol{\phi}_h) &+ Pr_B[(\nabla\times\mathbf{B}_h^{n+1/2}, \nabla\times\boldsymbol{\phi}_h) \\
&+ (\nabla\cdot\mathbf{B}_h^{n+1/2}, \nabla\cdot\boldsymbol{\phi}_h)] \\
&+ d(\mathbf{u}_h^{n+1/2}, \boldsymbol{\phi}_h, \mathcal{I}(\mathbf{B}_h^{n+1/2})) \\
&= (\mathbf{k}^{n+1/2}, \boldsymbol{\phi}_h)_\Gamma \quad \forall \boldsymbol{\phi}_h \in \mathbf{Y}_h ,\\
(\mathcal{D}\theta_h^n, \psi_h) &+ (\nabla\theta_h^{n+1/2}, \nabla\psi_h) \\
&+ c_2(\mathcal{I}(\mathbf{u}_h^{n+1/2}), \theta_h^{n+1/2}, \psi_h) \\
&= (f_2^{n+1/2}, \psi_h) \quad \forall \psi_h \in Z_h ,
\end{aligned}\right. \tag{14}$$

for $n = 1, \ldots, N$, where $\mathbf{u}_h^{n+1/2}$, $\mathbf{B}_h^{n+1/2}$, $\theta_h^{n+1/2}$ and $p_h^{n+1/2}$ are the intermediate variables defined by $\mathbf{u}_h^{n+1/2} := \mathbf{u}_h^{n+1} + \mathbf{u}_h^n$, $\mathbf{B}_h^{n+1/2} := \mathbf{B}_h^{n+1} + \mathbf{B}_h^n$, $\theta_h^{n+1/2} := \theta_h^{n+1} + \theta_h^n$ and $p_h^{n+1/2} := p_h^{n+1} + p_h^n$, respectively.

3.1. Stability analysis

In this section, we demonstrate the unconditional energy stability of the decoupled scheme proposed in Section 2. We first recall a few basic facts and some notation that are needed below. Let us define the discrete trace spaces of $\mathbb{X}_h$, $\mathbb{Y}_h$ and $\mathbb{Z}_h$ by

$$\begin{aligned}\Lambda_h(\Gamma) \;:=\; & \{\mathbf{g}_h \in \mathbf{H}^{\frac{1}{2}}(\Gamma): \text{ there exists } \\ & \mathbf{v}_h \in \mathbb{X}_h \text{ such that } \lambda_h|_{\partial K\cap\Gamma} \\ & = \mathbf{v}_h|_{\partial K\cap\Gamma} \;\; \forall\, K \in \mathcal{T}_h \\ & \text{ and } \partial K \cap \Gamma \neq \emptyset\},\end{aligned}$$

$$\begin{aligned}\widehat{\Lambda}_h(\Gamma) \;:=\; & \{q_h \in H^{\frac{1}{2}}(\Gamma): \text{ there exists } \\ & \mathbf{C}_h \in \mathbb{Y}_h \text{ such that } q_h|_{\partial K\cap\Gamma} \\ & = \mathbf{C}_h\cdot\mathbf{n}|_{\partial K\cap\Gamma} \;\; \forall\, K \in \mathcal{T}_h \\ & \text{ and } \partial K \cap \Gamma \neq \emptyset\}\end{aligned}$$

and

$$\begin{aligned}\widetilde{\Lambda}_h(\Gamma) \;:=\; & \{\widetilde{q}_h \in H^{\frac{1}{2}}(\Gamma): \text{ there exists } \\ & \phi_h \in \mathbb{Z}_h \text{ such that } \widetilde{q}_h|_{\partial K\cap\Gamma} \\ & = \phi_h|_{\partial K\cap\Gamma} \;\; \forall\, K \in \mathcal{T}_h \\ & \text{ and } \partial K \cap \Gamma \neq \emptyset\}.\end{aligned}$$

Moreover, we define

$$\Lambda_{h,0}(\Gamma) := \{\lambda_h \in \Lambda_h(\Gamma): \int_\Gamma \lambda_h\cdot\mathbf{n}\, ds = 0\,\}$$

and

$$\widehat{\Lambda}_{h,0}(\Gamma) := \{\lambda_h \in \widehat{\Lambda}_h(\Gamma): \int_\Gamma \lambda_h\, ds = 0\,\}.$$

Then there exists a discrete extension operator $E_h : \Lambda_{h,0}(\Gamma) \to \mathbb{V}_h$ such that $E_h(\mathbf{g}_h)|_\Gamma = \mathbf{g}_h$ and $\|E_h(\mathbf{g}_h)\|_1 \le C\|\mathbf{g}_h\|_{1/2,\Gamma}$, see [10, 28]. Similarly, we can define discrete extension operators $\widehat{E}_h$ and $\widetilde{E}_h$ such that $\widehat{E}_h(q_h)\cdot\mathbf{n}|_\Gamma = q_h$ and $\widetilde{E}_h(\widetilde{q}_h)|_\Gamma = \widetilde{q}_h$.
In order to prove, we first define suitable boundary extensions. Let $(E_h(\mathbf{g}_h^n), \widehat{E}_h(q_h^n), \widetilde{E}_h(\widetilde{q}_h^n)) \in \mathbf{V}_{h,g_h} \times \mathbf{Y}_{h,q_h^n} \times Z_{h,\widetilde{q}_h^n}$ be the extension of $(\mathbf{g}_h^n, q_h^n, \widetilde{q}_h^n)$ for each $n \ge 0$. Set $\boldsymbol{\zeta}_h^n = \mathbf{u}_h^n - E_h(\mathbf{g}_h^n)$, $\boldsymbol{\xi}_h^n = \mathbf{B}_h^n - \widehat{E}_h(q_h^n)$ and $\chi_h^n = \theta_h^n - \widetilde{E}_h(\widetilde{q}_h^n)$ so that $(\boldsymbol{\zeta}_h^n, \boldsymbol{\xi}_h^n, \chi_h^n) \in \mathbf{V}_h \times \mathbf{Y}_h \times Z_h$.
We make the following assumptions about the extension operators $E_h(\mathbf{g}_h^n), \widehat{E}_h(q_h^n), \widetilde{E}_h(\widetilde{q}_h^n)$.

Assumption A4.

The extension operators satisfy

$$\begin{aligned}(i) \quad & |c_1(\mathcal{I}(\boldsymbol{\zeta}_h^{n+1/2}), E_h(\mathbf{g}_h^{n+1/2}), \boldsymbol{\zeta}_h^{n+1/2})| \\ \le\; & \delta(\|\nabla\boldsymbol{\zeta}_h^{n-1/2}\| + \|\nabla\boldsymbol{\zeta}_h^{n-3/2}\|)\|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|\end{aligned}$$

and

$$\begin{aligned}& |d(E_h(\mathbf{g}_h^{n+1/2}), \boldsymbol{\xi}_h^{n+1/2}, \mathcal{I}(\boldsymbol{\xi}_h^{n+1/2}))| \\ \le\; & \delta^*(\|\nabla\times\boldsymbol{\xi}_h^{n-1/2}\| + \|\nabla\times\boldsymbol{\xi}_h^{n-3/2}\|) \\ & \|\nabla\times\boldsymbol{\xi}_h^{n+1/2}\|,\end{aligned}$$

$$\begin{aligned}(ii) \quad & |Sd(\mathcal{I}(\boldsymbol{\xi}_h^{n+1/2}), \widehat{E}_h(q_h^{n+1/2})), \boldsymbol{\zeta}_h^{n+1/2})| \\ \le\; & \delta^{**}(\|\nabla\times\boldsymbol{\xi}_h^{n-1/2}\| + \|\nabla\times\boldsymbol{\xi}_h^{n-3/2}\|) \\ & \|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|,\end{aligned}$$

$$\begin{aligned}(iii) \quad & |c_2(\mathcal{I}(\boldsymbol{\zeta}_h^{n+1/2}), \widetilde{E}_h(\widetilde{q}_h^{n+1/2}), \chi_h^{n+1/2})| \\ \le\; & \delta^{***}(\|\nabla\boldsymbol{\zeta}_h^{n-1/2}\| + \|\nabla\boldsymbol{\zeta}_h^{n-3/2}\|) \\ & \|\nabla\chi_h^{n+1/2}\|.\end{aligned}$$

Theorem 1. *Suppose assumption (A4) holds and let* $\{(\mathbf{g}_h^n, q_h^n, \widetilde{q}_h^n)\}_{n=0}^N$ *satisfies* $(\mathbf{g}_h, q_h, \widetilde{q}_h) \in l^4(\Lambda_{h,0}(\Gamma)) \times l^4(\widehat{\Lambda}_{h,0}(\Gamma)) \times l^4(\widetilde{\Lambda}_{h,0}(\Gamma))$ *and* $(\mathcal{D}\mathbf{g}_h, \mathcal{D}q_h, \mathcal{D}\widetilde{q}_h) \in l^2(\Lambda_{h,0}(\Gamma)) \times l^2(\widehat{\Lambda}_{h,0}(\Gamma)) \times l^2(\widetilde{\Lambda}_{h,0}(\Gamma))$, *and let* $\mathbf{f}_1 \in l^2(\mathbf{H}^{-1}(\Omega))$, $f_2 \in l^2(H^{-1}(\Omega))$ *and* $\mathbf{k} \in l^2(H^{-1/2}(\Gamma))$. *Suppose that* $(\mathbf{u}_h^i, \mathbf{B}_h^i, \theta_h^i) \in \mathbf{V}_{h,g_h^i} \times Y_{h,q_h^i} \times Z_{h,\widetilde{q}_h^i}$ *for* $i = 0, 1$ *are such that* $\|\mathbf{u}_h^2\|^2 + \Delta t\sum_{i=0}^{1}\|\mathbf{u}_h^{i+1/2}\|_1^2 < \infty$, $\|\mathbf{B}_h^2\|^2 + \Delta t\sum_{i=0}^{1}\|\mathbf{B}_h^{i+1/2}\|_1^2 < \infty$ *and* $\|\theta_h^2\|^2 + \Delta t\sum_{i=0}^{1}\|\theta_h^{i+1/2}\|_1^2 < \infty$ *as* $h, \Delta t \to 0$. *Then the solutions* $(\mathbf{u}_h^n, \mathbf{B}_h^n, \theta_h^n)$ *of (14) satisfies* $\|\mathbf{u}_h\|_{l^\infty(L^2(\Omega))} + \|\nabla\mathbf{u}_h\|_{l^2(L^2(\Omega))} < M_1$, $\|\mathbf{B}_h\|_{l^\infty(L^2(\Omega))} + \|\nabla\mathbf{B}_h\|_{l^2(L^2(\Omega))} < M_2$ *and* $\|\theta_h\|_{l^\infty(L^2(\Omega))} + \|\nabla\theta_h\|_{l^2(L^2(\Omega))} < M_3$, *for some constants* $M_1, M_2, M_3 > 0$.

Proof. Substituting $\mathbf{u}_h^n = \boldsymbol{\zeta}_h^n + E_h(\mathbf{g}_h^n)$, $\theta_h^n = \chi_h^n + \widetilde{E}_h(\widetilde{q}_h^n)$ and $\mathbf{B}_h^n = \boldsymbol{\xi}_h^n + \widehat{E}_h(q_h^n)$ into (14), then setting $(\mathbf{v}_h, \boldsymbol{\phi}_h, \psi_h) = (\boldsymbol{\zeta}_h^{n+1/2}, \boldsymbol{\xi}_h^{n+1/2}, \chi_h^{n+1/2})$ and using the skew-symmetry of $c_1(\cdot,\cdot,\cdot)$ and $c_2(\cdot,\cdot,\cdot)$, we obtain

$$
\left\{
\begin{aligned}
(\mathcal{D}\boldsymbol{\zeta}_h^n, \ & \boldsymbol{\zeta}_h^{n+1/2}) + Pr_\theta\|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|^2 \\
+ \ & Sd(\mathcal{I}(\mathbf{B}_h^{n+1/2}), \boldsymbol{\xi}_h^{n+1/2}, \boldsymbol{\zeta}_h^{n+1/2}) \\
\leq \ & (\mathbf{f}_1^{n+1/2}, \boldsymbol{\zeta}_h^{n+1/2}) - (\mathcal{D}E_h(\mathbf{g}_h^n), \boldsymbol{\zeta}_h^{n+1/2}) \\
+ \ & e(\mathcal{I}(\chi_h^{n+1/2}), \boldsymbol{\zeta}_h^{n+1/2}) \\
- \ & Pr_\theta(\nabla E_h(\mathbf{g}_h^{n+1/2}), \nabla\boldsymbol{\zeta}_h^{n+1/2}) \\
+ \ & e(\mathcal{I}(\widetilde{E}_h(\widetilde{q}_h^{n+1/2})), \boldsymbol{\zeta}_h^{n+1/2}) \\
- \ & c_1(\mathcal{I}(E_h(\mathbf{g}_h^{n+1/2})), E_h(\mathbf{g}_h^{n+1/2}), \boldsymbol{\zeta}_h^{n+1/2}) \\
- \ & Sd(\mathcal{I}((\widehat{E}_h(q_h^{n+1/2})), \widehat{E}_h(q_h^{n+1/2})), \boldsymbol{\zeta}_h^{n+1/2}) \\
- \ & c_1(\mathcal{I}(\boldsymbol{\zeta}_h^{n+1/2}), E_h(\mathbf{g}_h^{n+1/2}), \boldsymbol{\zeta}_h^{n+1/2}) \\
- \ & Sd(\mathcal{I}(\boldsymbol{\xi}_h^{n+1/2}), \widehat{E}_h(q_h^{n+1/2})), \boldsymbol{\zeta}_h^{n+1/2}) \\
=: \ & \textstyle\sum_{i=1}^9 A_i^n \\
(\mathcal{D}\boldsymbol{\xi}_h^n, \ & \boldsymbol{\xi}_h^{n+1/2}) + Pr_B[\|\nabla\times\boldsymbol{\xi}_h^{n+1/2}\|^2 \\
+ \ & \|\nabla\cdot\boldsymbol{\xi}_h^{n+1/2}\|^2] \\
+ \ & d(\boldsymbol{\zeta}_h^{n+1/2}, \boldsymbol{\xi}_h^{n+1/2}, \mathcal{I}(\mathbf{B}_h^{n+1/2})) \\
\leq \ & (\mathbf{k}^{n+1/2}, \boldsymbol{\xi}_h^{n+1/2})_\Gamma - (\mathcal{D}\widehat{E}_h(q_h^n), \boldsymbol{\xi}_h^{n+1/2}) \\
- \ & Pr_B(\nabla\times\widehat{E}_h(q_h^{n+1/2}), \nabla\times\boldsymbol{\xi}_h^{n+1/2}) \\
- \ & Pr_B(\nabla\cdot\widehat{E}_h(q_h^{n+1/2}), \nabla\cdot\boldsymbol{\xi}_h^{n+1/2}) \\
- \ & d(E_h(\mathbf{g}_h^{n+1/2}), \boldsymbol{\xi}_h^{n+1/2}, \mathcal{I}(\widehat{E}_h(q_h^{n+1/2}))) \\
- \ & d(E_h(\mathbf{g}_h^{n+1/2}), \boldsymbol{\xi}_h^{n+1/2}, \mathcal{I}(\boldsymbol{\xi}_h^{n+1/2})) \\
(\mathcal{D}\chi_h^n, \ & \chi_h^{n+1/2}) + \|\nabla\chi_h^{n+1/2}\|^2 \leq (f_2^{n+1/2}, \chi_h^{n+1/2}) \\
- \ & (\mathcal{D}\widetilde{E}_h(\widetilde{q}_h^n), \chi_h^{n+1/2}) \\
- \ & (\nabla\widetilde{E}_h(\widetilde{q}_h^{n+1/2}), \nabla\chi_h^{n+1/2}) \\
- \ & c_2(\mathcal{I}(E_h(\mathbf{g}_h^{n+1/2})), \widetilde{E}_h(\widetilde{q}_h^{n+1/2}), \chi_h^{n+1/2}) \\
- \ & c_2(\mathcal{I}(\boldsymbol{\zeta}_h^{n+1/2}), \widetilde{E}_h(\widetilde{q}_h^{n+1/2}), \chi_h^{n+1/2})\,.
\end{aligned}
\right. \tag{15}
$$

Let us next bound each term on the right-hand side of $(15)_1$ except the last two. The first five terms can be estimated using Cauchy/Duality and Young's inequalities to obtain

$$
\begin{aligned}
|\sum_{i=1}^5 \ & A_i^n| \leq C[\|\mathbf{f}_1^{n+1/2}\|_{-1}^2 + \|\nabla E_h(\mathbf{g}_h^{n+1/2})\|^2 \\
+ \ & \|\mathcal{I}(\widetilde{E}_h(\widetilde{q}_h^{n+1/2}))\|^2 + \|\mathcal{D}E_h(\mathbf{g}_h^n)\|_{-1}^2] \\
+ \ & \tfrac{Pr_\theta}{18}\|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|^2 + \tfrac{9}{2Pr_\theta}\|\mathcal{I}(\chi_h^{n+1/2})\|^2\,.
\end{aligned}
$$

We estimate A_6^n and A_7^n using Hölder's, Gagliardo-Nirenberg and Young's inequalities as follows

$$
\begin{aligned}
|A_6^n| = \ & |c_1(\mathcal{I}(E_h(\mathbf{g}_h^{n+1/2})), E_h(\mathbf{g}_h^{n+1/2}), \boldsymbol{\zeta}_h^{n+1/2})| \\
\leq \ & C\|\mathcal{I}(E_h(\mathbf{g}_h^{n+1/2}))\|_{L^4(\Omega)} \\
& \Big[\|\nabla E_h(\mathbf{g}_h^{n+1/2})\|\|\boldsymbol{\zeta}_h^{n+1/2}\|_{L^4(\Omega)} \\
+ \ & \|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|\|E_h(\mathbf{g}_h^{n+1/2})\|_{L^4(\Omega)}\Big] \\
\leq \ & C\|\mathcal{I}(E_h(\mathbf{g}_h^{n+1/2}))\|_1\|E_h(\mathbf{g}_h^{n+1/2})\|_1 \\
& \|\nabla\boldsymbol{\zeta}_h^{n+1/2}\| \\
\leq \ & C\sum_{i=0}^2\|E_h(\mathbf{g}_h^{n-i+1/2})\|_1^4 \\
+ \ & \tfrac{Pr_\theta}{18}\|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|^2
\end{aligned}
$$

and

$$
\begin{aligned}
|A_7^n| = \ & |Sd(\mathcal{I}(\widehat{E}_h(q_h^{n+1/2})), \widehat{E}_h(q_h^{n+1/2})), \boldsymbol{\zeta}_h^{n+1/2})| \\
\leq \ & C\|\mathcal{I}(\widehat{E}_h(q_h^{n+1/2}))\|_{L^4(\Omega)} \\
& \|\nabla\times\widehat{E}_h(q_h^{n+1/2}))\|\|\boldsymbol{\zeta}_h^{n+1/2}\|_{L^4(\Omega)} \\
\leq \ & C\sum_{i=0}^2\|\widehat{E}_h(q_h^{n-i+1/2})\|_1^4 \\
+ \ & \tfrac{Pr_\theta}{18}\|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|^2\,.
\end{aligned}
$$

Collecting these estimates in $(15)_1$, we obtain

$$
\begin{aligned}
(\mathcal{D}\boldsymbol{\zeta}_h^n, \ & \boldsymbol{\zeta}_h^{n+1/2}) + \tfrac{11Pr_\theta}{18}\|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|^2 \\
+ \ & Sd(\mathcal{I}(\mathbf{B}_h^{n+1/2}), \boldsymbol{\xi}_h^{n+1/2}, \boldsymbol{\zeta}_h^{n+1/2}) \\
\leq \ & C[\|\mathbf{f}_1^{n+1/2}\|_{-1}^2 + \|\mathcal{D}E_h(\mathbf{g}_h^n)\|_{-1}^2 \\
+ \ & \|\mathbf{g}_h^{n+1/2}\|_{\frac{1}{2},\Gamma}^2 + \sum_{i=1}^2\|\widetilde{q}_h^{n-i+1/2}\|_{\frac{1}{2},\Gamma}^2 \\
+ \ & \sum_{i=0}^2(\|q_h^{n-i+1/2}\|_{\frac{1}{2},\Gamma}^4 + \|\mathbf{g}_h^{n-i+1/2}\|_{\frac{1}{2},\Gamma}^4)] \\
+ \ & \frac{9}{2Pr_\theta}\|\mathcal{I}(\chi_h^{n+1/2})\|^2 \\
- \ & c_1(\mathcal{I}(\boldsymbol{\zeta}_h^{n+1/2}), E_h(\mathbf{g}_h^{n+1/2}), \boldsymbol{\zeta}_h^{n+1/2}) \\
- \ & Sd(\mathcal{I}(\boldsymbol{\xi}_h^{n+1/2}), \widehat{E}_h(q_h^{n+1/2}), \boldsymbol{\zeta}_h^{n+1/2})\,.
\end{aligned} \tag{16}
$$

We employ similar arguments to bound the terms on the right-hand-side of $(15)_2$ and $(15)_3$ to obtain

$$
\begin{aligned}
&(\mathcal{D}\boldsymbol{\xi}_h^n\,,\,\boldsymbol{\xi}_h^{n+1/2}) + \tfrac{Pr_B}{2}[\|\nabla\times\boldsymbol{\xi}_h^{n+1/2}\|^2\\
&\quad + \;\|\nabla\cdot\boldsymbol{\xi}_h^{n+1/2}\|^2]\\
&\quad + \;d(\boldsymbol{\zeta}_h^{n+1/2},\boldsymbol{\xi}_h^{n+1/2},\mathcal{I}(\mathbf{B}_h^{n+1/2}))\\
&\quad \le \;C[\|\mathbf{k}^{n+1/2}\|^2_{-\frac{1}{2},\Gamma}\\
&\quad + \;\|q_h^{n+1/2}\|^2_{\frac{1}{2},\Gamma} + \|\mathcal{D}\widehat{E}_h(q_h^n)\|^2_{-1}\\
&\quad + \;\|\mathbf{g}_h^{n+1/2}\|^4_{\frac{1}{2},\Gamma} + \sum_{i=1}^{2}\|q_h^{n-i+1/2}\|^4_{\frac{1}{2},\Gamma}]\\
&\quad - \;d(E_h(\mathbf{g}_h^{n+1/2}),\boldsymbol{\xi}_h^{n+1/2},\mathcal{I}(\boldsymbol{\xi}_h^{n+1/2}))
\end{aligned} \tag{17}
$$

and

$$
\begin{aligned}
&(\mathcal{D}\chi_h^n\,,\,\chi_h^{n+1/2}) + \tfrac{1}{2}\|\nabla\chi_h^{n+1/2}\|^2 \le C[\|f_2^{n+1/2}\|^2_{-1}\\
&\quad + \;\|\mathcal{D}\widetilde{E}_h(\widetilde{q}_h^n)\|^2_{-1}\\
&\quad + \;\|\widetilde{q}_h^{n+1/2}\|^2_{\frac{1}{2},\Gamma} + \|\widetilde{q}_h^{n+1/2}\|^4_{\frac{1}{2},\Gamma}\\
&\quad + \;\sum_{i=1}^{2}\|\mathbf{g}_h^{n-i+1/2}\|^4_{\frac{1}{2},\Gamma}]\\
&\quad - \;c_2(\mathcal{I}(\boldsymbol{\zeta}_h^{n+1/2}),\widetilde{E}_h(\widetilde{q}_h^{n+1/2}),\chi_h^{n+1/2})\,.
\end{aligned} \tag{18}
$$

Finally we estimate the last terms in (16)-(18) using assumption (A4) and Young's inequality to obtain

$$
\begin{aligned}
&|c_1(\;\mathcal{I}(\boldsymbol{\zeta}_h^{n+1/2}),E_h(\mathbf{g}_h^{n+1/2}),\boldsymbol{\zeta}_h^{n+1/2})|\\
&\quad \le \;\tfrac{Pr_\theta}{18}\|\nabla\boldsymbol{\zeta}_h^n\|^2\\
&\quad + \;\tfrac{Pr_\theta}{9}(\|\nabla\boldsymbol{\zeta}_h^{n-3/2}\|^2 + \|\nabla\boldsymbol{\zeta}_h^{n-1/2}\|^2)\\
&|d(\;E_h(\mathbf{g}_h^{n+1/2}),\boldsymbol{\xi}_h^{n+1/2},\mathcal{I}(\boldsymbol{\xi}_h^{n+1/2}))|\\
&\quad \le \;\tfrac{Pr_B}{8}\|\nabla\times\boldsymbol{\xi}_h^{n+1/2}\|^2\\
&\quad + \;\tfrac{Pr_B}{16}(\|\nabla\times\boldsymbol{\xi}_h^{n-3/2}\|^2 + \|\nabla\times\boldsymbol{\xi}_h^{n-1/2}\|^2)\\
&|Sd(\;\mathcal{I}(\boldsymbol{\xi}_h^n),\widehat{E}_h(q_h^{n+1/2})),\boldsymbol{\zeta}_h^{n+1/2})|\\
&\quad \le \;\tfrac{Pr_\theta}{18}\|\nabla\times\boldsymbol{\zeta}_h^{n+1/2}\|^2\\
&\quad + \;\tfrac{Pr_B S}{9}(\|\nabla\times\boldsymbol{\xi}_h^{n-3/2}\|^2 + \|\nabla\times\boldsymbol{\xi}_h^{n-1/2}\|^2)\\
&|c_2(\;\mathcal{I}(\boldsymbol{\zeta}_h^{n+1/2}),\widetilde{E}_h(\widetilde{q}_h^{n+1/2}),\chi_h^{n+1/2})|\\
&\quad \le \;\tfrac{1}{18}\|\nabla\chi_h^{n+1/2}\|^2\\
&\quad + \;\tfrac{Pr_\theta^2}{9\epsilon}(\|\nabla\boldsymbol{\zeta}_h^{n-3/2}\|^2 + \|\nabla\boldsymbol{\zeta}_h^{n-1/2}\|^2)\,,
\end{aligned} \tag{19}
$$

where ϵ is a suitably chosen positive constant. Employing these estimates in (16)-(18), we obtain

$$
\begin{aligned}
&(\mathcal{D}\boldsymbol{\zeta}_h^n\,,\,\boldsymbol{\zeta}_h^{n+1/2}) + \tfrac{Pr_\theta}{2}\|\nabla\boldsymbol{\zeta}_h^{n+1/2}\|^2\\
&\quad + \;Sd(\mathcal{I}(\mathbf{B}_h^{n+1/2}),\boldsymbol{\xi}_h^{n+1/2},\boldsymbol{\zeta}_h^{n+1/2})\\
&\quad \le \;C[\|\mathbf{f}_1^{n+1/2}\|^2_{-1}\\
&\quad + \;\|\mathcal{D}E_h(\mathbf{g}_h^n)\|^2_{-1} + \|\mathbf{g}_h^{n+1/2}\|^2_{\frac{1}{2},\Gamma}\\
&\quad + \;\sum_{i=0}^{2}(\|\overline{q}_h^{n-i}\|^4_{\frac{1}{2},\Gamma} + \|\overline{\mathbf{g}}_h^{n-i}\|^4_{\frac{1}{2},\Gamma})]\\
&\quad + \;\frac{9}{2Pr_\theta}\|\mathcal{I}(\chi_h^{n+1/2})\|^2\\
&\quad + \;\textstyle\sum_{i=1}^{2}\|\widetilde{q}_h^{n-i+1/2}\|^2_{\frac{1}{2},\Gamma}\\
&\quad + \;\tfrac{Pr_\theta}{9}(\|\boldsymbol{\zeta}_h^{n-3/2}\|_1^2 + \|\boldsymbol{\zeta}_h^{n-1/2}\|_1^2)\\
&\quad + \;\tfrac{Pr_B S}{9}(\|\boldsymbol{\xi}_h^{n-3/2}\|_1^2 + (\|\boldsymbol{\xi}_h^{n-1/2}\|_1^2)\,,
\end{aligned}
$$

$$
\begin{aligned}
&(\mathcal{D}\boldsymbol{\xi}_h^n\,,\,\boldsymbol{\xi}_h^{n+1/2}) + \tfrac{5Pr_B}{8}[\|\nabla\times\boldsymbol{\xi}_h^{n+1/2}\|^2\\
&\quad + \;\|\nabla\cdot\boldsymbol{\xi}_h^{n+1/2}\|^2]\\
&\quad + \;d(\boldsymbol{\zeta}_h^{n+1/2},\boldsymbol{\xi}_h^{n+1/2},\mathcal{I}(\mathbf{B}_h^{n+1/2}))\\
&\quad \le \;C[\|\mathbf{k}^{n+1/2}\|^2_{-\frac{1}{2},\Gamma} + \|q_h^{n+1/2}\|^2_{\frac{1}{2},\Gamma}\\
&\quad + \;\|\mathcal{D}\widehat{E}_h(q_h^n)\|^2_{-1} + \|\mathbf{g}_h^{n+1/2}\|^4_{\frac{1}{2},\Gamma}\\
&\quad + \;\sum_{i=1}^{2}\|q_h^{n-i+1/2}\|^4_{\frac{1}{2},\Gamma}] + \frac{Pr_B}{16}(\|\boldsymbol{\xi}_h^{n-3/2}\|_1^2\\
&\quad + \;\|\boldsymbol{\xi}_h^{n-1/2}\|_1^2)
\end{aligned}
$$

$$
\begin{aligned}
&(\mathcal{D}\chi_h^n\,,\,\chi_h^{n+1/2}) + \tfrac{4}{9}\|\nabla\chi_h^{n+1/2}\|^2 \le C[\|f_2^{n+1/2}\|^2_{-1}\\
&\quad + \;\|\mathcal{D}\widetilde{E}_h(\widetilde{q}_h^n)\|^2_{-1}\\
&\quad + \;\|\widetilde{q}_h^{n+1/2}\|^2_{\frac{1}{2},\Gamma} + \|\widetilde{q}_h^{n+1/2}\|^4_{\frac{1}{2},\Gamma}\\
&\quad + \;\sum_{i=1}^{2}\|\mathbf{g}_h^{n-i+1/2}\|^4_{\frac{1}{2},\Gamma})]\\
&\quad + \;\tfrac{Pr_\theta^2}{9\epsilon}(\|\nabla\boldsymbol{\zeta}_h^{n-3/2}\|^2 + \|\nabla\boldsymbol{\zeta}_h^{n-1/2}\|^2)\,.
\end{aligned} \tag{20}
$$

Now summing each of the inequalities in (20) from n = 2 to m, using the skew symmetry of $d(\cdot,\cdot,\cdot)$ and the telescoping property, we obtain that

$$\begin{aligned}
[\|\boldsymbol{\zeta}_h^m\|^2 &+ S\|\xi_h^m\|^2 + \|\chi_h^m\|^2] \\
&+ \Delta t Pr_\theta \sum_{n=2}^{m} \|\nabla \boldsymbol{\zeta}_h^{n+1/2}\|^2 \\
&+ 7\Delta t Pr_B S \lambda_m \textstyle\sum_{n=2}^m \|\boldsymbol{\xi}_h^{n+1/2}\|_1^2 \\
&+ \Delta t \textstyle\sum_{n=2}^m \|\nabla \chi_h^{n+1/2}\|^2 \le M\,,
\end{aligned} \tag{21}$$

for some constant $M > 0$ by the assumptions. The required stability bound follows by setting $(\boldsymbol{\zeta}_h^n, \boldsymbol{\xi}_h^n, \chi_h^n) = (\mathbf{u}_h^n, \mathbf{B}_h^n, \theta_h^n) - (E_h(\mathbf{g}_h^n), \widehat{E}_h(q_h^n), \widetilde{E}_h(\widetilde{q}_h^n))$ and applying triangle inequality. □

3.2. Error analysis

In this section we discuss the accuracy and convergence of the decoupled Crank-Nicolson scheme. In the subsequent analysis, we will assume the boundary data is independent of time for simplicity.

Theorem 2. *Suppose that the assumption (A1)-(A3) hold with a positive number h_0 and a positive integer k, that the solution $(\mathbf{u}, \mathbf{B}, p, \theta)$ of (1)-(3) satisfy $\mathbf{u} \in \mathcal{C}([0,T]; \mathbf{V}_g) \cap H^1(0,T;\mathbf{H}^{k+1}(\Omega)) \cap H^3(0,T;\mathbf{L}^2(\Omega))$, $\mathbf{B} \in \mathcal{C}([0,T];\mathbf{H}^1_{n,q}) \cap H^1(0,T;\mathbf{H}^{k+1}(\Omega)) \cap H^3(0,T;\mathbf{L}^2(\Omega))$, $\theta \in \mathcal{C}([0,T];H^1_{n,\widehat{q}}) \cap H^1(0,T;H^{k+1}(\Omega)) \cap H^3(0,T;L^2(\Omega))$, $p \in \mathcal{C}([0,T];L^2_0(\Omega) \cap H^k(\Omega))$ and that the initial conditions $(\mathbf{u}_h^i, \mathbf{B}_h^i, \theta_h^i)$, $i = 0,1$ satisfy $\sum_{i=0}^1 \|\mathbf{u}_h^i - \mathbf{u}(t_i)\| + S\|\mathbf{B}_h^i - \mathbf{B}(t_i)\| + \|\theta_h^i - \theta(t_i)\| \le ch^k$. Then, for any $h \in (0, h_0]$ the approximate solutions $(\mathbf{u}_h, \mathbf{B}_h, \theta_h)$ of (14) satisfy the following error estimates*

$$\|\mathbf{u} - \mathbf{u}_h\|_{l^\infty(L^2(\Omega)) \cap l^2(\mathbf{H}^1(\Omega))} \le C(\Delta t^2 + h^k)\,,$$

$$\|\mathbf{B} - \mathbf{B}_h\|_{l^\infty(L^2(\Omega)) \cap l^2(\mathbf{H}^1(\Omega))} \le C(\Delta t^2 + h^k)$$

and

$$\|\theta - \theta_h\|_{l^\infty(L^2(\Omega)) \cap l^2(H^1(\Omega))} \le C(\Delta t^2 + h^k)\,.$$

for some constant C independent of the mesh size h and time step Δt.

Proof. Let $(P_h^s \mathbf{u}(t_n), P_h^s p(t_n))$ be the Stokes projection of $(\mathbf{u}(t_n), p(t_n))$, let $P_h^m \mathbf{B}(t_n)$ be the Maxwell projection of $\mathbf{B}(t_n)$ and let $P_h^r \theta(t_n)$ be the Ritz projection of $\theta(t_n)$. Let $(\mathbf{e}_{1h}^n, e_{2h}^n, \mathbf{e}_{3h}^n, e_{4h}^n)$ be the errors defined by $\mathbf{e}_{1h}^n := \mathbf{u}_h^n - P_h^s \mathbf{u}(t_n)$, $e_{2h}^n := p_h^n - P_h^s p(t_n)$, $\mathbf{e}_{3h}^n := \mathbf{B}_h^n - P_h^m \mathbf{B}(t_n)$ and $e_{4h}^n := \theta_h^n - P_h^r \theta(t_n)$. We first subtract (1) from (14) and obtain

$$\begin{aligned}
(\mathcal{D}\mathbf{u}_h^n - \partial_t \mathbf{u}(t_{n+1/2}), \mathbf{v}_h) &+ Pr_\theta(\nabla \mathbf{u}_h^{n+1/2}, \nabla \mathbf{v}_h) \\
&+ b(\mathbf{v}_h, p_h^{n+1/2}) = < \aleph_h^n, \mathbf{v}_h > \\
&+ Pr_\theta(\nabla \mathbf{u}(t_{n+1/2}), \nabla \mathbf{v}_h) \\
&+ b(\mathbf{v}_h, p(t_{n+1/2})\,, \\
0 &= b(\mathbf{u}_h^{n+1/2} - \mathbf{u}(t_{n+1/2}), r_h)\,, \\
(\mathcal{D}\mathbf{B}_h^{n+1/2} &- \partial_t \mathbf{B}(t_{n+1/2}), \phi_h) \\
&+ Pr_B[(\nabla \times \mathbf{B}_h^{n+1/2}, \nabla \times \boldsymbol{\phi}_h) \\
&+ (\nabla \cdot \mathbf{B}_h^{n+1/2}, \nabla \cdot \boldsymbol{\phi}_h)] \\
&= Pr_B[(\nabla \times \mathbf{B}(t_{n+1/2}), \nabla \times \boldsymbol{\phi}_h) \\
&+ (\nabla \cdot \mathbf{B}(t_{n+1/2}), \nabla \cdot \boldsymbol{\phi}_h)] \\
&+ < \widehat{\aleph}_h^n, \boldsymbol{\phi}_h >\,, \\
(\mathcal{D}\theta_h^n - \partial_t \theta(t_{n+1/2}), \psi_h) &+ (\nabla \theta_h^{n+1/2}, \nabla \psi_h) \\
&= < \widetilde{\aleph}_h^n, \psi_h > + (\nabla \theta(t_{n+1/2}), \nabla \psi_h)
\end{aligned}$$

for all $\mathbf{v}_h \in \mathbf{X}_h$, $r_h \in Q_h$, $\boldsymbol{\phi}_h \in \mathbf{Y}_h$, $\psi_h \in Z_h$, at each time step n, where $\aleph_h^n$, $\widehat{\aleph}_h^n$ and $\widetilde{\aleph}_h^n$ are defined by

$$\begin{aligned}
< \aleph_h^n, \mathbf{v}_h > &:= c_1(\mathbf{u}(t_{n+1/2}), \mathbf{u}(t_{n+1/2}), \mathbf{v}_h) \\
&- c_1(\mathcal{I}(\mathbf{u}_h^{n+1/2}), \mathbf{u}_h^{n+1/2}, \mathbf{v}_h) \\
&+ e(\mathcal{I}(\theta_h^{n+1/2}) - \theta(t_{n+1/2}), \mathbf{v}_h) \\
&+ S\, d(\mathbf{B}(t_{n+1/2}), \mathbf{B}(t_{n+1/2}), \mathbf{v}_h) \\
&- S\, d(\mathcal{I}(\mathbf{B}_h^{n+1/2}), \mathbf{B}_h^{n+1/2}, \mathbf{v}_h)\,,
\end{aligned}$$

$$\begin{aligned}
< \widehat{\aleph}_h^n, \boldsymbol{\phi}_h > &:= d(\mathbf{u}(t_{n+1/2}), \boldsymbol{\phi}_h, \mathbf{B}(t_{n+1/2})) \\
&- d(\mathbf{u}_h^{n+1/2}, \boldsymbol{\phi}_h, \mathcal{I}(\mathbf{B}_h^{n+1/2}))
\end{aligned}$$

and

$$\begin{aligned}
< \widetilde{\aleph}_h^n, \psi_h > &:= c_2(\mathbf{u}(t_{n+1/2}), \theta(t_{n+1/2}), \psi_h) \\
&- c_2(\mathcal{I}(\mathbf{u}_h^{n+1/2}), \theta_h^{n+1/2}, \psi_h)\,.
\end{aligned}$$

Using the definition of Stokes, Maxwell and Ritz projections, we obtain the basic error equations of the method

$$
\begin{aligned}
(\mathcal{D}\mathbf{e}_{1h}^n\,,\ \ \mathbf{v}_h) &+ Pr_\theta(\nabla \mathbf{e}_{1h}^{n+1/2}, \nabla \mathbf{v}_h)\\
&+\ \ b(\mathbf{v}_h, e_{2h}^{n+1/2}) =< \aleph_h^n, \mathbf{v}_h >\\
&+\ \ (\partial_t \mathbf{u}(t_{n+1/2}) - \mathcal{D}P_h^s\mathbf{u}(t_n), \mathbf{v}_h)\\
&b(\mathbf{e}_{1h}^{n+1/2}, r_h) = 0\\
(\mathcal{D}\mathbf{e}_{3h}^n\,,\ \ \boldsymbol{\phi}_h) &+ Pr_B[(\nabla\times \mathbf{e}_{3h}^{n+1/2}, \nabla\times\boldsymbol{\phi}_h)\\
&+\ \ (\nabla\cdot \mathbf{e}_{3h}^{n+1/2}, \nabla\cdot\boldsymbol{\phi}_h)]\\
&=\ \ (\partial_t \mathbf{B}(t_{n+1/2}) - \mathcal{D}P_h^m\mathbf{B}(t_n), \boldsymbol{\phi}_h)\\
&+\ \ < \widehat{\aleph}_h^n, \boldsymbol{\phi}_h >
\end{aligned}
$$

$$
\begin{aligned}
(\mathcal{D}e_{4h}^n\,,\ \ \psi_h) &+ (\nabla e_{4h}^{n+1/2}, \nabla\psi_h) =< \widetilde{\aleph}_h^n, \psi_h >\\
&+\ \ (\partial_t\theta(t_{n+1/2}) - \mathcal{D}P_h^r\theta(t_n), \psi_h)\,,
\end{aligned}
\tag{22}
$$

for all $\mathbf{v}_h\in\mathbf{X}_h$, $r_h\in Q_h$, $\boldsymbol{\phi}_h\in\mathbf{Y}_h$, $\psi_h\in Z_h$,. We next split the nonlinear terms $<\aleph_h^n,\mathbf{v}_h>$, $<\widehat{\aleph}_h^n,\boldsymbol{\phi}_h>$ and $<\widetilde{\aleph}_h^n,\psi_h>$ on the right-hand side of (22) into several terms as follows:

$$
\begin{aligned}
&<\widehat{\aleph}_h^n, \boldsymbol{\phi}_h>\\
&=d((\mathbf{u}(t_{n+1/2}) - P_h^s\mathbf{u}(t_{n+1/2})), \boldsymbol{\phi}_h, \mathbf{B}(t_{n+1/2}))\\
&+d(P_h^s\mathbf{u}(t_{n+1/2}), \boldsymbol{\phi}_h, \mathbf{B}(t_{n+1/2}) - \mathcal{I}(\mathbf{B}(t_{n+1/2})))\\
&+d(P_h^s\mathbf{u}(t_{n+1/2}), \boldsymbol{\phi}_h, \mathcal{I}(\mathbf{B}(t_{n+1/2})\\
&-P_h^m\mathbf{B}(t_{n+1/2})))\\
&-d(P_h^s\mathbf{u}(t_{n+1/2})), \boldsymbol{\phi}_h, \mathcal{I}(\mathbf{e}_{3h}^{n+1/2}))\\
&-d(\mathbf{e}_{1h}^{n+1/2}, \boldsymbol{\phi}_h, \mathcal{I}(\mathbf{e}_{3h}^{n+1/2}))\\
&-d(\mathbf{e}_{1h}^{n+1/2}, \boldsymbol{\phi}_h, \mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2})))\\
&=:\textstyle\sum_{i=1}^4 < \widehat{\aleph}_i^n, \nabla\times\boldsymbol{\phi}_h >\\
&-d(\mathbf{e}_{1h}^{n+1/2}, \boldsymbol{\phi}_h, \mathcal{I}(\mathbf{e}_{3h}^{n+1/2}))\\
&-d(\mathbf{e}_{1h}^{n+1/2}, \boldsymbol{\phi}_h, \mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2})))\,,
\end{aligned}
$$

$$
\begin{aligned}
&<\widetilde{\aleph}_h^n, \psi_h>\\
&=c_2(\mathbf{u}(t_{n+1/2}), \theta(t_{n+1/2}) - P_h^r\theta(t_{n+1/2}), \psi_h)\\
&+c_2(\mathbf{u}(t_{n+1/2}) - \mathcal{I}(\mathbf{u}(t_{n+1/2}))\\
&\ ,\ P_h^r\theta(t_{n+1/2}), \psi_h)\\
&+c_2(\mathcal{I}(\mathbf{u}(t_{n+1/2})) - \mathcal{I}(P_h^s\mathbf{u}(t_{n+1/2}))\\
&\ ,\ P_h^r\theta(t_{n+1/2}), \psi_h)\\
&-c_2(\mathcal{I}(\mathbf{e}_{1h}^{n+1/2}), P_h^r\theta(t_{n+1/2}), \psi_h)\\
&-c_2(\mathcal{I}(\mathbf{e}_{1h}^{n+1/2}), e_{4h}^{n+1/2}, \psi_h)\\
&-c_2(\mathcal{I}(P_h^s\mathbf{u}(t_{n+1/2})), e_{4h}^{n+1/2}, \psi_h)\\
&=:\ \ \textstyle\sum_{i=1}^6 < \widetilde{\aleph}_i^n, \psi_h >
\end{aligned}
$$

and

$$
\begin{aligned}
&<\ \aleph_h^n, \mathbf{v}_h>\\
&=c_1(\mathbf{u}(t_{n+1/2}), \mathbf{u}(t_{n+1/2}) - P_h^s\mathbf{u}(t_{n+1/2}), \mathbf{v}_h)\\
&+c_1(\mathbf{u}(t_{n+1/2}) - \mathcal{I}(\mathbf{u}(t_{n+1/2})), P_h^s\mathbf{u}(t_{n+1/2}), \mathbf{v}_h)\\
&+c_1(\mathcal{I}(\mathbf{u}(t_{n+1/2})) - \mathcal{I}(P_h^s\mathbf{u}(t_{n+1/2}), P_h^s\mathbf{u}(t_{n+1/2}), \mathbf{v}_h)\\
&-c_1(\mathcal{I}(\mathbf{e}_{1h}^{n+1/2}), P_h^s\mathbf{u}(t_{n+1/2}), \mathbf{v}_h)\\
&-c_1(\mathcal{I}(P_h^s(\mathbf{u}(t_{n+1/2}))), \mathbf{e}_{1h}^{n+1/2}, \mathbf{v}_h)\\
&-c_1(\mathcal{I}(\mathbf{e}_{1h}^{n+1/2}), \mathbf{e}_{1h}^{n+1/2}, \mathbf{v}_h)\\
&+S(\mathbf{B}(t_{n+1/2})\times(\nabla\times(\mathbf{B}(t_{n+1/2})\\
&-P_h^m\mathbf{B}(t_{n+1/2}))), \mathbf{v}_h)\\
&+S((\mathbf{B}(t_{n+1/2}) - \mathcal{I}(\mathbf{B}(t_{n+1/2})))\\
&\times(\nabla\times P_h^m\mathbf{B}(t_{n+1/2})), \mathbf{v}_h)\\
&+S(\mathcal{I}(\mathbf{B}(t_{n+1/2}) - P_h^m\mathbf{B}(t_{n+1/2}))\\
&\times(\nabla\times P_h^m\mathbf{B}(t_{n+1/2})), \mathbf{v}_h)\\
&-S(\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\times(\nabla\times P_h^m\mathbf{B}(t_{n+1/2})), \mathbf{v}_h)\\
&+e(\mathcal{I}(e_{4h}^{n+1/2}), \mathbf{v}_h)\\
&+e(\mathcal{I}(P_h^r\theta(t_{n+1/2}) - \theta(t_{n+1/2})), \mathbf{v}_h)\\
&-S(\mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2}))\times(\nabla\times\mathbf{e}_{3h}^{n+1/2}), \mathbf{v}_h)\\
&-S(\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\times(\nabla\times\mathbf{e}_{3h}^{n+1/2}), \mathbf{v}_h)\\
&=:\textstyle\sum_{i=1}^{12} < \aleph_i^n, \mathbf{v}_h>\\
&-S(\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\times(\nabla\times\mathbf{e}_{3h}^{n+1/2}), \mathbf{v}_h)\\
&-S(\mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2}))\times(\nabla\times\mathbf{e}_{3h}^{n+1/2}), \mathbf{v}_h)\,.
\end{aligned}
$$

Notice $<\ \aleph_5^n, \mathbf{e}_{1h}^{n+1/2}\ \ >=<\ \aleph_6^n, \mathbf{e}_{1h}^{n+1/2}\ \ >=< \widetilde{\aleph}_5^n, \mathbf{e}_{4h}^{n+1/2}\ \ >=< \widetilde{\aleph}_6^n, \mathbf{e}_{4h}^{n+1/2}\ \ >= 0$ due to skew-symmetry of tri-linear forms $c_1(\cdot,\cdot,\cdot)$ and $c_2(\cdot,\cdot,\cdot)$, respectively. Therefore, setting $\mathbf{v}_h = \mathbf{e}_{1h}^{n+1/2}$,$\boldsymbol{\phi}_h = \mathbf{e}_{3h}^{n+1/2}$, $\psi_h = e_{4h}^{n+1/2}$ into (22) we can write it as

$$
\begin{cases}
(\mathcal{D}\mathbf{e}_{1h}^n, \mathbf{e}_{1h}^{n+1/2}) + Pr_\theta\|\nabla\mathbf{e}_{1h}^{n+1/2}\|^2 \\
\quad =(\partial_t\mathbf{u}(t_{n+1/2}) - \mathcal{D}P_h^s\mathbf{u}(t_n), \mathbf{e}_{1h}^{n+1/2}) \\
\quad +\sum_{i=1}^4 < \aleph_i^n, \mathbf{e}_{1h}^{n+1/2} > \\
\quad +\sum_{i=7}^{12} < \aleph_i^n, \mathbf{e}_{1h}^{n+1/2} > \\
\quad -S(\mathcal{I}(\mathbf{e}_{3h}^{n+1/2}) \times (\nabla\times\mathbf{e}_{3h}^{n+1/2}), \mathbf{e}_{1h}^{n+1/2}) \\
\quad -S(\mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2})) \times (\nabla\times\mathbf{e}_{3h}^{n+1/2}) \\
\quad , \|\mathbf{e}_{1h}^{n+1/2}) , \\
(\mathcal{D}\mathbf{e}_{3h}^n, \mathbf{e}_{3h}^{n+1/2}) + Pr_B[\|\nabla\times\mathbf{e}_{3h}^{n+1/2}\|^2 \\
\quad +\|\nabla\cdot\mathbf{e}_{3h}^{n+1/2}\|^2] \\
\quad =(\partial_t\mathbf{B}(t_{n+1/2}) - \mathcal{D}P_h^m\mathbf{B}^n, \mathbf{e}_{3h}^{n+1/2}) \\
\quad +\sum_{i=1}^4 < \widehat{\aleph}_h^n, \mathbf{e}_{3h}^{n+1/2} > , \\
\quad +(\mathbf{e}_{1h}^{n+1/2} \times \mathcal{I}(\mathbf{e}_{3h}^{n+1/2}), \nabla\times\mathbf{e}_{3h}^{n+1/2}) \\
\quad +(\mathbf{e}_{1h}^{n+1/2} \times \mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2})) \\
\quad , \nabla\times\mathbf{e}_{3h}^{n+1/2}) , \\
(\mathcal{D}e_{4h}^n, e_{4h}^{n+1/2}) + \|\nabla e_{4h}^{n+1/2}\|^2 \\
\quad =\sum_{i=1}^4 < \widetilde{\aleph}_h^n, e_{4h}^{n+1/2} > \\
\quad +(\partial_t\theta(t_{n+1/2}) - \mathcal{D}P_h^r\theta(t_n), e_{4h}^{n+1/2}) .
\end{cases}
\tag{23}
$$

By Cauchy-Schwarz inequality, triangle inequality and Lemma 2, we have

$$
\begin{aligned}
&(\partial_t\mathbf{u}(t_{n+1/2}) - \mathcal{D}P_h^s\mathbf{u}(t_n), \mathbf{e}_{1h}^{n+1/2}) \\
&\leq C\left\{(\Delta t)^{3/2}\|\partial_t^3\mathbf{u}\|_{L^2(t_n,t_{n+1};\mathbf{L}^2(\Omega))}\right. \\
&\left.+ \ \frac{h^k}{\sqrt{\Delta t}}\|(\partial_t\mathbf{u}, \partial_t p)\|_{L^2(t_n,t_{n+1};(\mathbf{H}^{k+1}\times H^k)(\Omega)}\right\} \\
&\cdot \ \|\mathbf{e}_{1h}^{n+1/2}\| ,
\end{aligned}
\tag{24}
$$

$$
\begin{aligned}
&(\partial_t\mathbf{B}(t_{n+1/2}) - \mathcal{D}P_h^m\mathbf{B}(t_n), \mathbf{e}_{3h}^{n+1/2}) \\
\leq \quad & C\left\{(\Delta t)^{3/2}\|\partial_t^3\mathbf{B}\|_{L^2(t_n,t_{n+1};\mathbf{L}^2(\Omega))}\right. \\
&\left.+ \ \frac{h^k}{\sqrt{\Delta t}}\|\partial_t\mathbf{B}\|_{L^2(t_n,t_{n+1};\mathbf{H}^{k+1}(\Omega))}\right\} \|\mathbf{e}_{3h}^{n+1/2}\|
\end{aligned}
\tag{25}
$$

and

$$
\begin{aligned}
&(\partial_t\theta(t_{n+1/2}) - \mathcal{D}P_h^r\theta(t_n), \mathbf{e}_{4h}^{n+1/2}) \\
&\leq C\left\{(\Delta t)^{3/2}\|\partial_t^3\theta\|_{L^2(t_n,t_{n+1};\mathbf{L}^2(\Omega))}\right. \\
&\left.+ \ \frac{h^k}{\sqrt{\Delta t}}\|\partial_t\theta\|_{L^2(t_n,t_{n+1};H^{k+1}(\Omega))}\right\} \|e_{4h}^{n+1/2}\| .
\end{aligned}
\tag{26}
$$

Using Hölders inequality, Gagliardo-Nirenberg inequality and Lemma 1, we obtain

$$
\begin{aligned}
&| < \aleph_1^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&\leq c^*\|\mathbf{u}(t_{n+1/2})\|_1\|\mathbf{u}(t_{n+1/2}) - P_h^s\mathbf{u}(t_{n+1/2})\|_1 \\
&\cdot\| \mathbf{e}_{1h}^{n+1/2}\|_1 \\
&\leq c^*h^k\|(\mathbf{u}, p)\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times\mathbf{H}^k)}\|\mathbf{e}_{1h}^{n+1/2}\| ,
\end{aligned}
$$

$$
\begin{aligned}
&| < \aleph_2^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&\leq c^*\|\mathbf{u}(t_{n+1/2}) - \mathcal{I}(\mathbf{u}(t_{n+1/2}))\| \\
&\cdot(\|\nabla P_h^s\mathbf{u}(t_{n+1/2})\|_{L^3} \\
&+ \|P_h^s\mathbf{u}(t_{n+1/2})\|_\infty)\|\mathbf{e}_{1h}^{n+1/2}\|_1 \\
&\leq c^*(\Delta t)^{3/2}\|\partial_t^2\mathbf{u}\|_{L^2(t_n,t_{n+1};L^2(\Omega))}\|\mathbf{e}_{1h}^{n+1/2}\|_1 ,
\end{aligned}
$$

$$
\begin{aligned}
&| < \aleph_3^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&\leq c^*\|\mathcal{I}(\mathbf{u}(t_{n+1/2})) - \mathcal{I}(P_h^s\mathbf{u}(t_{n+1/2}))\|_1 \\
&\cdot(\|P_h^s\mathbf{u}(t_{n+1/2})\|_\infty \\
&+ \|\nabla P_h^s\mathbf{u}(t_{n+1/2})\|_{L^3})\|\mathbf{e}_{1h}^{n+1/2}\| \\
&\leq c^*h^k\|(\mathbf{u}, p)\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times\mathbf{H}^k)}\|\mathbf{e}_{1h}^{n+1/2}\| ,
\end{aligned}
$$

$$
\begin{aligned}
&| < \aleph_4^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&\leq c^*\|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\| \\
&\cdot \quad (\|P_h^s\mathbf{u}(t_{n+1/2})\|_\infty + \|\nabla P_h^s\mathbf{u}(t_{n+1/2})\|_{L^3}) \\
&\cdot \quad \|\mathbf{e}_{1h}^{n+1/2}\|_1 \\
&\leq c^*(\|\mathbf{e}_{1h}^n\| + \|\mathbf{e}_{1h}^{n-1}\|)\|\mathbf{e}_{1h}^{n+1/2}\|_1 .
\end{aligned}
$$

We estimate $\aleph_7^n - \aleph_{12}^n$ using Hölders inequality, Gagliardo-Nirenberg inequality and Lemma 1 as follows

$$
\begin{aligned}
&| < \aleph_7^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&= |S(\mathbf{B}(t_{n+1/2}) \\
&\times (\nabla\times(\mathbf{B}(t_{n+1/2} - P_h^m\mathbf{B}(t_{n+1/2}))), \mathbf{e}_{1h}^{n+1/2})| \\
&\leq C\|\mathbf{B}(t_{n+1/2})\|_\infty \\
&\cdot \quad \|\mathbf{B}(t_{n+1/2}) - P_h^m\mathbf{B}(t_{n+1/2})\|_1\|\mathbf{e}_{1h}^{n+1/2}\| \\
&\leq c\, h^k\|\mathbf{B}\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1})})\|\mathbf{e}_{1h}^{n+1/2}\| ,
\end{aligned}
$$

$$
\begin{aligned}
&| < \aleph_8^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&= |S((\mathbf{B}(t_{n+1/2}) - \mathcal{I}(\mathbf{B}(t_{n+1/2}))) \\
&\times (\nabla \times P_h^m \mathbf{B}(t_{n+1/2})), \mathbf{e}_{1h}^{n+1/2})| \\
&\leq C\|\mathbf{B}(t_{n+1/2} - \mathcal{I}(\mathbf{B}(t_{n+1/2}))\| \\
&\cdot \ \ \|\nabla \times P_h^m \mathbf{B}(t_{n+1/2})\|_{L^3} \|\mathbf{e}_{1h}^{n+1/2}\|_1 \\
&\leq c^*(\Delta t)^{3/2} \|\partial_t^2 \mathbf{B}(t_{n+1/2})\|_{L^2(t_n,t_{n+1};L^2(\Omega))} \\
&\cdot \ \ \|\mathbf{e}_{1h}^{n+1/2}\|_1 \,,
\end{aligned}
$$

$$
\begin{aligned}
&| < \aleph_9^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&= |S(\mathcal{I}(\mathbf{B}(t_{n+1/2}) - P_h^m \mathbf{B}(t_{n+1/2})) \\
&\times (\nabla \times P_h^m \mathbf{B}(t_{n+1/2})), \mathbf{e}_{1h}^{n+1/2})| \\
&\leq C\|\mathcal{I}(\mathbf{B}(t_{n+1/2}) - P_h^m \mathbf{B}(t_{n+1/2}))\| \\
&\cdot \ \ |\nabla \times P_h^m \mathbf{B}(t_{n+1/2})\|_{L^3} \|\mathbf{e}_{1h}^{n+1/2}\|_1 \\
&\leq \ \ ch^k \|\mathbf{B}(t_{n+1/2})\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1})}) \\
&\cdot \ \ \|\mathbf{e}_{1h}^{n+1/2}\|_1 \,,
\end{aligned}
$$

$$
\begin{aligned}
&| < \aleph_{10}^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&= |S(\mathcal{I}(\mathbf{e}_{3h}^{n+1/2}) \\
&\times (\nabla \times P_h^m \mathbf{B}(t_{n+1/2})), \mathbf{e}_{1h}^{n+1/2})| \\
&\leq \ \ C\|\nabla P_h^m \mathbf{B}(t_{n+1/2})\|_{L^3} \\
&\cdot \ \ \|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\| \|\mathbf{e}_{1h}^{n+1/2}\|_1 \,,
\end{aligned}
$$

$$
| < \aleph_{11}^n, \mathbf{e}_{1h}^{n+1/2} > | \leq C\|\mathcal{I}(e_{4h}^{n+1/2})\| \|\mathbf{e}_{1h}^{n+1/2}\| \,,
$$

and

$$
\begin{aligned}
| < \aleph_{12}^n, \mathbf{e}_{1h}^{n+1/2} > | &\leq C\|\mathcal{I}(P_h^r \theta(t_{n+1/2}) - \theta(t_{n+1/2}))\| \\
&\cdot \ \ \|\mathbf{e}_{1h}^{n+1/2}\| \,.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&\sum_{i=1}^{4} | \ \ < \aleph_i^n, \mathbf{e}_{1h}^{n+1/2} > | + \sum_{i=7}^{12} | < \aleph_i^n, \mathbf{e}_{1h}^{n+1/2} > | \\
&\leq c\{h^k \|(\mathbf{u}, p)\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times\mathbf{H}^k)} \\
&+ h^k \|\theta\|_{\mathcal{C}([t_n,t_{n+1}];H^{k+1})} \\
&+ (\Delta t)^{3/2} \|(\partial_t^2 \mathbf{u}, \partial_t^2 \mathbf{B})\|_{L^2(t_n,t_{n+1};L^2(\Omega))} \\
&+ \|h^k \mathbf{B}\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times\mathbf{H}^k)} \\
&+ \|\mathbf{e}_{1h}^n\| + \|\mathbf{e}_{1h}^{n-1}\| \\
&+ \|\mathbf{e}_{3h}^n\| + \|\mathbf{e}_{3h}^{n-1}\| + \|e_{4h}^n\| + \|e_{4h}^{n-1}\|\} \\
&\cdot \ \ \|\mathbf{e}_{1h}^{n+1/2}\|_1 \,.
\end{aligned} \tag{27}
$$

We can estimate $\widehat{\aleph}_1^n - \widehat{\aleph}_4^n$ similarly using Hölders inequality, Gagliardo-Nirenberg inequality and Lemma 1-2 as follows

$$
\begin{aligned}
&| < \widehat{\aleph}_1^n, \mathbf{e}_{3h}^{n+1/2} > | \\
&\leq c\|\mathbf{u}(t_{n+1/2}) - P_h^s \mathbf{u}(t_{n+1/2})\| \\
&\cdot \ \ \|\mathbf{B}(t_{n+1/2})\|_\infty \|\nabla \times \mathbf{e}_{3h}^{n+1/2}\| \\
&\leq ch^k \|(\mathbf{u}, p)\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times\mathbf{H}^k)} \\
&\cdot \ \ \|\nabla \times \mathbf{e}_{3h}^{n+1/2}\| \,,
\end{aligned}
$$

$$
\begin{aligned}
&| < \widehat{\aleph}_2^n, \mathbf{e}_{3h}^{n+1/2} > | \\
&\leq c\|P_h^s \mathbf{u}(t_{n+1/2})\|_\infty \\
&\cdot \ \ \|\mathbf{B}(t_{n+1/2}) - \mathcal{I}\mathbf{B}(t_{n+1/2})\| \\
&\cdot \ \ \|\nabla \times \mathbf{e}_{3h}^{n+1/2}\| \\
&\leq \ \ c\{(\Delta t)^{3/2} \|\partial_t^2 \mathbf{B}\|_{L^2(t_n,t_{n+1};L^2(\Omega))} \\
&\cdot \ \ \|\nabla \times \mathbf{e}_{3h}^{n+1/2}\|
\end{aligned}
$$

$$
\begin{aligned}
&| < \widehat{\aleph}_3^n, \mathbf{e}_{3h}^{n+1/2} > | \\
&\leq c\|P_h^s \mathbf{u}(t_{n+1/2})\|_\infty \\
&\cdot \ \ \|\mathcal{I}(\mathbf{B}(t_{n+1/2}) - P_h^m \mathbf{B}(t_{n+1/2}))\| \\
&\cdot \ \ \|\nabla \times \mathbf{e}_{3h}^{n+1/2}\| \\
&\leq \ \ ch^k \|\mathbf{B}\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1})} \|\nabla \times \mathbf{e}_{3h}^{n+1/2}\|
\end{aligned}
$$

$$
\begin{aligned}
&| < \widehat{\aleph}_4^n, \mathbf{e}_{3h}^{n+1/2} > | \\
&\leq c\|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\| \|P_h^s \mathbf{u}(t_{n+1/2})\|_\infty \|\nabla \times \mathbf{e}_{3h}^{n+1/2})\| \,.
\end{aligned}
$$

Therefore, we have

$$\begin{aligned}\sum_{i=1}^{4} | &\ \ \widehat{\aleph}_i^n, \mathbf{e}_{3h}^{n+1/2} > | \\ &\leq c\{(\Delta t)^{3/2}\|\partial_t^2 \mathbf{B}\|_{L^2(t_n,t_{n+1};L^2(\Omega))} \\ &+ h^k \|(\mathbf{u},p)\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times H^k)} \\ &+ h^k\|\mathbf{B}\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1})}\| \\ &+ \|\mathbf{e}_{3h}^n\| + \|\mathbf{e}_{3h}^{n-1}\|\}\|\nabla\times \mathbf{e}_{3h}^{n+1/2}\|\end{aligned} \tag{28}$$

Estimating $\widetilde{\aleph}_1^n - \widetilde{\aleph}_4^n$ similarly, we obtain

$$\begin{aligned}\sum_{i=1}^{4} | &\ < \widetilde{\aleph}_i^n, e_{4h}^{n+1/2} > | \\ &\leq c\{(\Delta t)^{3/2}\|\partial_t^2 \mathbf{u}\|_{L^2(t_n,t_{n+1};L^2(\Omega))} \\ &+ h^k\|(\mathbf{u},p,\theta)\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times H^k\times H^{k+1})} \\ &+ \|\mathbf{e}_{1h}^n\| + \|\mathbf{e}_{1h}^{n-1}\|\}\|\nabla \mathbf{e}_{4h}^{n+1/2}\|\end{aligned} \tag{29}$$

Employing (24)-(29) into (23) and using Young's inequality, we obtain

$$\left\{\begin{aligned}(\mathcal{D}(\mathbf{e}_{1h}^n)\,,&\ \ \mathbf{e}_{1h}^{n+1/2}) + \tfrac{Pr_\theta}{4}\|\nabla \mathbf{e}_{1h}^{n+1/2}\|^2 \\ &\leq \Upsilon_1^n + c\left\{\|\mathbf{e}_{1h}^n\|^2 + \|\mathbf{e}_{1h}^{n-1}\|^2\right. \\ &+ \|\mathbf{e}_{3h}^n\|^2 + \|\mathbf{e}_{3h}^{n-1}\|^2 + \|\mathbf{e}_{4h}^n\|^2 \\ &+ \left.\|\mathbf{e}_{4h}^{n-1}\|^2\right\} \\ &- S(\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\times(\nabla\times \mathbf{e}_{3h}^{n+1/2}), \mathbf{e}_{1h}^{n+1/2}) \\ &- S(\mathcal{I}(P_h^m \mathbf{B}(t_{n+1/2}))\times(\nabla\times\mathbf{e}_{3h}^{n+1/2}) \\ &,\ \ \mathbf{e}_{1h}^{n+1/2}) \\ (\mathcal{D}(\mathbf{e}_{3h}^n)\,,&\ \ \mathbf{e}_{3h}^{n+1/2}) + \tfrac{Pr_B}{4}[\|\nabla\times\mathbf{e}_{3h}^{n+1/2}\|^2 \\ &+ \|\nabla\cdot\mathbf{e}_{3h}^{n+1/2}\|^2] \leq \Upsilon_2^n \\ &+ c\left\{\|\mathbf{e}_{3h}^n\|^2 + \|\mathbf{e}_{3h}^{n-1}\|^2\right\} \\ &+ (\mathbf{e}_{1h}^{n+1/2}\times\mathcal{I}(\mathbf{e}_{3h}^{n+1/2}), (\nabla\times\mathbf{e}_{3h}^{n+1/2})) \\ &+ (\mathbf{e}_{1h}^{n+1/2}\times\mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2})), \nabla\times\mathbf{e}_{3h}^{n+1/2})\,, \\ (\mathcal{D}(e_{4h}^n)\,,&\ \ e_{4h}^{n+1/2}) + \|\nabla e_{4h}^{n+1/2}\|^2 \leq \Upsilon_3^n \\ &+\ \ c\left\{\|\mathbf{e}_{1h}^n\|^2 + \|\mathbf{e}_{1h}^{n-1}\|^2\right\}\,,\end{aligned}\right. \tag{30}$$

where

$$\begin{aligned}\Upsilon_1^n := c\Big\{&(\Delta t)^3\|\partial_t^3\mathbf{u}\|^2_{L^2(t_n,t_{n+1};L^2(\Omega))} \\ &+\ \tfrac{h^{2k}}{\Delta t}\|(\partial_t\mathbf{u},\partial_t p)\|^2_{L^2(t_n,t_{n+1};\mathbf{H}^{k+1}\times\mathbf{H}^k)} \\ &+\ h^{2k}\|(\mathbf{u},p)\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times\mathbf{H}^k)} \\ &+\ (\Delta t)^3\|(\partial_t^2\mathbf{u},\partial_t^2\mathbf{B})\|^2_{L^2(t_n,t_{n+1};L^2(\Omega))} \\ &+\ h^{2k}\|\mathbf{B}\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1})}\Big\}\end{aligned}$$

$$\begin{aligned}\Upsilon_2^n := c\Big\{&(\Delta t)^3\|\partial_t^3\mathbf{B}\|^2_{L^2(t_n,t_{n+1};L^2(\Omega))} \\ &+\ \tfrac{h^{2k}}{\Delta t}\|\partial_t\mathbf{B}\|^2_{L^2(t_n,t_{n+1};\mathbf{H}^{k+1})} \\ &+\ h^{2k}\|(\mathbf{u},p)\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times\mathbf{H}^k)} \\ &+\ (\Delta t)^3\|\partial_t^2\mathbf{B}\|^2_{L^2(t_n,t_{n+1};L^2(\Omega))} \\ &+\ h^{2k}\|\mathbf{B}\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1})}\ \Big\}\,,\end{aligned}$$

$$\begin{aligned}\Upsilon_3^n := c\left\{\right.&(\Delta t)^3\|\partial_t^3\theta\|_{L^2(t_{n-1},t_{n+1};L^2(\Omega))} \\ &+\ \tfrac{h^{2k}}{\Delta t}\|\partial_t\theta\|_{L^2(t_n,t_{n+1};H^{k+1})} \\ &+\ h^{2k}\|(\mathbf{u},p,\theta)\|_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times H^k\times H^{k+1})} \\ &+\ (\Delta t)^3\|\partial_t^2\mathbf{u}\|^2_{L^2(t_n,t_{n+1};L^2(\Omega))}\left.\right\}\,.\end{aligned}$$

We next add the three equations in (30) and use the identity $(\mathbf{A}\times\mathbf{B},\nabla\times\mathbf{C}) = (\mathbf{B}\times(\nabla\times\mathbf{C}),\mathbf{A})$ to obtain

$$\begin{aligned}(\mathcal{D}(\mathbf{e}_{1h}^n)\,,&\ \ \mathbf{e}_{1h}^{n+1/2}) + S(\mathcal{D}(\mathbf{e}_{3h}^n),\mathbf{e}_{3h}^{n+1/2}) \\ &+ (\mathcal{D}(\mathbf{e}_{4h}^n),\mathbf{e}_{4h}^{n+1/2}) \\ &+ \tfrac{S\,Pr_B}{4}[\|\nabla\times\mathbf{e}_{3h}^{n+1/2}\|^2 \\ &+ \|\nabla\cdot\mathbf{e}_{3h}^{n+1/2}\|^2] + \tfrac{Pr_\theta}{4}\|\nabla\mathbf{e}_{1h}^{n+1/2}\|^2 \\ &+ \|\nabla e_{4h}^{n+1/2}\|^2 \\ &\leq c[\|\mathbf{e}_{3h}^{n-1}\|^2 + \|\mathbf{e}_{3h}^n\|^2 + \|\mathbf{e}_{1h}^{n-1}\|^2 \\ &+ \|\mathbf{e}_{1h}^n\|^2 + \|e_{4h}^{n-1}\|^2 + \|e_{4h}^n\|^2] \\ &+ \Upsilon^n\,,\end{aligned} \tag{31}$$

where

$$\begin{aligned}\Upsilon^n &:= \textstyle\sum_{i=1}^3 \Upsilon_i^n \\ &= c\Big[(\Delta t)^3\|(\partial_t^3\mathbf{u}, \partial_t^3\mathbf{B}, \partial_t^3\theta)\|^2_{L^2(t_n,t_{n+1};\mathbf{L}^2(\Omega))} \\ &+ (\Delta t)^3\|(\partial_t^2\mathbf{u}, \partial_t^2\mathbf{B})\|^2_{L^2(t_n,t_{n+1};L^2(\Omega))} \\ &+ \tfrac{h^{2k}}{\Delta t}\|(\partial_t\mathbf{u}, \partial_t\mathbf{B}, \partial_t\theta, \partial_t p)\|^2_{L^2(t_n,t_{n+1};\mathbf{H}^{k+1}\times H^k)} \\ &+ h^{2k}\|(\mathbf{u},p,\theta)\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times\mathbf{H}^k)} \\ &+ h^{2k}\|\mathbf{B}\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}}\ \Big] .\end{aligned}$$

From the assumptions on the solution $(\mathbf{u}, p, \mathbf{B}, \theta)$ it holds that

$$\Delta t\sum_{n=1}^N \Upsilon^n \le c((\Delta t)^4 + h^{2k})\,. \tag{32}$$

Therefore summing (31) from $n = 1$ to m and the discrete Grönwall inequality (Lemma 3), we have that

$$\begin{aligned}[\|\mathbf{e}_{1h}^m\|^2 &+ S\|\mathbf{e}_{3h}^m\|^2 + \|e_{4h}^m\|^2] \\ &+ Pr_\theta\Delta t\sum_{n=1}^m\|\nabla\mathbf{e}_{1h}^{n+1/2}\|^2 \\ &+ \Delta t\sum_{n=1}^m\|\nabla e_{4h}^{n+1/2}\|^2 \\ &+ S\,Pr_B\Delta t\sum_{n=1}^m[\|\nabla\times\mathbf{e}_{3h}^{n+1/2}\|^2 \\ &+ \|\nabla\cdot\mathbf{e}_{3h}^{n+1/2}\|^2] \\ &\le c((\Delta t)^4 + h^{2k})\,.\end{aligned} \tag{33}$$

The required error estimate now follows from (33) and triangle inequality. □

Theorem 3. *Under the assumptions in Theorem 2, the approximate pressure p_h of (14) satisfies*

$$\|p - p_h\|_{l^2(L^2(\Omega))} \le \frac{c}{\sqrt{\Delta t}}(\Delta t^2 + h^k)\,,$$

for some constant c independent of mesh size h and time step Δt.

Proof. From $(22)_1$ and the inf-sup condition it holds that

$$\begin{aligned}\|\mathbf{e}_{2h}^{n+1/2}\| &\le \tfrac{1}{\beta}\sup_{\mathbf{v}_h\in\mathbf{X}_h}\frac{b(\mathbf{v}_h, \mathbf{e}_{2h}^{n+1/2})}{\|\mathbf{v}_h\|_1} \\ &\le \tfrac{1}{\beta}\sup_{\mathbf{v}_h\in\mathbf{X}_h}\frac{1}{\|\mathbf{v}_h\|_1}\{-(\mathcal{D}\mathbf{e}_{1h}^n, \mathbf{v}_h) \\ &- Pr_\theta(\nabla\mathbf{e}_{1h}^{n+1/2}, \nabla\mathbf{v}_h) \\ &+ (\partial_t\mathbf{u}(t_{n+1/2}) - \mathcal{D}P_h^s\mathbf{u}(t_n), \mathbf{v}_h) \\ &+ < \aleph_h^n, \mathbf{v}_h > \\ &\le c\Big\{\|\mathcal{D}\mathbf{e}_{1h}^n\| + \|\nabla\mathbf{e}_{1h}^{n+1/2}\| \\ &+ \|\partial_t\mathbf{u}(t_{n+1/2}) - \mathcal{D}P_h^s\mathbf{u}(t_n)\|_{X_h{}^*} \\ &+ \textstyle\sum_{i=1}^{12}\|\aleph_i^n\|_{X_h{}^*} \\ &+ \|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\times(\nabla\times\mathbf{e}_{3h}^{n+1/2})\|_{X_h{}^*} \\ &+ \|\mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2})) \\ &\times(\nabla\times\mathbf{e}_{3h}^{n+1/2})\|_{X_h{}^*}\Big\}\,.\end{aligned} \tag{34}$$

We start estimating $\|\aleph_5^n\|_{X_h{}^*}$, $\|\aleph_6^n\|_{X_h{}^*}$, $\|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\times(\nabla\times\mathbf{e}_{3h}^{n+1/2})\|_{X_h{}^*}$ and $\|\mathcal{I}(P_h^m\mathbf{B}(t_{n+1/2}))\times(\nabla\times\mathbf{e}_{3h}^{n+1/2})\|_{X_h{}^*}$ below. First, by Hölder's and Gagliardo-Nirenberg inequalities, we obtain

$$\begin{aligned}| < \aleph_5^n, \mathbf{v}_h > | &\le c(\|\mathcal{I}(P_h^s(\mathbf{u}(t_{n+1/2}))\|_\infty \\ &+ \|\nabla(\mathcal{I}(P_h^s(\mathbf{u}(t_{n+1/2})))\|_{L^3}) \\ &\cdot\ \|\mathbf{e}_{1h}^{n+1/2}\|\|\mathbf{v}_h\|_1\end{aligned}$$

and

$$\begin{aligned}&| \ < \mathcal{I}(P_h^m(\mathbf{B}(t_{n+1/2}))\times(\nabla\times\mathbf{e}_{3h}^{n+1/2}), \mathbf{v}_h > | \\ &\le C\|\mathcal{I}(P_h^m(\mathbf{B}(t_{n+1/2}))\|_\infty\|\nabla\times\mathbf{e}_{3h}^{n+1/2}\|\|\mathbf{v}_h\|\,.\end{aligned}$$

Before estimating the other two terms, notice that by the inverse estimate (Assumption (A3)) and (33), we obtain

$$\begin{aligned}\|\mathbf{e}_{1h}^{n+1/2}\|_1 &\le c^*\min\{h^{-1}\|\mathbf{e}_{1h}^{n+1/2}\|, \|\mathbf{e}_{1h}^{n+1/2}\|_1\} \\ &\le c\min\{h^{-1}(\Delta t^2 + h^k) \\ &,\ (\Delta t)^{-1}(\Delta t^2 + h^k)\} \\ &\le c\,.\end{aligned} \tag{35}$$

Similarly, we can show

$$\|\nabla\times\mathbf{e}_{3h}^{n+1/2}\| \le c\,. \tag{36}$$

Therefore, by Hölder's, Gagliardo-Nirenberg inequalities and (35)-(36), we obtain

$$
\begin{aligned}
| < \aleph_6^n, \mathbf{v}_h > | &\le c \|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1 \|\mathbf{e}_{1h}^{n+1/2}\|_1 \\
&\cdot \ \|\mathbf{v}_h\|_1 \\
&\le c^* \|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1 \|\mathbf{v}_h\|_1
\end{aligned}
$$

and

$$
\begin{aligned}
&< \mathcal{I}(\mathbf{e}_{3h}^{n+1/2}) \times (\nabla \times \mathbf{e}_{3h}^{n+1/2}), \mathbf{v}_h > \\
&\qquad \le c \|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\|_1 \|\nabla \times \mathbf{e}_{3h}^{n+1/2}\| \\
&\qquad \cdot \ \|\mathbf{v}_h\|_1 \\
&\qquad \le c \|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\|_1 \|\mathbf{v}_h\|_1 \,.
\end{aligned}
$$

Estimating other terms in (34) as we did in the proof of Theorem 2, we obtain

$$
\begin{aligned}
\|\mathbf{e}_{2h}^{n+1/2}\| \le c \Big\{ &\|\mathcal{D}\mathbf{e}_{1h}^n\| + \|\nabla \mathbf{e}_{1h}^{n+1/2}\| \\
&+ \|\nabla \times \mathbf{e}_{3h}^{n+1/2}\| + \|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\| \\
&+ \|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1 + \|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\|_1 \\
&+ \|\mathcal{I}(e_{4h}^{n+1/2})\| + \|\mathbf{e}_{1h}^{n+1/2}\| \\
&+ (\Delta t)^{3/2} + h^k + \tfrac{h^k}{\sqrt{\Delta t}} \,. \Big\}
\end{aligned} \tag{37}
$$

The required error estimate now follows from last inequality by using Theorem 2 and triangle inequality. □

The error estimate for the pressure in the previous theorem can be improved under stronger regularity properties of the solution. To this end, we next derive optimal order error estimates for the time derivatives of velocity, magnetic field and temperature.

Corollary 1. *Suppose the assumptions of Theorem 2 hold. Moreover, assume* $\mathbf{u}, \mathbf{B} \in H^2(0,T;\mathbf{H}^1(\Omega))$ *and* $\theta \in H^2(0,T;H^1(\Omega))$. *In addition, assume the initial conditions* $(\mathbf{u}_h^i, \mathbf{B}_h^i, \theta_h^i)\,, i = 0,1$ *satisfy* $\sum_{i=0}^1 \|\mathbf{u}(t_i) - \mathbf{u}_h^i\|_1\,,$ $\sum_{i=0}^1 \|\mathbf{B}(t_i) - \mathbf{B}_h^i\|_1\,,$ $\sum_{i=0}^1 \|\theta(t_i) - \theta_h^i\|_1 \le ch^k$ *and* $b(\mathbf{u}_h^i, r_h) = 0\,,\quad \forall r_h \in Q_h$. *Then for any* $h \in (0,h_0]$ *the approximate velocity* $\mathbf{u}_h^n$, *magnetic field* $\mathbf{B}_h^n$ *and temperature* θ_h^n *satisfy*

$$
\|\partial_t \mathbf{u} - \mathcal{D}\mathbf{u}_h\|_{l^2(L^2(\Omega))} \le c(\Delta t^2 + h^k)\,,
$$

$$
\|\partial_t \mathbf{B} - \mathcal{D}\mathbf{B}_h\|_{l^2(L^2(\Omega))} \le c(\Delta t^2 + h^k)\,,
$$

and

$$
\|\partial_t \theta - \mathcal{D}\theta_h\|_{l^2(L^2(\Omega))} \le c(\Delta t^2 + h^k)\,,
$$

for some constant c independent of the mesh size h and time step Δt *Moreover, we have*

$$
\|\mathbf{u} - \mathbf{u}_h\|_{l^\infty(H^1(\Omega))} \le c(\Delta t^2 + h^k)\,,
$$

$$
\|\theta - \theta_h\|_{l^\infty(H^1(\Omega))} \le c(\Delta t^2 + h^k)\,,
$$

and

$$
\|\mathbf{B} - \mathbf{B}_h\|_{l^\infty(H^1(\Omega))} \le c(\Delta t^2 + h^k)
$$

for some constant c independent of the mesh size h and time step Δt.

Proof. Putting $\mathbf{v}_h = \mathcal{D}(\mathbf{e}_{1h}^n)$, $\phi_h = \mathcal{D}(\mathbf{e}_{3h}^n)$, $\psi_h = \mathcal{D}(e_{4h}^n)$ into (22) and splitting the nonlinear terms as in the proof of Theorem 2, we obtain

$$
\left\{
\begin{aligned}
&\| \ \mathcal{D}(\mathbf{e}_{1h}^n)\|^2 + Pr_\theta \mathcal{D}(\|\nabla \mathbf{e}_{1h}^n\|^2) \\
&= (\partial_t \mathbf{u}(t_{n+1/2}) - \mathcal{D}(P_h^s \mathbf{u}(t_n)), \mathcal{D}(\mathbf{e}_{1h}^n)) \\
&= \textstyle\sum_{i=1}^{14} < \aleph_i^n, \mathcal{D}(\mathbf{e}_{1h}^n) > , \\
& \\
&\| \ \mathcal{D}(\mathbf{e}_{3h}^n)\|^2 + Pr_B[\mathcal{D}(\|\nabla \times \mathbf{e}_{3h}^n\|^2) \\
&+ \mathcal{D}(\|\nabla \cdot \mathbf{e}_{3h}^n\|^2)] \\
&= (\partial_t \mathbf{B}(t_{n+1/2}) - \mathcal{D}(P_h^m \mathbf{B}(t_n)), \mathcal{D}(\mathbf{e}_{3h}^n)) \\
&+ \textstyle\sum_{i=1}^{6} < \widehat{\aleph}_i^n, \mathcal{D}(\mathbf{e}_{3h}^n > , \\
& \\
&\| \ \mathcal{D}(e_{4h}^n)\|^2 + \mathcal{D}(\|\nabla e_{4h}^n\|^2) \\
&= (\partial_t \theta(t_{n+1/2}) - \mathcal{D}(P_h^r \theta(t_n)), \mathcal{D}(e_{4h}^n)) \\
&+ \textstyle\sum_{i=1}^{6} < \widetilde{\aleph}_i^n, \mathcal{D}(e_{4h}^n) > \ .
\end{aligned}
\right. \tag{38}
$$

Let us start estimating $< \aleph_i^n, \mathcal{D}(\mathbf{e}_{1h}^n) >$ for $i = 1, \ldots, 14$. First using Hölder's inequality and Gagliardo-Nirenberg inequality, we obtain

$$
\begin{aligned}
&| \ < \aleph_1^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\
&\le c(\|\mathbf{u}(t_{n+1/2})\|_\infty + \|\nabla \mathbf{u}(t_{n+1/2})\|_{L^3}) \\
&\cdot \ \|\mathbf{u}(t_{n+1/2}) - P_h^s \mathbf{u}(t_{n+1/2})\|_1 \|\mathcal{D}(\mathbf{e}_{1h}^n)\|\,,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \aleph_2^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\
&\le c \|\mathbf{u}(t_{n+1/2}) - \mathcal{I}(\mathbf{u}(t_{n+1/2}))\|_1 \\
&\cdot \ (\|P_h^s \mathbf{u}(t_{n+1/2})\|_\infty + \|\nabla P_h^s \mathbf{u}(t_{n+1/2})\|_{L^3}) \\
&\cdot \ \|\mathcal{D}(\mathbf{e}_{1h}^n)\|\,,
\end{aligned}
$$

$$\begin{aligned}&| \ < \aleph_3^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq c(\|P_h^s\mathbf{u}(t_{n+1/2})\|_\infty + \|\nabla P_h^s\mathbf{u}(t_{n+1/2})\|_{L^3}) \\ &\cdot \ \|\mathcal{I}(\mathbf{u}(t_{n+1/2}) - P_h^s\mathbf{u}(t_{n+1/2}))\|_1\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,,\end{aligned}$$

$$\begin{aligned}&| \ < \aleph_4^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq c(\|P_h^s\mathbf{u}(t_{n+1/2})\|_\infty + \|\nabla P_h^s\mathbf{u}(t_{n+1/2})\|_{L^3}) \\ &\cdot \ \|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,,\end{aligned}$$

and

$$\begin{aligned}&| \ < \aleph_5^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq c(\|\mathcal{I}(P_h^s\mathbf{u}(t_{n+1/2}))\|_\infty + \|\nabla\mathcal{I}(P_h^s\mathbf{u}(t_{n+1/2}))\|_{L^3}) \\ &\cdot \ \|\mathbf{e}_{1h}^{n+1/2}\|_1\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,.\end{aligned}$$

From the inverse inequality (Assumption (A3)) and Gagliardo-Nirenberg inequality, it follows that

$$\|\phi_h\|_\infty + \|\nabla\phi_h\|_{L^3(\Omega)} \leq ch^{-\frac{d}{6}}\|\phi_h\|_1 \ \forall\phi_h \in X^h \,. \tag{39}$$

Using (39), we estimate $< \aleph_6^n, \mathcal{D}(\mathbf{e}_{1h}^n) >$ as below

$$\begin{aligned}&| \ < \aleph_6^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq [\|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_\infty + \|\nabla\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_{L^3}] \\ &\cdot \ \|\mathbf{e}_{1h}^{n+1/2}\|_1\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \\ &\leq c^*\|\mathbf{e}_{1h}^{n+1/2}\|_1\|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1 h^{-\frac{d}{6}} \\ &\cdot \ \|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,.\end{aligned} \tag{40}$$

Alternatively, we can estimate $< \aleph_6^n, \mathcal{D}(\mathbf{e}_{1h}^n) >$ as follows

$$\begin{aligned}&| \ < \aleph_6^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &= |\tfrac{1}{2\Delta t}c_1(\mathcal{I}(\mathbf{e}_{1h}^{n+1/2}), \mathbf{e}_{1h}^n, \mathbf{e}_{1h}^{n-1})| \\ &+ |\tfrac{1}{2\Delta t}c_1(\mathcal{I}(\mathbf{e}_{1h}^{n+1/2}), \mathbf{e}_{1h}^{n-1}, \mathbf{e}_{1h}^n)| \\ &\leq \tfrac{c^*}{\Delta t}\|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1\|\mathbf{e}_{1h}^n\|_1\|\mathbf{e}_{1h}^{n-1}\|_1 \,.\end{aligned} \tag{41}$$

Combining (40) and (41), we have

$$\begin{aligned}| < \aleph_6^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | &\leq c\gamma_n\|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1[\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \\ &+ \|\mathbf{e}_{1h}^{n-1}\|_1] \,,\end{aligned} \tag{42}$$

where

$$\gamma_n := \min\{h^{-\frac{d}{6}}, (\Delta t)^{-\frac{1}{2}}\}\|\mathbf{e}_{1h}^{n+1/2}\|_1 \,. \tag{43}$$

Estimating other terms as before, we obtain

$$\begin{aligned}&| \ < \aleph_7^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq c\|\mathbf{B}(t_{n+1/2})\|_\infty\|\mathbf{B}(t_{n+1/2}) - P_h^m\mathbf{B}(t_{n+1/2})\|_1 \\ &\cdot \ \|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,,\end{aligned}$$

$$\begin{aligned}&| \ < \aleph_8^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq c\|\mathbf{B}(t_{n+1/2}) - \mathcal{I}(\mathbf{B})(t_{n+1/2})\|_1 \\ &\cdot \ \|\nabla\times P_h^m\mathbf{B}(t_{n+1/2})\|_{L^3(\Omega)}\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,,\end{aligned}$$

$$\begin{aligned}&| \ < \aleph_9^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq c\|\mathcal{I}(\mathbf{B}(t_{n+1/2}) - P_h^m\mathbf{B}(t_{n+1/2}))\|_1 \\ &\cdot \ \|\nabla\times P_h^m\mathbf{B}\|_{L^3(\Omega)}\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,,\end{aligned}$$

$$\begin{aligned}&| \ < \aleph_{10}^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq c\|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\|_1\|\nabla\times P_h^m\mathbf{B}(t_{n+1/2})\|_{L^3(\Omega)} \\ &\cdot \ \|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,,\end{aligned}$$

$$| < \aleph_{11}^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \leq c\|\mathcal{I}(e_{4h}^{n+1/2})\|\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,,$$

$$\begin{aligned}&| \ < \aleph_{12}^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | \\ &\leq C\|\mathcal{I}(P_h^r\theta(t_{n+1/2}) - \theta(t_{n+1/2}))\| \\ &\cdot \ \|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,,\end{aligned}$$

$$\begin{aligned}| < \aleph_{13}^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | &\leq c\|\mathcal{I}(P_h^m\mathbf{B})\|_\infty \\ &\cdot \ \|\nabla\times\mathbf{e}_{3h}^{n+1/2}\|\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \,.\end{aligned}$$

Estimating as we did with $< \aleph_6^n, \mathcal{D}(\mathbf{e}_{1h}^n) >$, we get

$$\begin{aligned}| < \aleph_{14}^n, \mathcal{D}(\mathbf{e}_{1h}^n) > | &\leq c\widehat{\gamma}_n\|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\|_1 \\ &\cdot \ [\|\mathcal{D}(\mathbf{e}_{1h}^n)\| \\ &+ \textstyle\sum_{i=0}^1 \|\mathbf{e}_{1h}^{n-i}\|_1] \,,\end{aligned}$$

where

$$\widehat{\gamma}_n := \min\{h^{-\frac{d}{6}}, (\Delta t)^{-\frac{1}{2}}\}\|\mathbf{e}_{3h}^{n+1/2}\|_1 \,.$$

Let us next start estimating $\widehat{\aleph}_1 - \widehat{\aleph}_6$. First, we rewrite them using integration by parts formula and then we estimate them using Hölder's inequality and Gagliardo-Nirenberg inequality

$$
\begin{aligned}
&| \ < \widehat{\aleph}_1^n, \nabla \times \mathcal{D}(\mathbf{e}_{3h}^n) > | \\
&\leq c[\|\mathbf{B}(t_{n+1/2})\|_\infty + \|\nabla \times \mathbf{B}(t_{n+1/2})\|_{L^3}] \\
&\cdot \ \|\mathbf{u}(t_{n+1/2}) - P_h^s(\mathbf{u}(t_{n+1/2}))\|_1 \|\mathcal{D}(\mathbf{e}_{3h}^n)\|,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \widehat{\aleph}_2^n, \nabla \times \mathcal{D}(\mathbf{e}_{3h}^n) > | \\
&\leq c[\|P_h^s \mathbf{u}(t_{n+1/2})\|_\infty + \|\nabla P_h^s(\mathbf{u}(t_{n+1/2}))\|_{L^3(\Omega)}] \\
&\cdot \ \|\mathbf{B}(t_{n+1/2}) - \mathcal{I}(\mathbf{B}(t_{n+1/2}))\|_1 \\
&\cdot \ \|\mathcal{D}(\mathbf{e}_{3h}^n)\|,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \widehat{\aleph}_3^n, \nabla \times \mathcal{D}(\mathbf{e}_{3h}^n) > | \\
&\leq c[\|P_h^s \mathbf{u}\|_\infty + \|\nabla P_h^s(\mathbf{u}(t_{n+1/2})\|_{L^3(\Omega)}] \\
&\cdot \ \|\mathcal{I}(\mathbf{B}(t_{n+1/2}) - P_h^m \mathbf{B}(t_{n+1/2}))\|_1 \|\mathcal{D}(\mathbf{e}_{3h}^n)\|,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \widehat{\aleph}_4^n, \nabla \times \mathcal{D}(\mathbf{e}_{3h}^n) > | \\
&\leq c[\|P_h^s \mathbf{u}(t_{n+1/2})\|_\infty + \|\nabla P_h^s \mathbf{u}(t_{n+1/2})\|_{L^3(\Omega)}] \\
&\cdot \ \|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\|_1 \|\mathcal{D}(\mathbf{e}_{3h}^n)\|,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \widehat{\aleph}_6^n, \nabla \times \mathcal{D}(\mathbf{e}_{3h}^n) > | \\
&\leq c[\|\mathcal{I}(P_h^m \mathbf{B}(t_{n+1/2}))\|_\infty \\
&+ \|\nabla \mathcal{I}(P_h^m \mathbf{B}(t_{n+1/2}))\|_{L^3(\Omega)}] \\
&\cdot \ \|\mathbf{e}_{1h}^{n+1/2}\|_1 \|\mathcal{D}(\mathbf{e}_{3h}^n)\|.
\end{aligned}
$$

Estimating as we did with $< \aleph_{14}^n, \mathcal{D}(\mathbf{e}_{1h}^n) >$, we get

$$
\begin{aligned}
&| \ < \widehat{\aleph}_5^n, \nabla \times \mathcal{D}(\mathbf{e}_{3h}^n) > | \\
&\leq c\gamma_n \|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\|_1 [\|\mathcal{D}(\mathbf{e}_{3h}^n)\| \\
&+ \textstyle\sum_{i=0}^1 \|\mathbf{e}_{3h}^{n-i}\|_1],
\end{aligned}
$$

where γ_n is defined as in (43). Finally, we estimate $\widetilde{\aleph}_1 - \widetilde{\aleph}_6$ as follows

$$
\begin{aligned}
&| \ < \widetilde{\aleph}_1^n, \mathcal{D}(e_{4h}^n) > | \\
&\leq c(\|\mathbf{u}(t_{n+1/2})\|_\infty + \|\nabla \times \mathbf{u}(t_{n+1/2})\|_{L^3}) \\
&\cdot \ \|\theta(t_{n+1/2}) - P_h^r(\theta(t_{n+1/2}))\|_1 \|\mathcal{D}(e_{4h}^n)\|,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \widetilde{\aleph}_2^n, \mathcal{D}(e_{4h}^n) > | \\
&\leq c(\|P_h^r \theta(t_{n+1/2})\|_\infty + \|\nabla P_h^r(\theta(t_{n+1/2}))\|_{L^3(\Omega)}) \\
&\cdot \ \|\mathbf{u}(t_{n+1/2}) - \mathcal{I}(\mathbf{u}(t_{n+1/2}))\|_1 \|\mathcal{D}(e_{4h}^n)\|,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \widetilde{\aleph}_3^n, \mathcal{D}(\mathbf{e}_{3h}^n) > | \\
&\leq c\|\mathcal{I}((\mathbf{u}(t_{n+1/2}) - P_h^s(\mathbf{u}(t_{n+1/2}))\|_1 \\
&\cdot \ (\|P_h^r \theta(t_{n+1/2})\|_\infty + \|\nabla P_h^r(\theta(t_{n+1/2}))\|_{L^3(\Omega)}) \\
&\cdot \ \|\mathcal{D}(e_{4h}^n)\|,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \widetilde{\aleph}_4^n, \mathcal{D}(e_{4h}^n) > | \\
&\leq c(\|P_h^r \theta(t_{n+1/2})\|_\infty + \|\nabla P_h^r \theta(t_{n+1/2})\|_{L^3(\Omega)}) \\
&\cdot \ \|\mathcal{I}\mathbf{e}_{1h}^{n+1/2}\|_1 \|\mathcal{D}(e_{4h}^n)\|,
\end{aligned}
$$

$$
\begin{aligned}
&| \ < \widetilde{\aleph}_6^n, \mathcal{D}(e_{4h}^n) > | \\
&\leq c[\|\mathcal{I}(P_h^s \mathbf{u}(t_{n+1/2}))\|_\infty \\
&+ \|\nabla \mathcal{I}(P_h^s \mathbf{u}(t_{n+1/2}))\|_{L^3(\Omega)}] \\
&\cdot \ \|e_{4h}^{n+1/2}\|_1 \|\mathcal{D}(e_{4h}^n)\|.
\end{aligned}
$$

Estimating as we did with $< \aleph_{14}^n, \mathcal{D}(\mathbf{e}_{1h}^n) >$, we get

$$
\begin{aligned}
&| \ < \widetilde{\aleph}_5^n, \mathcal{D}(e_{4h}^n) > | \\
&\leq c\widetilde{\gamma}_n \|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1 [\|\mathcal{D}(e_{4h}^n)\| \\
&+ \|e_{4h}^{n-1}\|_1],
\end{aligned}
$$

where $\widetilde{\gamma}_n := \min\{h^{-\frac{d}{6}}, (\Delta t)^{-\frac{1}{2}}\}\|e_{4h}^{n+1/2}\|_1$. Employing these estimates in (38), we can write it as

$$
\left\{
\begin{array}{rcl}
\frac{1}{2}\|\mathcal{D}(\mathbf{e}_{1h}^n)\|^2 & + & \frac{Pr_\theta}{2}\mathcal{D}(\|\nabla \mathbf{e}_{1h}^n\|^2) \\
& \leq & c(\gamma_n^2 \|\mathcal{I}(\mathbf{e}_{1h}^{n+1/2})\|_1^2 \\
& + & \widehat{\gamma}_n^2 \|\mathcal{I}(\mathbf{e}_{3h}^{n+1/2})\|_1^2 \\
& + & \alpha_n), \\
\frac{1}{2}\|\mathcal{D}(\mathbf{e}_{3h}^n)\|^2 & + & \frac{Pr_B}{2}[\mathcal{D}(\|\nabla \times \mathbf{e}_{3h}^n\|^2) \\
& + & \mathcal{D}(\|\nabla \cdot \mathbf{e}_{3h}^n\|^2)] \\
& \leq & c\left\{\widehat{\alpha}_n + \gamma_n^2 \|\mathcal{I}(\mathbf{e}_{3h}^n)\|_1^2\right\}, \\
\frac{1}{2}\|\mathcal{D}(e_{4h}^n)\|^2 & + & \frac{1}{2}\mathcal{D}(\|\nabla e_{4h}^n\|^2) \\
& \leq & c\left\{\widetilde{\alpha}_n + \widetilde{\gamma}_n^2 \|\mathcal{I}(\mathbf{e}_{1h}^n)\|_1^2\right\},
\end{array}
\right. \tag{44}
$$

where

$$
\begin{aligned}
\alpha_n &:= (\Delta t)^3\|\partial_t^3\mathbf{u}\|^2_{L^2(t_n,t_{n+1};\mathbf{L}^2(\Omega))}\\
&+ \tfrac{h^{2k}}{\Delta t}\|(\partial_t\mathbf{u},\partial_t p)\|^2_{L^2(t_n,t_{n+1};\mathbf{H}^{k+1}\times H^k)}\\
&+ h^{2k}\|(\mathbf{u},p)\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times H^k)}\\
&+ h^{2k}\|\mathbf{B}\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1})}\\
&+ h^{2k}\|\theta\|^2_{\mathcal{C}([t_n,t_{n+1}];H^{k+1})}\\
&+ (\Delta t)^3\|\partial_t^2\mathbf{B}\|^2_{L^2(t_n,t_{n+1};\mathbf{H}^1(\Omega))}\\
&+ (\Delta t)^3\|\partial_t^2\mathbf{u}\|^2_{L^2(t_n,t_{n+1};\mathbf{H}^1(\Omega))}\\
&+ \textstyle\sum_{i=0}^1[\|\mathbf{e}_{1h}^{n-i}\|_1^2+\|\mathbf{e}_{4h}^{n-i}\|_1^2]+\|\mathbf{e}_{3h}^{n+1/2}\|_1^2\\
&+ \|\mathbf{e}_{1h}^{n+1/2}\|_1^2,\\
\widehat{\alpha}_n &:= (\Delta t)^3\|\partial_t^3\mathbf{B}\|^2_{L^2(t_n,t_{n+1};L^2(\Omega))}\\
&+ \tfrac{h^{2k}}{\Delta t}\|\partial_t\mathbf{B}\|^2_{L^2(t_n,t_{n+1};\mathbf{H}^{k+1})}\\
&+ h^{2k}\|(\mathbf{u},p)\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times H^k)}\\
&+ h^{2k}\|\mathbf{B}\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1})}\\
&+ (\Delta t)^3\|\partial_t^2\mathbf{B}\|^2_{L^2(t_n,t_{n+1};\mathbf{H}^1(\Omega))}\\
&+ \textstyle\sum_{i=0}^2[\|\mathbf{e}_{1h}^{n-i}\|_1^2+\|\mathbf{e}_{3h}^{n-i}\|_1^2],\\
\widetilde{\alpha}_n &:= h^{2k}\|\theta\|^2_{\mathcal{C}([t_n,t_{n+1}];H^{k+1})}\\
&+ (\Delta t)^3\|\partial_t^2\theta\|^2_{L^2(t_n,t_{n+1};H^1(\Omega))}\\
&+ \tfrac{h^{2k}}{\Delta t}\|\partial_t\theta\|^2_{L^2(t_n,t_{n+1};H^{k+1})}\\
&+ (\Delta t)^3\|\partial_t^3\theta\|^2_{L^2(t_n,t_{n+1};L^2(\Omega))}\\
&+ h^{2k}\|(\mathbf{u},p)\|^2_{\mathcal{C}([t_n,t_{n+1}];\mathbf{H}^{k+1}\times H^k)}\\
&+ \textstyle\sum_{i=0}^1\|\mathbf{e}_{1h}^{n-i}\|_1^2+\|e_{4h}^{n+1/2}\|_1^2.
\end{aligned}
$$

Notice that by (33) and (43), we have that

$$
\begin{aligned}
\Delta t\sum_{i=1}^N\gamma_i^2 &\le \min\{h^{-\frac{d}{3}},(\Delta t)^{-2}\}\Delta t\sum_{i=1}^N\|\mathbf{e}_{1h}^i\|_1^2\\
&\le c\min\{h^{-\frac{d}{3}},(\Delta t)^{-2}\}(h^{2k}+(\Delta t)^4)\\
&\le c\min\{h^{2k-\frac{d}{3}}+(\Delta t)^2\}\\
&\le c\,.
\end{aligned}
\tag{45}
$$

Similarly, we can show that

$$
\Delta t\sum_{i=1}^N\widehat{\gamma}_i^2\le c\quad\text{and}\quad\Delta t\sum_{i=1}^N\widetilde{\gamma}_i^2\le c\,. \tag{46}
$$

Using the regularity properties of the solution $(\mathbf{u},p,\theta)$ and (33), we obtain

$$
\begin{aligned}
&\Delta t\textstyle\sum_{i=1}^N\alpha_i\,,\quad \Delta t\sum_{i=1}^N\widehat{\alpha}_i\quad\text{and}\\
&\Delta t\textstyle\sum_{i=1}^N\widetilde{\alpha}_i\le c((\Delta t)^4+h^{2k})\,.
\end{aligned}
\tag{47}
$$

Summing (44) from $n=1$ to m and the assumptions about initial conditions $(\mathbf{u}_h^i,\mathbf{B}_h^i,\theta_h^i)$, $i=0,1$, we obtain

$$
\left\{
\begin{aligned}
\|\nabla\mathbf{e}_{1h}^m\|^2 &+ \tfrac{2}{Pr_\theta}\Delta t\textstyle\sum_{i=1}^m\|\mathcal{D}(\mathbf{e}_{1h}^n)\|^2\\
&\le c\Big\{\tfrac{4}{Pr_\theta}\Delta t\textstyle\sum_{i=1}^m\gamma_i^2\|\mathcal{I}(\mathbf{e}_{1h}^i)\|_1^2\\
&+ \tfrac{4}{Pr_\theta}\Delta t\textstyle\sum_{i=1}^m\widehat{\gamma}_i^2\|\mathcal{I}(\mathbf{e}_{3h}^i)\|_1^2\\
&+ (\Delta t)^4+h^{2k}\Big\}\,,\\
\|\nabla\times\mathbf{e}_{3h}^m\|^2 &+ \|\nabla\cdot\mathbf{e}_{3h}^m\|^2\\
&+ \tfrac{2}{Pr_B}\Delta t\textstyle\sum_{i=1}^m\|\mathcal{D}(\mathbf{e}_{3h}^n)\|^2\\
&\le c\Big\{\tfrac{4}{Pr_\theta}\Delta t\textstyle\sum_{i=1}^m\gamma_i^2\|\mathcal{I}(\mathbf{e}_{3h}^i)\|_1^2\\
&+ (\Delta t)^4+h^{2k}\Big\}\,,\\
\|\nabla e_{4h}^m\|^2 &+ 2\Delta t\textstyle\sum_{i=1}^m\|\mathcal{D}(e_{4h}^n)\|^2\\
&\le c\Big\{4\Delta t\textstyle\sum_{i=1}^m\widetilde{\gamma}_i^2\|\mathcal{I}(\mathbf{e}_{1h}^i)\|_1^2\\
&+ (\Delta t)^4+h^{2k}\Big\}\,.
\end{aligned}
\right.
\tag{48}
$$

The required results now follows from (45), (46) and (48). □

Corollary 2. *Suppose the assumptions of Corollary 3.3 hold. Then the approximate pressure $p_h^{n+1/2}$ in (14) satisfies*

$$
\|p-p_h\|_{l^2(L^2(\Omega))}\le c(\Delta t^2+h^k)\,.
$$

Proof. We provide only a sketch of the proof of this Corollary as it is similar to the proof of Theorem 2. It follows from (3.35) that

$$
\Delta t\|\mathcal{D}\mathbf{e}_{1h}^n\|^2\le c((\Delta t)^4+h^{2k})\,. \tag{49}
$$

Therefore using (49) in (37), we obtain the required estimate. □

4. Numerical results

In this section, we present a numerical example to illustrate the theoretical results of the previous section. We set $\Omega:=(0,1)\times(0,1)$ and choose the standard piecewise quadratic finite space for approximating the magnetic field and temperature. We also choose the Taylor-Hood element pair, i.e., continuous piecewise-quadratic and continuous piecewise linear finite element space for the fluid velocity and pressure approximations, respectively. Uniform triangular meshes are created by first dividing the rectangular domain Ω into identical small squares and then dividing each square into two triangles. We set the exact solutions to

$$\mathbf{u} = ((y + y^2)e^{-t}, (x + x^2)e^{-t})$$

$$\mathbf{B} = ((\sin(y) + y)e^{-t}, (\sin(x) + x^2)e^{-t})$$

$$p = (x + y)e^{-t}$$

$$\theta = (1 + xy)e^{-t}\,.$$

The right-hand side data in the MHD system, initial conditions and boundary conditions are then chosen correspondingly. For simplicity, we set the parameters Pr_θ, S, Pr_B, Ra equal to 1.0. In order to determine the order of convergence α with respect to the time step Δt, we fix the spatial spacing h and use the following approximation

$$\alpha \approx \log_2 \frac{\|\mathbf{v}_{h,\Delta t}(x,t_N) - \mathbf{v}_{h,\frac{\Delta t}{2}}(x,t_N)\|}{\|\mathbf{v}_{h,\frac{\Delta t}{2}}(x,t_N) - \mathbf{v}_{h,\frac{\Delta t}{4}}(x,t_N)\|}\,. \tag{50}$$

A set of values of α are listed in Table 4.1 with a fixed spacing $h = 1/32$ and varying time step $\Delta t = 1/20, 1/40, 1/80, 1/160, 1/320$, which clearly suggest the concerned orders of convergence in time are all $\mathcal{O}(\Delta t^2)$ for the decoupled scheme. Thus, the numerical experiments clearly suggest that the orders of convergence in time in error estimates in Theorem 2 for the L^2- norm of $\mathbf{u}$, $\mathbf{B}$ and θ are optimal.

Table 1. Convergence order of $O(\Delta t^\alpha)$ of the partitioned scheme at time $t_N = 1.0$, with the fixed spacing $h = \frac{1}{32}$

Δt	$\|\mathbf{u}(t_n) - \mathbf{u}_h^n\|$	Order
1/20	4.13475×10^{-5}	-
1/40	1.0724423×10^{-5}	1.9469
/80	0.2699941×10^{-5}	1.9899
1/160	0.0675874×10^{-5}	1.9981
1/320	0.0169062×10^{-5}	1.9992

Δt	$\|\mathbf{B}(t_n) - \mathbf{B}_h^n\|$	Order
1/20	3.92644×10^{-5}	-
1/40	0.9977026×10^{-5}	1.97654
/80	0.2512598×10^{-5}	1.98943
1/160	0.0630024×10^{-5}	1.9957
1/320	0.0157597×10^{-5}	1.99916

Δt	$\|\theta(t_n) - \theta_h^n\|$	Order
1/20	3.659835×10^{-5}	-
1/40	0.9312186×10^{-5}	1.9745867
/80	0.2344775×10^{-5}	1.98967
1/160	0.0588082×10^{-5}	1.99536
1/320	0.0147111×10^{-5}	1.99911

References

[1] Barleon, L., Casal, V., and Lenhart, L. (1991). MHD flow in liquid-metal-cooled blankets. *Fusion Engineering and Design*, 14(3-4), 401-412.

[2] Davidson, P.A. (1999). Magneto-hydrodynamics in material processing. *Annual Review of Fluid Mechanics*, 31 , 273-300.

[3] DiBenedetto, E. (1993). *Degenerate Parabolic Equations*, Springer-Verlag, New York.

[4] El-Kaddah, N., Patel, A.D. and Natarajan, T.T. (1995). The electromagnetic filtration of molten aluminum using an induced-current separator. *Journal of the Minerals, Metals and Materials Society*, 47, 46-49.

[5] Ervin, V.J. and Heuer, N. (2004). Approximation of time-dependent, viscoelastic fluid flow: Crank-Nicolson, finite element approximation. *Numer. Methods Partial Differential Equations*, 20(2), 248-283.

[6] Girault, V. and Raviart, P.A. (1986). *Finite Element Method for Navier-Stokes Equations*, Springer, Berlin.

[7] Gerbeau, J.F. (2000). A stabilized finite element methods for incompressible methods magnetohydrodynamic equations. *Numer. Math.*, 87, 83-111.

[8] Guermond, J. L. and Minev, P.D. (2003). Mixed finite element approximation of an MHD problem involving conducting and insulating regions: The 3D case. *Numer. Methods Partial Differential Equations*, 19(6), 709-731.

[9] Gunzburger, M.D., Meir, A.J. and Peterson, J. (1991). On the existence, uniqueness, and finite element approximation of solutions of the equations of stationary, incompressible magnetohydrodynamics. *Math. Comp.*, 56(194), 523-563.

[10] Gunzburger, M.D. and Peterson, J. (1983). On conforming finite element methods for the inhomogeneous stationary Navier-Stokes equations. *Numer. Math.*, 42, 173-194.

[11] Hashizume, H. (2006). Numerical and experimental research to solve MHD problem in liquid-blanket system. *Fusion Engineering and Design*, 81(8), 1431-1438.

[12] Heywood, J.G. and Rannacher, R. (1990). Finite element approximation of the non-stationary Navier-Stokes problem, Part IV: Error for second order time discretization. *SIAM Journal of Numerical Analysis*, 27(2), 353-384.

[13] Hughes, W. F. and Young, F.J. (1966). *The electromagnetodynamicos of fluids.* Wiley, New York,

[14] Ingram, R. (2013). A new linearly extrapolated Crank-Nicolson time-stepping scheme for the Navier-Stokes

equations. *Mathematics of Computation*, 82(284), 1953-1973.

[15] Julien, K., Knobloch, E. and Tobias, S. (1999). Strongly nonlinear magnetoconvection in three dimensions. *Physica D*, 128, 105-129.

[16] Kherief, K. (1984). Quelques proprietes des equations de la magnetohydrodynamique stationnaires et devolution. Th'ese, Universite de Paris VII.

[17] Layton, W., Tran, H. and Trenchea, C. (2013). Stability of partitioned methods for magnetohydrodynamics flows at small magnetic Reynolds number. *Contemp. Math.*, 586, 231-238.

[18] Lin, T.F., Gilbert, J.B. and Roy, G.D. (1991). Analyses of magnetohydrodynamic propulsion with seawater for underwater vehicles. *Journal of Propulsion and Power*, 7(6), 1081-1083.

[19] Lifschitz, A.E. (1989). *Magnetohydrodynamics and Spectral Theory*. Kluwer, Dordrecht.

[20] Meir, A.J. and Schmidt, P.G. (1996). Variational methods for stationary MHD flow under natural interface conditions. *Nonlinear Analysis*, 26(4), 659-689.

[21] Meir, A.J. and Schmidt, P.G. (1999). Analysis and numerical approximation of a stationary MHD flow problem with nonideal boundary. *SIAM J. Numer. Anal.*, 36(4), 1304-1332.

[22] Monk, P. (2003). *Finite Element Methods for Maxwell's Equations*. Oxford University Press.

[23] Prohl, A. (2008). Convergent finite element discretizations of the nonstationary incompressible magnetohydrodynamics system. *M2AN Math. Model. Numer. Anal.*, 42, 1065-1087.

[24] Rappaz, J. and Touzani, R. (1991). On a two-dimensional magnetohydrodynamic problem. I) modelling and analysis. *Modelisation mathematique et Analyse numerique*, 26(2), 347-364.

[25] Ravindran, S.S. (2015). An extrapolated second order backward difference time-stepping scheme for the magnetohydrodynamics system. *Numerical Functional Analysis and Optimization*, submitted.

[26] Schmidt, P.G. (1999). A Galerkin method for time-dependent MHD flow with nonideal boundaries. *Commun. Appl. Anal.*, 3(3), 383-398.

[27] Schonbek, M.E., Schonbek, T.P. and Soli, E. (1996). Large-time behaviour of solutions to the magnetohydrodynamics equations. *Math. Ann.*, 304, 717-756.

[28] Scott, L.R. and Zhang, S. (1990). Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54, 483-493.

[29] Sermange, M. and Temam,R. (1983). Some mathematical questions related to the MHD equations. *Comm. Pure Appl. Math.*, 36, 635-664.

[30] Shercliff, J.A. (1965). *A textbook of magnetohydrodynamics*. Pergamon Press, Oxford.

[31] Smolentsev, S., Moreau, R., Buhler, L. and Mistrangelo, C. (2010). MHD thermofluid issues of liquid metal blankets: Phenomena and advances. *Fusion Engineering and Design*, 85(7-9), 1196-1205.

[32] Spitzer, K.H., Dubke, M. and Schwertfeger, K. (1986). Rotational electromagnetic stirring in continuous casting of round strands. *Metallurgical Transactions B*, 17, 119-131.

[33] Tone, F. (2009). On the long-time H2-stability of the implicit Euler scheme for the 2D magnetohydrodynamics equations. *J. Sci. Comput.*, 38, 331-348.

[34] Utech, H.P. and Flemings, M.C. (1967). Thermal convection in metal-crystal growth [in tin and aluminium-copper alloys] : Effect of a magnetic field, Proceedings of Internat. Conf. on Crystal Growth, Boston, June 1966, J. Phys. Chem. Solids, (Suppl. 1), 651-658.

[35] WalkerJ, S. (1980). *Large interaction parameter magnetohydrodynamics and applications in fusion reactor technology*. Fluid Mechanics in Energy Conversion (J. Buckmaster, ed.), SIAM, Philadelphia.

S.S. Ravindran *is currently a professor at the University of Alabama in Huntsville, which he joined in 1999. Prior to this appointment, he was an NRC research fellow in the Flow Modeling and Control Branch at NASA Langley Research Center. Previous to that, he was a visiting assistant professor in the Center for Research in Scientific Computation at North Carolina State University. As principal investigator of various research grants, he has conducted research for agencies such as the National Science Foundation, DOD, NASA Langley Research Center and NASA Marshall Space Flight Center. He has to his credit numerous refereed publications in prestegeous journals. He has also given invited lectures in France, Austria, Spain, India, Canada and the United States, and at professional societies such as SIAM, IEEE, AIAA, and ASME. His scientific expertise has been recognized by over 1400 citations of his publications and, by invitations to consult by industry and government labs and to serve in the Editorial Boards of a number of international journals.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.1, pp.63-72 (2018)
http://doi.org/10.11121/ijocta.01.2018.00417

RESEARCH ARTICLE

New travelling wave solutions for fractional regularized long-wave equation and fractional coupled Nizhnik-Novikov-Veselov equation

Özkan Güner*

Department of International Trade, Faculty of Economics and Administrative Sciences, Çankırı Karatekin University, Çankırı, Turkey
ozkanguner@karatekin.edu.tr

ARTICLE INFO

Article History:
Received 15 November 2016
Accepted 16 June 2017
Available 23 October 2017

Keywords:
Exact solution
Ansatz method
(G'/G) −expansion method
Fractional regularized long-wave equation
Fractional coupled Nizhnik-Novikov-Veselov equation

AMS Classification 2010:
34A08, 35R11, 83C15

ABSTRACT

In this paper, solitary-wave ansatz and the (G'/G) −expansion methods have been used to obtain exact solutions of the fractional regularized long-wave (RLW) and coupled Nizhnik-Novikov-Veselov (NNV) equation. As a result, three types of exact analytical solutions such as rational function solutions, trigonometric function solutions, hyperbolic function solutions are formally derived from these equations. Proposed methods are more powerful and can be applied to other fractional differential equations arising in mathematical physics.

1. Introduction

Fractional differential equations (FDEs) are the generalized form of classical differential equations of integer order. Researchers especially in applied mathematician and physicist became highly interested in obtaining exact solutions for nonlinear FDEs in recent decades. Nonlinear FDEs are frequently used to describe many problems of physical phenomena that may arise in various fields such as biology, physics, chemistry, engineering, heat transfer, applied mathematics, control theory, mechanics, signal processing, seismic wave analysis, finance, and many other fractional dynamical systems [1-3].

In the past several decades, new exact solutions may help to find new phenomena. So, variety of powerful analytical and numerical methods for solving differential equations of fractional order have been suggested such as the adomian decomposition method, the homotopy perturbation method, the variational iteration method, the finite difference method, the differential transform method, homotopy perturbation method, the homotopy analysis method, the sub-equation method, the first integral method, the (G'/G)-expansion method, the modified trial equation method, the functional variable method, the exp-function method, the simplest equation method, the exponential rational function method, ansatz method and others [4-31].

To solve mathematical problems, the transforms are an important methods. A variety of useful transforms for solving different problems appeared in the literature, such as the traveling wave transform, the Fourier transform and the others [32-41]. Recently, Li and He [42] suggested a fractional complex transform to convert FDEs into ordinary differential equations (ODEs).

*Corresponding Author

There are different kinds of fractional derivative operators. The most famous one is the Caputo definition that the function should be differentiable [43]. Recently, Jumarie derived definitions for the fractional derivative called modified Riemann–Liouville, which are suitable for continuous and non-differentiable functions. The order α of Jumarie's derivative is defined by [44]:

$$D_w^\alpha f(w) = \begin{cases} \frac{1}{\Gamma(1-\alpha)}\frac{d}{dw}\int_0^w \frac{(f(\theta)-f(0))}{(w-\theta)^\alpha}d\theta, 0<\alpha<1 \\ (f^{(\rho)}(w))^{(\alpha-\rho)}, \rho \le \alpha < \rho+1, \rho \ge 1. \end{cases} \quad (1)$$

Some properties of the fractional modified RL derivative are [45]

$$D_w^\alpha w^r = \frac{\Gamma(1+r)}{\Gamma(1+r-\alpha)}w^{r-\alpha}, \quad (2)$$

$$D_w^\alpha (c) = 0, \quad (3)$$

$$D_w^\alpha\{af(w)+bg(w)\} = aD_w^\alpha f(w)+bD_w^\alpha g(w), \quad (4)$$

where a, b and c are constants.

We take into consideration the following general nonlinear FDE of the type

$$H(u, D_t^\alpha u, D_x^\beta u, D_t^{2\alpha}u, D_t^\alpha D_x^\beta u, D_x^{2\beta}u, ...) = 0, \quad (5)$$

where $0 < \alpha, \beta < 1$, H is a polynomial of u, u is an unknown function and D^α partial fractional derivatives of u.

The traveling wave variable

$$u(x,t) = U(\theta), \quad \theta = \frac{\varepsilon x^\alpha}{\Gamma(1+\alpha)} - \frac{\tau t^\alpha}{\Gamma(1+\alpha)}, \quad (6)$$

where $\tau \ne 0$ and $\varepsilon \ne 0$ are constants. Applying the fractional chain rule

$$D_t^\alpha u = \sigma_t \frac{dU}{d\xi}D_t^\alpha \theta \quad D_x^\alpha u = \sigma_x \frac{dU}{d\xi}D_x^\alpha \theta \quad (7)$$

where σ_t and σ_x are called the sigma indexes [46,28] and we can choose $\sigma_t = \sigma_x = L$, where L is a constant.

When we substitute, (6) with (2) and (7) into (5), we can get Eq.(5) in the following NODE;

$$\Psi(U, U', U'', U''', ..., U^{(n)}, ...) = 0, \quad (8)$$

where $U^{(n)}$ is the n^{th} derivative of U with respect to θ.

2. Description of the ansatz method for solving FDEs

For bright solitons, the starting hypothesis is in the form [47,48]

$$u(x,t) = A\,\text{sech}^p\,\theta \quad (9)$$

and

$$\theta = \frac{kx^\alpha}{\Gamma(1+\alpha)} - \frac{ct^\alpha}{\Gamma(1+\alpha)} \quad (10)$$

where A, k and c are nonzero constants. From the ansatz given above with two equalities, it is possible to obtain necessary derivatives. Then, the obtained derivatives are substituted in the Eq.(5) and we collect all terms with the same order of necessary terms. Then by equating each coefficient of the resulting polynomial to zero, we obtain a set of algebraic equations for; A, k and c. Finally solving the system of equations we can get exact solution of Eq.(5) [49-52].

2.1. Applications of the proposed method

Example 1: The space-time fractional RLW equation has the form [53]

$$D_t^\alpha u + vD_x^\alpha u + auD_x^\alpha u - \tau D_t^\alpha D_x^{2\alpha} u = 0, \quad (11)$$

where α describing the order of the fractional derivatives $0 < \alpha \le 1$ and v, a and τ are all constants that describe the behavior of the undular bore [54]. The RLW equation is modeled to govern a large number of physical phenomena such as nonlinear transverse waves in shallow water, ion acoustic and magneto hydrodynamic waves in plasma and phonon packets in nonlinear crystals. Eq.(11) was first put forward as a model for small amplitude long waves on the surface of water in channel by Peregine [55], and later by Benjamin et al. [56]. This equation is considered as an alternative to the KdV equation. Abdel-Salam and Hassan solved Eq.(11) by the fractional auxiliary sub-equation expansion method [53]. Abdel-Salam and Yousif, have obtained abundant types of exact analytical solutions including generalized trigonometric and hyperbolic functions solutions of this equation with the fractional Riccati expansion method in [57]. Analytical solutions of fractional RLW equation is very low. When $\alpha = 1$, equation (11) is called the RLW equation. Conversely, many researchers focus on numerical methods to obtain approximate solutions of RLW equation. For example, Esen and Kutluay solved the equation by a lumped Galerkin method in [58]. Dag et al. have applied least square quadratic B-spline and cubic B-spline finite element method to obtain new analytical solutions of RLW equation in [59,60]. Saka et al. [61,62] solved this equation by quintic B-spline collocation and B-spline collocation algorithms methods. In [63], the variational iteration method successfully applied to finding the solution of the RLW equation by Yusufoglu and Bekir.

In order to solve Eq.(11), we use the traveling wave transformation

$$\begin{gathered} u(x,t) = U(\theta), \\ \theta = \frac{kx^{\alpha}}{\Gamma(1+\alpha)} - \frac{ct^{\alpha}}{\Gamma(1+\alpha)}, \end{gathered} \tag{12}$$

where $k \neq 0$ and $c \neq 0$ are constants. When we substitute (12) with (2) and (6) into (11) and by integrating once and setting the constants of integration to be zero, the Eq. (11) can carry to an ODE

$$(kv - c)U + \tfrac{ak}{2}U^2 + \tau ck^2L^2U'' = 0, \tag{13}$$

where $U' = \frac{dU}{d\theta}$. The solitary wave ansatz for the bright soliton solution, the hypothesis is (9) and (10). From (9) and (10), it is possible to get

$$\frac{d^2U(\theta)}{d\theta^2} = Ap^2 \operatorname{sech}^p \theta - Ap(p+1) \operatorname{sech}^{p+2} \theta, \tag{14}$$

and

$$U^2(\theta) = A^2 \operatorname{sech}^{2p} \theta. \tag{15}$$

Thus, substituting the ansatz (14) and (15) into Eq.(13), yields to

$$\begin{aligned} &(kv - c)A \operatorname{sech}^p \theta + \frac{ak}{2} A^2 \operatorname{sech}^{2p} \theta \\ &+ \tau ck^2L^2Ap^2 \operatorname{sech}^p \theta \\ &- \tau ck^2L^2Ap(p+1) \operatorname{sech}^{p+2} \theta = 0. \end{aligned} \tag{16}$$

Now, from (16), equating the exponents $p+2$ and $2p$ leads to

$$p = 2. \tag{17}$$

From (16), setting the coefficients of $\operatorname{sech}^{p+2}\theta$ and $\operatorname{sech}^{2p}\theta$ terms to zero, we get

$$\frac{ak}{2}A^2 - \tau ck^2L^2Ap(p+1) = 0, \tag{18}$$

by use (17) and after some calculations, we have

$$A = \frac{12\tau kcL^2}{a}, \ a \neq 0. \tag{19}$$

We find, from setting the coefficients of $\operatorname{sech}^p \theta$ terms in Eq.(16) to zero

$$(kv - c)A + \tau ck^2L^2Ap^2 = 0, \tag{20}$$

also we obtain

$$c = \frac{vk}{1 - 4\tau k^2L^2}. \tag{21}$$

From (21) it is important to note that

$$4\tau k^2L^2 \neq 1. \tag{22}$$

Thus finally, bright soliton solution of (11) is given by:

$$\begin{aligned} u(x,t) \;=\; & \tfrac{12\tau kcL^2}{a} \times \\ & \operatorname{sech}^2\left(\tfrac{kx^{\alpha}}{\Gamma(1+\alpha)} - \tfrac{vkt^{\alpha}}{(1-4\tau k^2L^2)\Gamma(1+\alpha)}\right). \end{aligned} \tag{23}$$

Example 2: Secondly, we consider the following the space-time fractional coupled Nizhnik-Novikov-Veselov (NNV) equation [64]

$$\begin{aligned} &D_t^{\alpha}u - AD_x^{3\alpha}u - BD_y^{3\alpha}u + 3AuD_x^{\alpha}v \\ &\quad +3AvD_x^{\alpha}u + 3BuD_y^{\alpha}w + 3BwD_y^{\alpha}u = 0, \\ &D_x^{\alpha}u - D_y^{\alpha}v = 0, \\ &D_y^{\alpha}u - D_x^{\alpha}w = 0, \end{aligned} \tag{24}$$

where $0 < \alpha \leq 1$, A and B are given constants satisfying $A+B \neq 0$, and u, v and w are the functions of (x, y, t). Yan has found three types of travelling wave solutions of equation (24) by using the fractional sub-equation method [64]. The (2+1)-dimensional NNV equation is an isotropic extension of the well-known (1+1)-dimensional KdV equation. In recent years, NNV equation have been studied several areas of physics including condense matter physics, optics, fluid mechanics, and plasma physics when $\alpha \to 1$ [65-67]. Darvishi et al. have applied exp-function method to obtain exact traveling wave solutions of classical NNV equation in [68]. Deng solved the equation by use of the extended hyperbolic function method in

[69]. In [70], Wazwaz et al. have investigated the bright soliton solutions with wave ansatz method. For our goal, we present the following transformation

$$\begin{aligned} u(x,y,t) = U(\theta), \theta = \tfrac{kx^\alpha}{\Gamma(1+\alpha)} + \tfrac{my^\alpha}{\Gamma(1+\alpha)} - \tfrac{nt^\alpha}{\Gamma(1+\alpha)}, \\ v(x,y,t) = V(\theta), \theta = \tfrac{kx^\alpha}{\Gamma(1+\alpha)} + \tfrac{my^\alpha}{\Gamma(1+\alpha)} - \tfrac{nt^\alpha}{\Gamma(1+\alpha)}, \\ w(x,y,t) = W(\theta), \theta = \tfrac{kx^\alpha}{\Gamma(1+\alpha)} + \tfrac{my^\alpha}{\Gamma(1+\alpha)} - \tfrac{nt^\alpha}{\Gamma(1+\alpha)}, \end{aligned} \tag{25}$$

where $k \neq 0$, $m \neq 0$ and $n \neq 0$ are constants.

Then by using of Eq. (25) with (2) and (7), Eq.(24) can be turned into an ODEs and by integrating once and setting the constants of integration to be zero, we obtain

$$\begin{aligned} &(Ak^3L^2 + Bm^3L^2)U'' - 3kA(UV) \\ &\quad -3mB(UW) + nU = 0, \\ &kU - mV = 0, \\ &mU - kW = 0, \end{aligned} \tag{26}$$

where $U' = \frac{dU}{d\theta}$ and $V' = \frac{dV}{d\theta}$. In order the start off with the solution hypothesis, the following ansatsz is assumed

$$u(x,y,t) = \lambda_1 \operatorname{sech}^p \theta, \tag{27}$$

and

$$v(x,y,t) = \lambda_2 \operatorname{sech}^s \theta, \tag{28}$$

and

$$w(x,y,t) = \lambda_3 \operatorname{sech}^r \theta, \tag{29}$$

where

$$\theta = \frac{kx^\alpha}{\Gamma(1+\alpha)} + \frac{my^\alpha}{\Gamma(1+\alpha)} - \frac{nt^\alpha}{\Gamma(1+\alpha)}. \tag{30}$$

Here in (27)-(30), λ_1, λ_2, λ_3, k and m are the free parameters of the solitons and n is the velocity of the soliton. The exponents p, s and r are unknown values will be find later. Now, from (27)-(29) and (30) it is possible to obtain

$$\begin{aligned} &(Ak^3L^2 + Bm^3L^2)\lambda_1 p^2 \operatorname{sech}^p \theta \\ &- (Ak^3L^2 + Bm^3L^2)\lambda_1 p(p+1) \operatorname{sech}^{p+2} \theta \\ &- 3kA\lambda_1\lambda_2 \operatorname{sech}^{p+s} \theta - 3mB\lambda_1\lambda_3 \operatorname{sech}^{p+r} \theta \\ &+ n\lambda_1 \operatorname{sech}^p \theta = 0, \end{aligned} \tag{31}$$

and

$$k\lambda_1 \operatorname{sech}^p \theta - m\lambda_2 \operatorname{sech}^s \theta = 0, \tag{32}$$

and

$$m\lambda_1 \operatorname{sech}^p \theta - k\lambda_3 \operatorname{sech}^r \theta = 0. \tag{33}$$

Now from (32) and (33) equating the exponents of $\operatorname{sech}\theta$ functions we have $p = s = r$. In (32) we obtain,

$$\lambda_2 = \frac{k\lambda_1}{m}. \tag{34}$$

Similarly in (34) that gives

$$\lambda_3 = \frac{m\lambda_1}{k}. \tag{35}$$

Now, equating the exponents of $\operatorname{sech}^{p+2}\theta$ or $\operatorname{sech}^{p+s}\theta$ and $\operatorname{sech}^{p+r}\theta$ functions in (31) with $p = s = r$, one gets

$$p + 2 = p + s = p + r, \tag{36}$$

so that

$$p = s = r = 2. \tag{37}$$

Setting the coefficients of $\operatorname{sech}^{p+2}\theta$ in (31), to zero gives

$$\begin{aligned} (Ak^3L^2 + Bm^3L^2)\lambda_1 p(p+1) + 3kA\lambda_1\lambda_2 \\ + 3mB\lambda_1\lambda_3 = 0, \end{aligned} \tag{38}$$

using Eqs. (34), (35), $p = 2$ and some calculations

$$\lambda_1 = -2kmL^2. \tag{39}$$

Again from (31), setting the coefficients of $\operatorname{sech}^p\theta$ terms to zero one obtains,

$$(Ak^3L^2 + Bm^3L^2)\lambda_1 p^2 + n\lambda_1 = 0, \tag{40}$$

which gives

$$n = -4L^2(Ak^3 + Bm^3). \tag{41}$$

Lastly, the bright soliton solution for space-time fractional coupled NNV equation is given by

$$u(x,y,t) = \lambda_1 \operatorname{sech}^2 \theta, \tag{42}$$

and

$$v(x,y,t) = \lambda_2 \operatorname{sech}^2 \theta, \tag{43}$$

and

$$w(x,y,t) = \lambda_3 \operatorname{sech}^2 \theta, \tag{44}$$

where the velocity of the solitons n is given in (41), free parameters λ_1, λ_2 and λ_3 are given by (39), (34) and (35) respectively.

3. Description of the $\left(\frac{G'}{G}\right)$ expansion method for solving FDEs

Suppose that the solution of ODE (8) can be expressed by a polynomial in (G'/G) as:

$$U = \sum_{i=0}^{z} a_i \left(\frac{G'}{G}\right)^i, \quad a_z \neq 0, \tag{45}$$

where $G = G(\xi)$ satisfies the second order LODE in the form [71]

$$\frac{d^2 G(\xi)}{d\xi^2} + \lambda \frac{dG(\xi)}{d\xi} + \mu G(\xi) = 0, \tag{46}$$

where $a_1, ..., a_z,\ \lambda$ and μ are constants will be determined later, z is the positive integer which can be determined by the homogeneous balance with the highest order derivatives and highest order nonlinear appearing in ODE (8). When we substitute (45) into (8) and use Eq.(46), we collect all terms with the same order of (G'/G) together. When we equate all coefficient of this polynomial to zero, it gives us a set of algebraic equations for $a_1, ..., a_z,\ \lambda,\ \tau,\ \varepsilon$ and μ by using Maple. Then substituting $a_1, ..., a_z,\ \lambda,\ \mu,\ \varepsilon,\ \tau$ and general solutions of Eq. (46) into (8) we can get a variety of exact solutions of the FDEs (5).

3.1. Applications of the proposed method

Example 1:

In order to solve Eq.(11) by the (G'/G) −expansion method, we use the traveling wave transformation (12) and with a similar approach in section 2, we get

$$(kv - c)U + \frac{ak}{2}U^2 + \tau c k^2 L^2 U'' + \xi_0 = 0, \tag{47}$$

where "U'" $= \frac{dU}{d\xi}$ and ξ_0 is an integral constant. Balancing U'' with U^2 in (47) gives

$$\begin{gathered} 2z = z + 2, \\ z = 2. \end{gathered} \tag{48}$$

Assume that it is possible to express solution of (47) by a polynomial in $\left(\frac{G'}{G}\right)$ as:

$$U(\xi) = a_0 + a_1 \left(\frac{G'}{G}\right) + a_2 \left(\frac{G'}{G}\right)^2, \quad a_2 \neq 0. \tag{49}$$

By using Eq.(46), from Eq.(49) we have

$$\begin{aligned} U''(\xi) = {} & 6a_2 \left(\frac{G'}{G}\right)^4 + (2a_1 + 10a_2\lambda)\left(\frac{G'}{G}\right)^3 \\ & + (8a_2\mu + 3a_1\lambda + 4a_2\lambda^2)\left(\frac{G'}{G}\right)^2 \\ & + (6a_2\lambda\mu + 2a_1\mu + a_1\lambda^2)\left(\frac{G'}{G}\right) \\ & + 2a_2\mu^2 + a_1\lambda\mu. \end{aligned} \tag{50}$$

When we substitute Eqs.(49) and (50) into Eq.(47), collecting the coefficients of $\left(\frac{G'}{G}\right)^i$ $(i = 0, ..., 4)$ and set them to zero we get a system. The solutions of this algebraic equations can be done by Maple which gives

$$\begin{aligned} a_0 &= \frac{c - vk - \tau c k^2 L^2 \lambda^2 - 8\tau c k^2 L^2 \mu}{ak}, \\ a_1 &= -\frac{12\lambda\tau c k L^2}{a}, \\ a_2 &= -\frac{12\tau c k L^2}{a}, \\ \xi_0 &= \frac{c^2k^4\tau^2L^4\left(8\lambda^2\mu - 16\mu^2 - \lambda^4\right) + c^2 - 2vkc + v^2k^2}{2ak}. \end{aligned} \tag{51}$$

where λ and μ are arbitrary constants. By using Eq.(51), expression (49) can be written as

$$\begin{aligned} U(\xi) \; = \; & \frac{c - vk - \tau c k^2 L^2 \lambda^2 - 8\tau c k^2 L^2 \mu}{ak} \\ & - \frac{12\lambda\tau c k L^2}{a}\left(\frac{G'}{G}\right) - \frac{12\tau c k L^2}{a}\left(\frac{G'}{G}\right)^2. \end{aligned} \tag{52}$$

When we substitute general solutions of Eq. (46) into Eq.(52) we have below travelling wave solutions of the equation as follows:

When $\lambda^2 - 4\mu > 0$,

$$U_1(\xi) = \frac{c-vk+2\tau ck^2L^2(\lambda^2-4\mu)}{ak} - \frac{3\tau ck^2L^2(\lambda^2-4\mu)}{ak} \times \left(\frac{C_1 \sinh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi + C_2\cosh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi}{C_1\cosh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi + C_2\sinh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi}\right)^2, \tag{53}$$

where $\xi = \frac{kx^\alpha}{\Gamma(1+\alpha)} - \frac{ct^\alpha}{\Gamma(1+\alpha)}$.
When $\lambda^2 - 4\mu < 0$,

$$U_2(\xi) = \frac{c-vk+2\tau ck^2L^2(\lambda^2-4\mu)}{ak} + \frac{3\tau ck^2L^2(\lambda^2-4\mu)}{ak} \left(\frac{-C_1\sin\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi + C_2\cos\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi}{C_1\cos\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi + C_2\sin\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi}\right)^2, \tag{54}$$

where $\xi = \frac{kx^\alpha}{\Gamma(1+\alpha)} - \frac{ct^\alpha}{\Gamma(1+\alpha)}$.
In particular, if $C_1 \neq 0,\ C_2 = 0,\ \lambda > 0,\ \mu = 0$, then U_1 and U_2 become

$$u_1(x,t) = \frac{c-vk+2\tau ck^2L^2\lambda^2}{ak} + \frac{3\tau ckL^2\lambda^2}{a}\operatorname{sech}^2\left(\frac{\lambda kx^\alpha}{2\Gamma(1+\alpha)} - \frac{\lambda ct^\alpha}{2\Gamma(1+\alpha)}\right). \tag{55}$$

When $\lambda^2 - 4\mu = 0$, we obtain rational function solution of Eq. (45)

$$u_2(x,t) = \frac{c-vk}{ak} - \frac{12\tau ckL^2}{a}\left(\frac{C_1}{C_1\left(\frac{kx^\alpha - ct^\alpha}{\Gamma(1+\alpha)}\right)+C_2}\right)^2. \tag{56}$$

Example 2:
Similarly, in order to solve Eq. (24) by the proposed method, suppose that the solutions of the Eq. (26) can be expressed by a polynomial in $\left(\frac{G'}{G}\right)$ as follows:

$$U(\xi) = \sum_{i=0}^{z} a_i \left(\frac{G'}{G}\right)^i, \quad a_z \neq 0, \tag{57}$$

$$V(\xi) = \sum_{i=0}^{r} b_i \left(\frac{G'}{G}\right)^i, \quad b_r \neq 0, \tag{58}$$

$$W(\xi) = \sum_{i=0}^{p} c_i \left(\frac{G'}{G}\right)^i, \quad c_p \neq 0. \tag{59}$$

By the same procedure as illustrated in example 1, the homogeneous balance between highest order derivatives and non-linear terms in (26) we get positive integers $z = 2,\ r = 2$ and $p = 2$. Consequently, we have:

$$U(\xi) = a_0 + a_1\left(\frac{G'}{G}\right) + a_2\left(\frac{G'}{G}\right)^2, a_2 \neq 0, \tag{60}$$

$$V(\xi) = b_0 + b_1\left(\frac{G'}{G}\right) + b_2\left(\frac{G'}{G}\right)^2, b_2 \neq 0, \tag{61}$$

$$W(\xi) = c_0 + c_1\left(\frac{G'}{G}\right) + c_2\left(\frac{G'}{G}\right)^2, c_2 \neq 0. \tag{62}$$

When we substitute Eqs. (50), (60)-(62) into Eq.(26), collecting the coefficients of $\left(\frac{G'}{G}\right)^i$ ($i = 0, ..., 3$) and apply the same procedure of example 1, we have

Case 1:

$$a_0 = a_2\mu, \quad a_1 = a_2\lambda, \quad a_2 = a_2,$$
$$b_0 = \frac{a_2^2\mu}{2m^2L^2}, \quad b_1 = \frac{a_2^2\lambda}{2m^2L^2}, \quad b_2 = \frac{a_2^2}{2m^2L^2},$$
$$c_0 = 2m^2L^2\mu, \quad c_1 = 2m^2L^2\lambda, \quad c_2 = 2m^2L^2,$$
$$k = \frac{a_2}{2mL^2}, \quad n = \frac{(4\mu-\lambda^2)(Aa_2^3+8Bm^6L^6)}{8m^3L^4}, \tag{63}$$

where λ and μ are arbitrary constants. Substituting Eq. (63) into Eqs.(60)-(62) yields

$$U(\xi) = a_2\mu + a_2\lambda\left(\frac{G'}{G}\right) + a_2\left(\frac{G'}{G}\right)^2, \tag{64}$$

$$V(\xi) = \frac{a_2^2\mu}{2m^2L^2} + \frac{a_2^2\lambda}{2m^2L^2}\left(\frac{G'}{G}\right) + \frac{a_2^2}{2m^2L^2}\left(\frac{G'}{G}\right)^2, \tag{65}$$

$$W(\xi) = 2m^2L^2\mu + 2m^2L^2\lambda\left(\frac{G'}{G}\right) + 2m^2L^2\left(\frac{G'}{G}\right)^2 \tag{66}$$

Then, when we substitute general solutions of Eq.(46) into Eqs.(64)-(66), we have two types of solutions of the Eqs.(24) as follows:

When $\lambda^2 - 4\mu > 0$,

$$U_1(\xi) = \frac{a_2\left(4\mu-\lambda^2\right)}{4}\left(1-\Omega_1^2\right), \quad (67)$$

$$V_1(\xi) = \frac{a_2^2\left(4\mu-\lambda^2\right)}{8m^2L^2}\left(1-\Omega_1^2\right), \quad (68)$$

$$W_1(\xi) = \frac{m^2L^2\left(4\mu-\lambda^2\right)}{2}\left(1-\Omega_1^2\right) \quad (69)$$

where

$$\Omega_1 = \frac{K_1\sinh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi+K_2\cosh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi}{K_1\cosh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi+K_2\sinh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi},$$
$$\xi = \frac{a_2x^\alpha}{2mL^2\Gamma(1+\alpha)}+\frac{my^\alpha}{\Gamma(1+\alpha)}-\frac{(4\mu-\lambda^2)(Aa_2^3+8Bm^6L^6)t^\alpha}{8m^3L^4\Gamma(1+\alpha)}$$

When $\lambda^2-4\mu<0$,

$$U_2(\xi) = \frac{a_2\left(4\mu-\lambda^2\right)}{4}\left(1+\Omega_2^2\right), \quad (70)$$

$$V_2(\xi) = \frac{a_2^2\left(4\mu-\lambda^2\right)}{8m^2L^2}\left(1+\Omega_2^2\right), \quad (71)$$

$$W_2(\xi) = \frac{m^2L^2\left(4\mu-\lambda^2\right)}{2}\left(1+\Omega_2^2\right) \quad (72)$$

where

$$\Omega_2 = \frac{-K_1\sin\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi+K_2\cos\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi}{K_1\cos\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi+K_2\sin\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi}$$

In particular, if $K_1\neq 0,\ K_2=0,\ \mu=0$ then $U_1(\xi),\ V_1(\xi)$ and $W_1(\xi)$ become

$$u_1(x,y,t) = -\frac{\lambda^2a_2}{4}\operatorname{sech}^2(\Phi), \quad (73)$$

$$v_1(x,y,t) = -\frac{\lambda^2a_2^2}{8m^2L^2}\operatorname{sech}^2(\Phi), \quad (74)$$

$$w_1(x,y,t) = -\frac{\lambda^2m^2L^2}{2}\operatorname{sech}^2(\Phi), \quad (75)$$

where

$$\Phi = \frac{\lambda a_2x^\alpha}{4mL^2\Gamma(1+\alpha)}+\frac{\lambda my^\alpha}{2\Gamma(1+\alpha)}+\frac{\lambda^3(Aa_2^3+8Bm^6L^6)t^\alpha}{16m^3L^4\Gamma(1+\alpha)}$$

Also if $K_1\neq 0,\ K_2=0,\ \mu=0$ then $U_2(\xi)$, $V_2(\xi)$ and $W_2(\xi)$ become, $u_1(x,y,t)$, $v_1(x,y,t)$ and $w_1(x,y,t)$.

Case 2:

$$a_0=\frac{a_2(2\mu+\lambda^2)}{6},\quad a_1=a_2\lambda,\quad a_2=a_2,$$
$$b_0=\frac{a_2^2(2\mu+\lambda^2)}{12m^2L^2},\quad b_1=\frac{a_2^2\lambda}{2m^2L^2},\quad b_2=\frac{a_2^2}{2m^2L^2},$$
$$c_0=\frac{2m^2L^2\mu+m^2L^2\lambda^2}{3},c_1=2m^2L^2\lambda,c_2=2m^2L^2,$$
$$k=\frac{a_2}{2mL^2},\quad n=\frac{(4\mu-\lambda^2)(Aa_2^3+8Bm^6L^6)}{8m^3L^4}, \quad (76)$$

where λ and μ are arbitrary constants. Substituting Eq.(76) into Eqs.(60)-(62), yields

$$U(\xi)=\frac{a_2(2\mu+\lambda^2)}{6}+a_2\lambda\left(\frac{G'}{G}\right)+a_2\left(\frac{G'}{G}\right)^2, \quad (77)$$

$$V(\xi)=\frac{a_2^2}{2m^2L^2}\left(\frac{(2\mu+\lambda^2)}{6}+\lambda\left(\frac{G'}{G}\right)+\left(\frac{G'}{G}\right)^2\right), \quad (78)$$

$$W(\xi)=m^2L^2\left(\frac{2\mu+\lambda^2}{3}+2\lambda\left(\frac{G'}{G}\right)+2\left(\frac{G'}{G}\right)^2\right). \quad (79)$$

When we substitute general solutions of Eq.(46) into Eqs.(77)-(79), we deduce the following traveling wave solutions:
When $\lambda^2-4\mu>0$,

$$U_3(\xi)=\frac{a_2\left(4\mu-\lambda^2\right)}{4}\left(\frac{1}{3}-\Omega_1^2\right), \quad (80)$$

$$V_3(\xi)=\frac{a_2^2\left(4\mu-\lambda^2\right)}{8m^2L^2}\left(\frac{1}{3}-\Omega_1^2\right), \quad (81)$$

$$W_3(\xi)=\frac{m^2L^2\left(4\mu-\lambda^2\right)}{2}\left(\frac{1}{3}-\Omega_1^2\right). \quad (82)$$

When $\lambda^2-4\mu<0$,

$$U_4(\xi)=\frac{a_2\left(4\mu-\lambda^2\right)}{4}\left(\frac{1}{3}+\Omega_2^2\right), \quad (83)$$

$$V_4(\xi)=\frac{a_2^2\left(4\mu-\lambda^2\right)}{8m^2L^2}\left(\frac{1}{3}+\Omega_2^2\right), \quad (84)$$

$$W_4(\xi)=\frac{m^2L^2\left(4\mu-\lambda^2\right)}{2}\left(\frac{1}{3}+\Omega_2^2\right). \quad (85)$$

In particular, if $K_1 \neq 0,\ K_2 = 0,\ \mu = 0$ then $U_3(\xi),\ V_3(\xi)$ and $W_3(\xi)$ become

$$u_2(x,y,t) = -\tfrac{\lambda^2 a_2}{4}\left(\frac{1}{3} - \tanh^2(\Phi)\right), \qquad (86)$$

$$v_2(x,y,t) = -\tfrac{a_2^2\lambda^2}{8m^2L^2}\left(\frac{1}{3} - \tanh^2(\Phi)\right), \qquad (87)$$

$$w_2(x,y,t) = -\tfrac{m^2L^2\lambda^2}{2}\left(\frac{1}{3} - \tanh^2(\Phi)\right). \qquad (88)$$

Also if $K_1 \neq 0,\ K_2 = 0,\ \mu = 0$ then $U_4(\xi)$, $V_4(\xi)$ and $W_4(\xi)$ become, $u_2(x,y,t)$, $v_2(x,y,t)$ and $w_2(x,y,t)$.

4. Conclusion

The ansatz and the (G'/G) expansion methods are used in this article to obtain some new exact solutions of the fractional regularized long-wave equation and the fractional coupled Nizhnik-Novikov-Veselov equation. The (G'/G) expansion method is more effective and more general than the ansatz method because it gives exact solutions in more general forms. These methods are quite proficient methods for obtaining new exact solutions of FDEs. The obtained solutions are new and the methods can be extended to solve problems of nonlinear FDEs arising in the theory of solitons and other areas. To our knowledge, these new solutions have not been reported in former literature.

References

[1] Miller, K. S. and Ross, B. (1993). *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, New York.

[2] Podlubny, I. (1999). *Fractional Differential Equations*, Academic Press, California.

[3] Kilbas, A. A., Srivastava, H. M. and Trujillo, J. J. (2006). *Theory and Applications of Fractional Differential Equations*, Elsevier, Amsterdam .

[4] Song, L., Wang, W. (2013). A new improved Adomian decomposition method and its application to fractional differential equations. *Applied Mathematical Modelling*. 37 (3) 1590–1598.

[5] Wang, Q. (2007). Homotopy perturbation method for fractional KdV equation. *Appl. Math. Comput.*. 190 1795-802.

[6] Wu, G.C. and Baleanu, D. (2013). Variational iteration method for fractional calculus - a universal approach by Laplace transform. *Advances in Difference Equations*. 2013 18.

[7] Cui, M. (2009). Compact finite difference method for the fractional diffusion equation. *Journal of Computational Physics*. 228 (20) 7792–7804.

[8] Khan, N.A., Ara, A. and Mahmood, A. (2012). Numerical solutions of time-fractional Burgers equations: a comparison between generalized differential transformation technique and homotopy perturbation method. *Int. J. Num. Meth. Heat & Fl. Flow*. 22 (2) 175-193.

[9] Song, L. and Zhang, H. (2007). Application of homotopy analysis method to fractional KdV-Burgers-Kuramoto equation. *Phys. Lett. A*. 367 88–94.

[10] El-Ajou, A., Odibat, Z., Momani, S. and Alawneh, A. (2010). Construction of analytical solutions to fractional differential equations using homotopy analysis method. *Int. J. Appl. Math.* 40 (2) 43-51.

[11] Tian, S.F. and Zhang, H.Q. (2012). On the integrability of a generalized variable-coefficient Kadomtsev-Petviashvili equation. *Journal of Physics A: Mathematical and Theoretical*. 45, 055203. .

[12] Tian, S.F. and Zhang, H.Q. (2014). On the integrability of a generalized variable-coefficient forced Korteweg-de Vries equation in fluids. *Stud. Appl. Math.*132, 212–246.

[13] Mohyud-Din, S., Yıldırım, A. and Yülüklü, E. (2012). Homotopy analysis method for space- and time-fractional KdV equation, *Int. J. Num. Meth. Heat & Fl. Flow*. 22 (7), 928-941.

[14] Sahoo, S. and Ray, S.S. (2015). Improved fractional sub-equation method for (3+1)-dimensional generalized fractional KdV–Zakharov–Kuznetsov equations. *Computers and Mathematics with Applications*. 70, 158–166.

[15] Zhang, S. and Zhang, H.Q. (2011). Fractional sub-equation method and its applications to nonlinear fractional PDEs. *Phys. Lett. A*. 375, 1069–1073.

[16] Tian, S.F. (2017). Initial–boundary value problems for the general coupled nonlinear Schrödinger equation on the interval via the Fokas method. *J. Differential Equations*, 262, 506-558.

[17] Tian, S.F. (2016). The mixed coupled nonlinear Schrödinger equation on the half-line via the Fokas method. *Proc. R. Soc. Lond. A*. 472, 20160588.

[18] Bekir, A., Guner, O. and Unsal, O. (2015). The first integral method for exact solutions of nonlinear fractional differential equations, *J. Comput. Nonlinear Dynam.* 10(2), 021020-5.

[19] Baleanu, D., Uğurlu, Y. and Kilic, B. (2015). Improved $(G\prime/G)-$ expansion method for the time-fractional biological population model and Cahn–Hilliard equation. *J. Comput. Nonlinear Dynam.* 10 (5) 051016.

[20] Tu, J.M., Tian, S.F., Xu, M.J., Ma, P.L. and Zhang, T.T. (2016). On periodic wave solutions with asymptotic behaviors to a (3+1)-dimensional generalized B-type Kadomtsev–Petviashvili equation in fluid dynamics. *Comput. & Math. Appl.* 72, 2486–2504.

[21] Bekir, A., Guner, O., Bhrawy, A.H. and Biswas, A. (2015). Solving nonlinear fractional differential equations using exp-function and (G'/G)-expansion methods. *Rom. Journ. Phys.* 60(3-4), 360-378.

[22] Bulut, H., Baskonus, H M. and Pandir, Y. (2013). The modified trial equation method for fractional wave equation and time fractional generalized Burgers equation. *Abstract and Applied Analysis*. 636802.

[23] Guner, O. and Eser, D. (2014). Exact solutions of the space time fractional symmetric regularized long wave equation using different methods. *Advances in Mathematical Physics*. 456804 .

[24] Zhang, S., Zong Q-A., Liu, D. and Gao, Q. (2010). A generalized exp-function method for fractional riccati differential equations. *Communications in Fractional Calculus*. 1, 48-51.

[25] Guner, O., Bekir, A. and Bilgil, H. (2015). A note on exp-function method combined with complex transform method applied to fractional differential equations. *Adv. Nonlinear Anal.* 4(3), 201–208.

[26] Kaplan, M., Bekir, A., Akbulut, A. and Aksoy, E. (2015). The modified simple equation method for nonlinear fractional differential equations. *Rom. Journ. Phys.* 60(9-10), 1374–1383.

[27] Kaplan, M., Akbulut, A. and Bekir, A. (2016). Solving space-time fractional differential equations by using modified simple equation method, *Commun. Theor. Phys.* 65(5), 563–568.

[28] Aksoy, E., Kaplan, M. and Bekir, A. (2016). Exponential rational function method for space-time fractional differential equations. *Waves in Random and Complex Media*. 26(2), 142-151.

[29] Guner, O. (2015). Singular and non-topological soliton solutions for nonlinear fractional differential equations. *Chin. Phys. B*. 24(10), 100201.

[30] Guner, O. and Bekir, A. (2016). Bright and dark soliton solutions for some nonlinear fractional differential equations. *Chin. Phys. B*. 25(3), 030203.

[31] Tu, J.M., Tian, S.F., Xu, M.J. and Zhang, T.T. (2016). Quasi-periodic waves and solitary waves to a generalized KdV-Caudrey-Dodd-Gibbon equation from fluid dynamics. *Taiwanese J. Math.* 20, 823-848.

[32] Tu, J.M., Tian, S.F., Xu, M.J. and Zhang, T.T. (2016). On Lie symmetries, optimal systems and explicit solutions to the Kudryashov–Sinelshchikov equation, *Appl. Math. Comput.* 275, 345–352.

[33] Lin, S.D. and Lu, C.H. (2013). Laplace transform for solving some families of fractional differential equations and its applications. *Advances in Difference Equations*. 137.

[34] Srivastava, H.M., Golmankhaneh, A.K., Baleanu, D. and Yang, X.J. (2014). Local fractional Sumudu transform with application to IVPs on Cantor sets. *Abstract and Applied Analysis*. 620529.

[35] Wang, X.B., Tian, S.F., Xua, M.J. and Zhang T.T. (2016). On integrability and quasi-periodic wave solutions to a (3+1)-dimensional generalized KdV-like model equation. *Appl. Math. Comput.* 283, 216–233.

[36] Arnous, A.H. and Mirzazadeh, M. (2015). Bäcklund transformation of fractional Riccati equation and its applications to the space–time FDEs. *Mathematical Methods in the Applied Sciences*. 38(18), 4673–4678.

[37] Feng, L.L., Tian, S.F., Wang, X.B. and Zhang, T.T. (2017). Rogue waves, homoclinic breather waves and soliton waves for the (2+1)-dimensional B-type Kadomtsev–Petviashvili equation. *Appl. Math. Lett.* 65, 90–97.

[38] Tian, S.F., Zhang, Y.F., Feng, B.L. and Zhang, H.Q. (2015). On the Lie algebras, generalized symmetries and Darboux transformations of the fifth-order evolution equations in shallow water. *Chin. Ann. Math.* 36B, 543–560.

[39] Tian, S.F., Wang, Z. and Zhang, H.Q. (2010). Some types of solutions and generalized binary Darboux transformation for the mKP equation with self-consistent sources. *J. Math. Anal. Appl.* 366, 646-662.

[40] Al-Shara, S. (2014). Fractional transformation method for constructing solitary wave solutions to some nonlinear fractional partial differential equations. *Applied Mathematical Sciences*. 8, 5751-5762.

[41] Xu, M.J., Tian, S.F., Tu, J.M. and Zhang T.T. (2016). Bäcklund transformation, infinite conservation laws and periodic wave solutions to a generalized (2+1)-dimensional Boussinesq equation. *Nonlinear Anal.: Real World Appl.* 31, 388–408.

[42] Li, Z.B. and He, J. H. (2011). Application of the fractional complex transform to fractional differential equations. *Nonlinear Sci. Lett. A*. 2121-126.

[43] Caputo, M. (1967). Linear models of dissipation whose Q is almost frequency independent II, Geophys. *J. Royal Astronom. Soc.* 13, 529-539.

[44] Jumarie, G. (2006). Modified Riemann–Liouville derivative and fractional Taylor series of nondifferentiable functions further results. *Comput. Math. Appl.* 51, 1367–1376.

[45] Jumarie, G. (2009). Table of some basic fractional calculus formulae derived from a modified Riemann-Liouvillie derivative for nondifferentiable functions. *Appl. Maths. Lett.*. 22, 378-385.

[46] He, J H., Elegan, S.K. and Li, Z.B. (2012). Geometrical explanation of the fractional complex transform and derivative chain rule for fractional calculus. *Phys. Lett. A*. 376, 257–259.

[47] Biswas, A. (2008). 1-soliton solution of the K(m, n) equation with generalized evolution. *Phys. Lett. A*. 372, 4601-4602.

[48] Triki, H., Wazwaz, A.M. (2009). Bright and dark soliton solutions for a K(m, n) equation with t-dependent coefficients. *Phys. Lett. A*. 373, 2162–2165.

[49] Bekir, A., Guner, O. (2013). Bright and dark soliton solutions of the (3+1)-dimensional generalized Kadomtsev–Petviashvili equation and generalized Benjamin equation. *Pramana J. Phys.* 81, 203.

[50] Triki, H., Milovic, D. and Biswas, A. (2013). Solitary waves and shock waves of the KdV6 equation. *Ocean Engineering*. 73, 119–125.

[51] Younis, M. and Ali, S. (2015). Bright, dark, and singular solitons in magneto-electro-elastic circular rod. *Waves in Random and Complex Media*. 25(4), 549-555.

[52] Bekir, A. and Guner, O. (2013). Topological (dark) soliton solutions for the Camassa–Holm type equations. *Ocean Engineering*. 74, 276–279.

[53] Abdel-Salam, E.A.B., Hassan, G.F. (2016). Solutions to class of linear and nonlinear fractional differential equations. *Commun. Theor. Phys.*. 65, 127–135.

[54] Peregrine, D.H. (1966). Calculations of the development of an undular bore. *J. Fluid Mech.* 25, 321–330.

[55] Peregrine, D.H. (1967). Long waves on a beach. *J. Fluid Mech.* 27, 815–827.

[56] Benjamin, T.B., Bona, J.L. and Mahony, J. (1972). Model equations for waves in nonlinear dispersive systems. *J. Philos. Trans. R. Soc. Lond.* 227, 47–78.

[57] Abdel-Salam, E.A.B., Yousif, E.A. (2013). Solution of nonlinear space-time fractional differential equations using the fractional riccati expansion method. *Mathematical Problems in Engineering*. 846283.

[58] Esen, A. and Kutluay, S. (2006). Application of a lumped Galerkin method to the regularized long wave equation. *Appl. Math. Comput.* 174, 833–845.

[59] Dag, I. (2000). Least square quadratic B-spline finite element method for the regularized long wave equation. *Comp. Meth. Appl. Mech. Eng.* 182, 205–215.

[60] Dag, I. and Ozer, M.N. (2001). Approximation of RLW equation by least square cubic B-spline finite element method. *Appl. Math. Model.* 25, 221–231.

[61] Saka, B., Dag, I. and Irk, D. (2008). Quintic B-spline collocation method for numerical solutions of the RLW equation. *Anziam J.* 49(3), 389–410.

[62] Saka, B., Sahin, A., Dag, I. (2011). B-spline collocation algorithms for numerical solution of the RLW equation. *Numer. Meth. Part. D. E.* 27, 581–607.

[63] Yusufoglu, E. and Bekir, A. (2007). Application of the variational iteration method to the regularized long wave equation. *Computers and Mathematics with Applications.* 54, 1154–1161.

[64] Liu, Y. and Yan, L. (2013). Solutions of fractional Konopelchenko-Dubrovsky and Nizhnik-Novikov-Veselov equations using a generalized fractional subequation method. *Abstract and Applied Analysis.* 839613.

[65] Hong, T., Wang, Y.Z. and Huo, Y.S. (1998). Bogoliubov quasiparticles carried by dark solitonic excitations in nonuniform Bose Einstein condensates. *Chin. Phys. Lett.* 15, 550 552.

[66] Das, G.C. (1997). Explosion of soliton in a multicomponent plasma. *Phys. Plasmas.* 4, 2095-2100.

[67] Lou, S.Y. (1999). A direct perturbation method: Nonlinear Schrodinger equation with loss. *Chin. Phys. Lett.* 16, 659-661.

[68] Shin, B.C., Darvishi, M.T. and Barati, A. (2009). Some exact and new solutions of the Nizhnik Novikov Vesselov equation using the Exp-function method. *Computers and Mathematics with Applications.* 58, 2147-2151.

[69] Deng, C. (2010). New abundant exact solutions for the $(2+1)$-dimensional generalized Nizhnik–Novikov–Veselov system. *Commun Nonlinear Sci Numer Simulat.* 15, 3349–3357

[70] Boubir, B., Triki, H. and Wazwaz, A.M. (2013). Bright solitons of the variants of the Novikov–Veselov equation with constant and variable coefficients. *Applied Mathematical Modelling.* 37, 420–431.

[71] Wang, M.L., Li X. and Zhang, J. (2008). The (G'/G)-expansion method and travelling wave solutions of nonlinear evolution equations in mathematical physics. *Phys. Lett. A.* 372(4), 417-423.

Özkan Güner *is currently an assistant professor at Cankiri Karatekin University. He obtained his Ph.D. degree from Eskisehir Osmangazi University. He has published more than 70 articles in ISI-listed journals with an h-index of 11. His current research interest mainly covers in exact solution of nonlinear partial and fractional differential equations.*

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.1, pp.73-83 (2018)
https://doi.org/10.11121/ijocta.01.2018.00412

RESEARCH ARTICLE

Integrated process planning, WSPT scheduling and WSLK due-date assignment using genetic algorithms and evolutionary strategies

Halil Ibrahim Demir [a] *, Onur Canpolat [a]

[a] *Department of Industrial Engineering, Sakarya University, Turkey*
hidemir@sakarya.edu.tr, onurcanpolat@sakarya.edu.tr

ARTICLE INFO

Article history:
Received: 7 November 2016
Accepted: 7 August 2017
Available Online: 2 November 2017

Keywords:
Process planning
Weighted scheduling
Weighted due-date assignment
Genetic algorithm
Random search

AMS Classification 2010:
90-08, 90B35, 78M32, 90C59

ABSTRACT

Process planning, scheduling and due-date assignment are three important manufacturing functions in our life. They all try to get local optima and there can be an enormous loss in overall performance value if they are handled separately. That is why they should be handled concurrently. Although integrated process planning and scheduling with due date assignment problem is not addressed much in the literature, there are numerous works on integrated process planning and scheduling and many works on scheduling with due date assignment. Most of the works in the literature assign common due date for the jobs waiting and due dates are determined without taking into account of the weights of the customers. Here process planning function is integrated with weighted shortest processing times (WSPT) scheduling and weighted slack (WSLK) due date assignment. In this study unique due dates are given to each customer and important customers get closer due dates. Integration of these three functions is tested for different levels of integration with genetic algorithms, evolutionary strategies, hybrid genetic algorithms, hybrid evolutionary strategies and random search techniques. Best combinations are found as full integration with genetic search and hybrid genetic search. Integration of these three functions provided substantial improvements in global performance.

1. Introduction

Traditionally process planning, scheduling and due date assignment are treated sequentially and separately. Independently predetermined process planning, scheduling and due date assignment can cause poor global performance and can be a poor input to the downstream functions. For example, independently predetermined process plans can be poor input to scheduling. Process planners can select same desired machines repeatedly, thus some machines may be starving. In this case, these plans may not be followed at the shop floor level. Independently predetermined scheduling without taking into consideration the due dates, may worsen global performance. As we may unnecessarily increase earliness and tardiness of some jobs. Independently given due dates can be unrealistic for the shop floor. They may be determined either too early or too late in which can worsen the production performance because of unnecessary tardiness, earliness or due dates.

Meanwhile, we should consider the importance of each customer because we may unnecessarily give very close due dates for the unimportant customer and give far due dates for very important customers. This situation may cause poor performance.

If we look at the literature we can find numerous works on integrated process planning and scheduling (IPPS) and scheduling with due date assignment (SWDDA). But if we search for works on integrated process planning, scheduling and due date assignment (IPPSDDA), only a few studies were found in the literature.

Merely scheduling part of the problem already belongs to the class of NP-hard problems, that is why heuristic solutions are required to solve the problem. We cannot find the optimum solution to the problem in a reasonable amount of time when it gets larger. That is why heuristic methods are used to solve this problem. When we integrate three functions, the problem becomes even more complex. For this reason, we applied genetic algorithms, evolutionary strategies and random search in the solution of the integrated

*Corresponding author

problem.

As we integrate more functions together overall solution becomes better and to prove this claim we integrated each function step by step. Finally, we integrated weighted shortest processing times (WSPT) scheduling with weighted slack (WSLK) due date assignment and process plan selection. We used WSPT because it is a popular dispatching rule that schedules shorter and important jobs first. Similarly, WSLK is a common due date assignment technique which adds slack to the processing times. In addition, we took into account weight of each job so that close due dates are given for important jobs and scheduled first and far due dates are given for unimportant jobs and scheduled later.

We used genetic algorithms (GA), hybrid genetic algorithms (R-GA), evolutionary strategies (ES), hybrid evolutionary strategies (R-ES), random search (RS) and ordinary solutions (OS) in the solution of the integrated problem. Problem is represented using chromosomes and first two genes are used for due date assignment and dispatching rules respectively. Remaining genes represents the selected route of jobs. Since the problem is NP-Hard, we used pure and hybrid metaheuristics in the solution. We also compared search results with ordinary solutions which are the initial results. We tried to prove the importance of search techniques and inferiority of initial random solutions. We also tried to observe superiority of directed genetic or evolutionary searches over undirected random search. Meanwhile, we tried to observe the power of hybrid searches which use random search initially and turn into directed search at the following iterations.

Let us give definitions of each function one by one; Society of Manufacturing Engineers has defined process planning as the systematic determination of the methods by which a product is to be manufactured economically and competitively. Zhang and Mallur defined production scheduling as a resource allocator, which considers timing information while allocating resources to the tasks [1]. Pinedo and Chao [2] defined scheduling as a proper allocation of resources that enables the company to optimize its objectives and achieve its goal. They also defined the job shop-scheduling environment as; n jobs to be processed on m machines to process these jobs. Each job processed in predetermined routes, visiting a number of machines. Job shop problems are seen in industries where orders have specified characteristics and order sizes are relatively small. According to Gordon et al. [3] "The scheduling problems involving due dates are of permanent interest. In a traditional production environment, a job is expected to be completed before its due date. In a just-in-time environment, a job is expected to be completed exactly at its due date."

There was a tremendous development in hardware, software and algorithm. With these developments, it became possible to solve problems which could not be solved earlier. After recent developments in computers, it is easier to prepare process plans using Computer Aided Process Planning (CAPP). As we mentioned earlier, these three functions effect each other and upstream decisions effect downstream functions and thus overall performance is affected. Poorly prepared process plans may cause unbalanced machine loading and reduce shop floor efficiency. Sometimes poorly prepared process plans are not followed on the shop floor. Since its easier to prepare process plans using CAPP, we may prepare alternative process plans and we can select among alternatives to balance workload at the shop floor. In case of contingencies such as machine breakdowns we can redirect jobs at the shop floor. This increases shop floor utilization and helps to balance it.

Customers are not equally important so we had better give close due dates to important customers and relatively far due dates to less important customers. The weighted due date assignment is not mentioned in IPPS and SWDDA problems in the literature. Besides assigning closer due dates for important customers we should also schedule important customers earlier as we did in this study. The problem should be solved in a reasonable amount of time, thus some powerful heuristics should be used. GA, R-GA, ES, R-ES and RS metaheuristics are utilized in this study.

After representing the problem as a chromosome we gave a higher probability of selection for dominant genes which are due date assignment and dispatching genes. Because these genes greatly affect the performance measure compared to any job route.

In the literature some works tried to minimize tardiness, some tried to minimize tardiness and earliness, some minimized maximum absolute lateness, and some minimized number of tardy jobs. Unlike these works, we tried to minimize the sum of weighted tardiness, earliness and due date related costs in this study.

Customers do not want long due dates, and far due dates can cause customer losses or price discounts and increase production costs. That is why we did not want to give far due dates unnecessarily, especially for the important customers. Conventionally tardiness is not desired. On the other hand in just-in-time (JIT) environment earliness is also undesired. Earliness means stock holding, spoilage, and some other earliness related costs. Tardiness means loss of customer goodwill, loss of customer permanently or may be a discount on the price. Thus we did not want any of these costs at our performance measure. Of course, we penalized these cost terms according to the weight of each job.

In this study, we did not want to give far due dates unnecessarily especially for important customers. We also wanted to give reasonable due dates so that we can keep our promises and reduce tardiness and earliness. We wanted jobs to be completed as close as possible to given reasonable due dates and tardiness is penalized more compared to earliness.

2. Literature survey

There are numerous works on IPPS and SWDDA or SWDWA problems in the literature. But there are only a few works on IPPSDDA problem. Demir and Taskin [4] worked on this problem for a Ph.D. thesis. Later Ceven and Demir [5] studied the benefit of integrating due date assignment with IPPS problem in a Master of Science thesis. Later Demir et al. [6] worked on the integration of process planning and due date assignment with ATC (Apparent Tardiness Cost) dispatching. Demir et al. [7] studied the integration of process planning and scheduling with SLK (Slack) due date assignment. In these studies unique due dates are determined for every customer.

Job shop scheduling with alternative process plans is integrated with due date determination. Concerning this research, we integrated WSPT dispatching with WSLK due date determination where alternative process plans are possible. As a distinct approach weighted SLK due date assignment method is used where weights of each customer are taken into account while determining unique due dates for each customer in this study. Important customers are given relatively closer due dates contrary to the relatively less important customers.

If we look at SWDDA problems we see that most of the works are done on common due date assignment. Parts of a product which are waiting to be assembled should be ready at the same time. But in this study, as mentioned above each customer has its own due dates.

Since job shop scheduling belongs to the NP-hard problem class, integrated problems are even harder to solve. For example, if we look at IPPS problems, exact solutions are only possible for very small problems. That is why genetic algorithms and its variants are more applicable for job shop scheduling problems or IPPS problems as they are utilized in this study. Zhu [8] and Wang and Li [9] used genetic algorithms and its variants in job shop scheduling.

"If we look at the literature we see that its hard to solve integrated problems. Some solutions are only possible for small problems. For IPPS in the literature, people use genetic algorithms, evolutionary algorithms or agent-based approach for integration, or they decompose problems because of the complexity of the problem. They decompose problems into loading and scheduling subproblems. They use mixed integer programming at the loading part and heuristics at the scheduling part" Demir et al. [7].

If we look at the early works on IPPS problem, we can see the following literature on this problem; Nasr and Elsayed [10], Hutchinson et al. [11], Chen and Khoshnevis [12], Zhang and Mallur [1], Brandimarte [13], Morad and Zalzala [14]. After these studies more works are done on IPPS such as: Ming and Mak [15], Tan and Khoshnevis [16], Kim et al. [17], Kumar and Rajotia [18], Lim and Zhang [19], Tan and Khoshnevis [20], Kumar and Rajotia [21], Moon et al. [22], Guo et al. [23], Leung et al. [24], Phanden et al. [25], Petrovic et al. [26], Zhang et al. [27], and Zhang and Wong [28].

Scheduling with due date assignment is also popular research topic which is extensively studied in the literature. Gordon et al. [3] presented a good literature survey and it will be useful to see this work before studying SWDDA problem. When we look at works on this problem most of them assigned unweighted due dates.

In this study, weighted due dates are assigned. Relatively important customers get closer due dates which greatly improve performance measure. At the IPPS and SWDDA problem some works penalize tardiness, some of them punish number of tardy jobs and some penalize both earliness and tardiness. In this study, we penalized all of the weighted due dates, weighted tardiness and weighted earliness cost. We give relatively close due dates for important customers and use WSLK as due date assignment and we schedule important customers first and used WSPT dispatching rule, so we save a lot from weighted due date related costs, weighted tardiness and earliness costs.

In the literature, we can observe works in two groups according to the number of machines. Single machine scheduling with due date assignment (SMSWDDA) and Multi-machine scheduling with due date assignment (MMSWDDA). Many works are on single machine problem and many of them on the multi-machine problem. At the former case, jobs are tried to be scheduled before a single machine and better due dates are tried to be found. At the latter case, jobs are tried to be scheduled on multiple machines and tried to be assigned better due dates. In our study, we have n jobs to be scheduled on m machines and each job will be given a unique due date according to processing time, the importance of the customer and given slack.

If we list some literature on SMSWDDA problem we can find following works; Panwalkar [29], Biskup and Jahnke [30], Cheng et al. [31], Cheng et al. [32], Lin et al. [33], Ying [34], Xia et al. [35], Gordon and Strusevich [36], and Li et al. [37].

When we list researchers on MMSWDDA problem we can give following works on this problem; Adamopolous and Pappis [38], Cheng and Kovalyov [39], Birman and Mosheiov[40] and Lauff and Werner [41]. Additionally following works can be given for SWDDA problem; Allaoua and Osmane [42], Yang et al. [43], Tuong and Soukhal [44], Li et al. [45], Li et al. [37], Vinod and Sridharan [46], Li et al. [47], Zhang and Wu [48], Yin et al. [49], Iranpoor et al. [50] Yin et al. [51], Shabtay [52], and Koulamas [53].

If we look at recent works we can see numerous works on scheduling with due window assignment (SWDWA) instead of SWDDA. In this studies, due windows are tried to be determined instead of a single point due date. The objective is to find better due windows with better starting point in time and length of windows. Jobs completed within due windows cause no cost and jobs completed out of windows cause tardiness or earliness costs. We can list works on this problem as; Mosheiov and Sarig [54], Cheng et al. [55], Zhao and Tang [56], Janiak et al. [57], Wang et al. [58], Ji et al. [59], Ji et al. [60] Yang et al. [61] and Liu et al. [62].

3. Problem studied

We studied IPPSDDA problem with different levels of integration. Alternative process plans, WSPT and service in random order (SIRO) dispatching rules and WSLK and Random (RDM) due date assignment rules are handled concurrently.

Four shop floors are studied. There are five alternative routes in relatively small shop floors. However, there are three alternative routes in relatively large shop floors. Because number of routes increases complexity and it takes more time to solve the problem. The number of alternative routes is limited to three in order to find a good solution in a reasonable amount of time.

Initially, SIRO dispatching rule and RDM due date assignment rule are used to represent unintegrated combination of the problem. RDM due date assignment is used to represent external due date assignment. WSLK is used to represent internal due date assignment. At the previous case, we try to optimize performance measure in case of given external due date. But at the second case, we assign due dates internally and we try to find best due dates which are the most suitable for us and optimize performance measure.

After unintegrated combination, we integrated WSPT scheduling with process plan selection. Here, due date assignment is still unintegrated and randomly determined. Later we tested the combination where due date assignment is integrated with process plan selection but here scheduling is unintegrated and jobs are scheduled according to SIRO. Finally, we integrated three functions and tested fully integrated combination.

As it is mentioned earlier, we have four shop floors. For example, at the smallest shop floor, we have 25 jobs and 5 machines. Each job has 5 alternative routes and there are 10 operations at each route. At the largest shop floor, there are 175 jobs and 35 machines. Each job has 3 alternative routes with 10 operations each. Processing times are determined as $\lfloor(12+z*6)\rfloor$ randomly. Here z is standard normal numbers and practically processing time of each operation changes in between 1 and 30 minutes.

Operations are assigned to machines randomly in each shop floor. Characteristics of each shop floor are as given in Table 1.

In this study, we assumed that shop floor works one shift that is 8 hours per day. So a shift makes 8*60=480 minutes. As a performance measure, we assumed sum of weighted tardiness, earliness and due dates. All terms are punished linearly according to weights of customers. We also assumed fixed cost if there is tardiness or earliness.

Table 1. Shop floors.

Shop floor	SF1 25x5x5	SF2 75x15x5	SF3 125x25x3	SF4 175x35x3
# of machines	5	15	25	35
# of Jobs	25	75	125	175
# of Routes	5	5	3	3
Processing Times	$\lfloor(12+z*6)\rfloor$			
# of op. per job	10			

Due dates are punished proportionally to the weight of the job and length of the due date. Due date length is multiplied with 8. Earliness is punished with 5 unit fixed cost and proportionally with multiplier 4 and multiplied with the weight of job. Tardiness is punished more compared to other terms. Because it is the most undesired component. Tardiness is punished with 10 unit fixed cost and proportionally with multiplier 12 and finally multiplied with the weight of job.

Punishment functions are given below where D is used for the due date of job j, E is used for earliness of job j and T is used for the tardiness of job j. PD is the penalty of due-date of job j, PE is the penalty of earliness and PT is the penalty of the tardiness of job j. Penalty (j) is the total penalty of job j and total penalty is the ultimate value that we want, which shows total punishment for all of the jobs;

$$\mathrm{PD(j)} = \text{weight (j)} * 8 * \left(\frac{\mathrm{D}}{480}\right) \tag{1}$$

$$\mathrm{PE(j)} = \text{weight (j)} * \left(5 + 4 * \left(\frac{\mathrm{E}}{480}\right)\right) \tag{2}$$

$$\mathrm{PT(j)} = \text{weight (j)} * \left(10 + 12 * \left(\frac{\mathrm{T}}{480}\right)\right) \tag{3}$$

$$\mathrm{Penalty(j)} = \mathrm{PD(j)} + \mathrm{PE(j)} + \mathrm{PT(j)} \tag{4}$$

$$\text{Total Penalty} = \sum_{\mathrm{j}} \mathrm{Penalty(j)} \tag{5}$$

4. Used techniques

We used five search techniques in this study. These techniques are GA, R-GA, ES, R-ES and RS. Results of searches are compared with OS and RS results to find out the power of directed and hybrid searches over undirected (random) search and over OS results. Solution techniques are explained as follows:

Genetic Algorithm (GA): One of the most powerful search technique in this study was genetic search. For each shop floor, we applied given number of genetic iterations. Two genetic operators which are crossover and mutation are used in genetic iterations. There are three populations used in this study. First is the main population, second is used for crossover population and the last is used as mutation population. We have 10 chromosomes in the main population and we produce 8 chromosomes using previous main population by using crossover operator for crossover population. Again by using previous main population we produce 5 more chromosomes by using mutation operator to get mutation population. According to the literature, the crossover is expected to be more powerful compared to the mutation operator. But as number of pairs of chromosomes selected for crossover and number of chromosomes selected for mutation increase, the time required to solve the problem also increases. On the one hand selecting chromosomes with better performance for crossover and mutation operators with higher probability improves the performance measure better and on the other hand, marginal improvement in selecting chromosomes with worse performance is poor compared the former case. So it is better to apply crossover with a higher rate and select better chromosomes with higher probability and not to select all of the chromosomes to reduce the time required to solve the problem. That's why we selected four pairs of chromosomes for crossover and selected five chromosomes for mutation. Using three populations which are previous main population, new crossover population and new mutation population we determine next main population and this makes one genetic iteration. Out of 23 chromosomes from three populations, we select best 10 chromosomes for next step main population. This search is also called directed search since updated best population is used to determine next step main population at every iteration.

Evolutionary Strategies (ES): In Germany, two students of Technical University of Berlin developed ES, while solving their optimization problem [63], [64]. Unlike GA, ES uses only mutation operator. To be fair in comparison in GA, R-GA, ES, R-ES and RS we use the same number of iterations and at every iteration, we produce 13 new solutions and we use same amount of new solutions.

Random Search (RS): We produce 13 brand new solutions in every iteration. To be fair amongst methods, we used the same number of iterations and the same amount of chromosomes at each iteration in all searches.

Hybrid Genetic Algorithms (R-GA): Initially random iterations are applied. Later iterations are turned into genetic iterations. It is very useful to apply random search in the beginning since it scans solution space faster and better. Afterwards, it becomes very poor to continue with random search and at this point, it is better to use a powerful directed search technique. For instance, if we produce a random number between 1 and 1000, the expected value becomes 500 where the marginal gain is 500. On the other hand, if we produce two random numbers and take their maximum then expected value of this maximum is 667 so marginal gain sharply reduced to 167. Furthermore, if we produce three random numbers and take the maximum of these three numbers then expected value becomes 750 and marginal gain further reduced 83. If we sort marginal gains in descending order; 500, 167, 83 are the marginal benefits of initial random iterations, respectively. So as its explained above initial random iterations are very useful but later it becomes drastically poor to apply.

Hybrid Evolutionary Strategies (R-ES): Initially random iterations are applied and later iterations are turned into evolutionary iterations in this search.

Iterations parameters for all search techniques are summarized in Table 4.

Ordinary Solution (OS): For every shop floor, we used only randomly produced chromosomes in the beginning as OS to represent how poor an ordinary solution can be.

CPU times required for pure and hybrid metaheuristics are given in Section 6.

If n is the number of jobs then we have n+2 genes at each chromosome. The first gene represents due date assignment methods and second gene is used for dispatching rules. Remaining genes are used to represent selected route of each job. In the relatively small shop floors, jobs can have 1 route out of 5 alternatives and at the relatively large shop floors, jobs have 3 alternative routes. A sample chromosome is given in Figure 1.

One thing we applied in this study is dominant genes. First and second genes affect results much more compared to other genes which are the route of jobs. For this reason, while applying GA, R-GA, ES, R-ES and RS first two genes are selected with a higher probability of crossover or mutation in GA and ES using dominant genes with higher probability improved efficiency of the solution techniques.

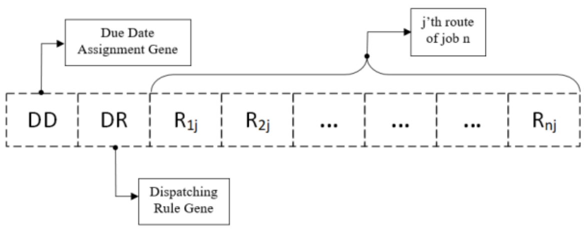


Figure 1. Sample chromosome.

Due dates are assigned using mainly two different type of rules. The first rule is weighted due date assignment rule WSLK which represents internal due date assignment and considers weights of each customer while assigning due dates. The second rule is RDM due date assignment rule that assigns due dates randomly which represents external due date assignment. With the constants used first gene takes one of four values. These rules are explained below in Table 2 and in Appendix A.

Table 2. Due-date assignment rules.

Method	Constant q_x	Rule no:
WSLK	$q_x = q_1, q_2, q_3$	1,2,3
RDM		4

Two different methods were used in order to dispatch. One is WSPT dispatching which schedules jobs according to the weights of the customers and processing times of the jobs. The other rule is SIRO which schedules jobs in random order. Dispatching rules are given and explained in Table 3 and in Appendix B.

Table 3. Dispatching rules.

Method	Rule no
WSPT	1
SIRO	2

Table 4. Iteration numbers for pure and hybrid searches.

	ES	R-ES Hybrid		RS	GA	R-GA Hybrid	
Shop Floor	ES Iter#	Random Iter#	ES Iter#	Random Iter#	GA Iter#	Random Iter#	GA Iter#
1	200	20	180	200	200	20	180
2	150	15	135	150	150	15	135
3	100	10	90	100	100	10	90
4	50	5	45	50	50	5	45

5. Solutions compared

SIRO-RDM (OS, RS, ES, R-ES, GA, R-GA): This is an unintegrated combination of the problem. Jobs are scheduled in random order and due dates are determined randomly. Random due date assignment represents exogenous due dates where we have no control over it.

WSPT-RDM (OS, RS, ES, R-ES, GA, R-GA): In this combination, a powerful dispatching rule WSPT is integrated with process planning, but due dates are still determined randomly. Here substantial improvement is obtained.

SIRO-WSLK (OS, RS, ES, R-ES, GA, R-GA): In this combination, this time WSLK due date assignment is integrated with process plan selection but this time dispatching is made in random order. Although this integration provides substantial improvement, unfortunately, SIRO dispatching sharply deteriorates the performance measure.

WSPT-WSLK (OS, RS, ES, R-ES, GA, R-GA): With this combination, we integrated process plan selection, scheduling and due-date assignment by using WSPT dispatching and WSLK due date assignment rules. This is the highest integration level.

All twenty-four solutions mentioned above are compared with one another. Different level of integrations are tested and searches are compared with each other and with ordinary solutions. Higher integration levels are found better and best results are obtained where three functions are integrated. Results are presented in Section 6 experiments and results section and conclusions are made in the final section.

6. Experiments and results

We coded the problem using C++ syntax. This program can perform genetic, evolutionary or random iterations while searching for better solutions, assign due dates and schedule jobs and evaluate performance measure for every given solution. At the first gene of the chromosomes, two rules which are WSLK and RDM rules are used and with the different constants used the first gene can take one of four values. At the second gene, two dispatching rules which are WSPT and SIRO rules are used. This gene can take one of two values. Remaining genes take values according to the selected route among given alternatives.

We ran the program using a laptop with 2.4 GHz processor with Intel i7 processor and 16 GB Ram with Borland C++ 5.02 compiler.

CPU times of the problems are given in Table 5. Since searches take time, CPU times for searches are given except ordinary solutions which require a negligible amount of time.

For different shop floors, given number of iterations are applied. At GA we applied genetic iterations which uses crossover and mutation operators. At the ES we applied only mutation operator. Since these are directed searches we use updated best solutions to generate new generations. On the other hand, RS is undirected search and we produce brand new random solutions at every iteration. We apply 200, 150, 100 and 50 iterations for each of the shop floors respectively.

If we look at Table 5 for shop floor 1, iterations took around 20 seconds approximately. For the second shop floor, it took approximately 200 seconds. At the third shop floor, it took in between 200 and 300 seconds. Finally, for the largest shop floor, 50 iterations took approximately 300 seconds.

First of all, we tested unintegrated solutions and we tested SIRO-RDM combination. Later, we integrated WSPT rule with process plan selection and we tested WSPT-RDM combinations. After that, we tested integration of WSLK rule with process planning. We tested SIRO-WSLK combinations. Finally, we integrated all three functions and this is the full integration and we tested WSPT-WSLK combinations. These solutions are explained in Section 5.

Four shop floors are tested for mentioned twenty-four types of solutions. For smallest shop floor we applied 200 genetic, evolutionary or random iterations. Iteration parameters can be found in Table 4. CPU times required to solve the shop floors for each search technique is given in Table 5. Smallest shop floor results are tabulated in Table 5 and illustrated in Figure 2. Results show that searches are found very useful and directed search outperformed undirected search. Furthermore higher integration is found better and fully integrated combinations are found the best. R-GA method gave the best result.

Similar results are obtained for the second shop floor. At this shop floor, we applied 150 genetic, evolutionary or random iterations and CPU times are given in Table 5. Results are listed in Table 5 and illustrated in Figure 3. Integrating functions are found to be useful. As integration level increases solutions become better and highest integration level with genetic search gave the best results. Directed searches outperformed undirected search.

Table 5. Comparison of twenty-four solutions for each shop floor.

Level of Integration (Combination)	Method	Shop Floor 1				Shop Floor 2				Shop Floor 3				Shop Floor 4			
		Best	Avg.	Worst	CPU	Best	Avg.	Worst	CPU	Best	Avg.	Worst	CPU	Best	Avg.	Worst	CPU
SIRO-RDM	OS	293	293	293	0	906	906	906	0	1413	1413	1413	0	2020	2020	2020	0
	ES	256	260	263	19	826	838	844	217	1315	1323	1329	277	1860	1871	1879	295
	R-ES	248	252	255	20	827	835	839	216	1322	1325	1327	277	1861	1875	1885	298
	GA	248	251	253	19	841	847	849	217	1309	1319	1323	274	1833	1872	1879	297
	R-GA	254	257	259	19	823	831	833	217	1303	1309	1314	274	1871	1882	1890	298
	RS	268	273	275	19	853	864	870	222	1355	1372	1378	286	1908	1925	1934	303
WSPT-RDM	OS	231	231	231	0	730	730	730	0	1153	1153	1153	0	1691	1691	1691	0
	ES	210	212	212	40	666	672	673	212	1084	1089	1092	274	1554	1560	1563	293
	R-ES	211	212	212	34	675	679	682	215	1086	1094	1096	275	1560	1567	1572	299
	GA	213	213	214	21	678	679	679	207	1101	1102	1103	274	1564	1564	1564	299
	R-GA	208	209	209	20	686	687	687	206	1102	1104	1106	271	1542	1543	1544	298
	RS	218	221	222	21	701	707	710	215	1124	1133	1140	284	1589	1600	1606	314
SIRO-WSLK	OS	322	322	322	0	982	982	982	0	1467	1467	1467	0	2104	2104	2104	0
	ES	256	271	275	23	853	858	862	230	1260	1267	1272	288	1770	1794	1807	305
	R-ES	258	265	268	24	851	860	866	229	1249	1269	1275	291	1785	1807	1821	303
	GA	260	264	267	23	845	850	854	226	1222	1252	1262	307	1770	1783	1792	302
	R-GA	255	266	269	22	846	853	857	225	1255	1264	1270	295	1782	1789	1795	305
	RS	267	279	283	22	868	884	890	230	1284	1294	1304	311	1790	1812	1825	307
WSPT-WSLK	OS	247	247	247	0	766	766	766	0	1120	1120	1120	0	1621	1621	1621	0
	ES	189	191	192	22	604	617	620	219	931	933	935	233	1363	1368	1373	306
	R-ES	189	192	193	22	621	626	630	231	937	942	945	228	1360	1371	1377	310
	GA	192	193	194	22	601	603	605	213	926	928	930	228	1351	1356	1359	308
	R-GA	187	191	193	22	619	621	622	219	931	932	933	235	1347	1350	1353	308
	RS	199	202	204	22	629	641	649	224	954	966	972	313	1379	1388	1392	309

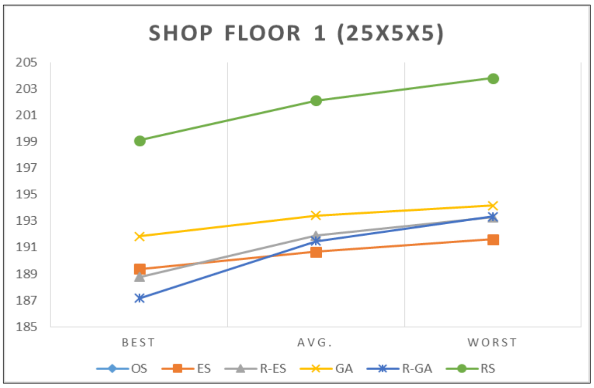


Figure 2. First shop floor results.

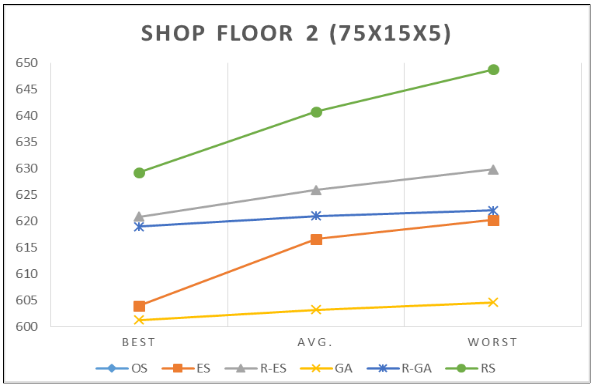


Figure 3. Second shop floor results.

At the third shop floor, GA gave the best result. At this shop floor, we applied 100 iterations and results are summarized in Table 5 and in Figure 4. Directed searches outperformed undirected search.

Last shop floor was the biggest shop floor. We applied

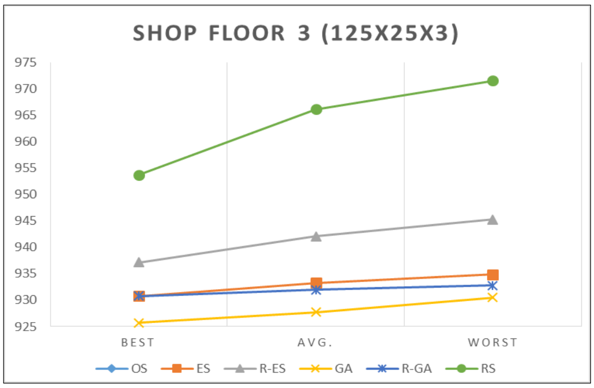


Figure 4. Third shop floor results.

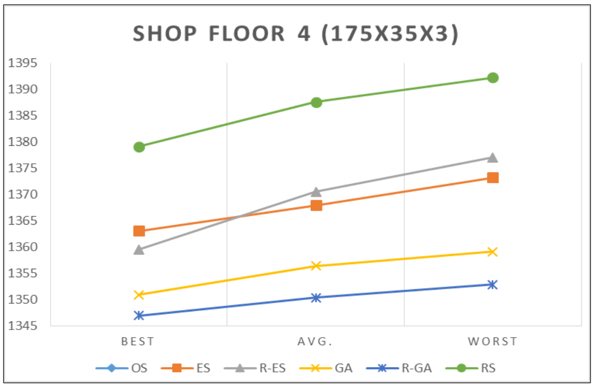


Figure 5. Fourth shop floor results.

50 iterations to find a good solution in a reasonable amount of time. According to the results listed in Table 5 and illustrated in Figure 5, highest integration level with R-GA is found best. According to the results, higher integration level gave better solutions. R-GA

outperformed all other searches.

If we consider all level of integrations and compare 24 combinations for each shop floor then there are 16 best results. 8 of these best results are obtained through GA and 7 of them is obtained through R-GA search and once we obtained the best result by using ES search. These results can be seen in Table 5.

7. Discussions and conclusion

In this study, we tested different integration level of process planning, weighted scheduling and weighted due date assignment. We used WSPT rule for weighted scheduling as it is a powerful dispatching rule. We applied WSLK rule as due date assignment rule. We considered weights of the jobs because they provide substantial improvements in the performance measure which is the sum of weighted tardiness, earliness and due date related costs. We tested different search techniques which are GA, R-GA, ES, R-ES and RS. We compared search techniques with each other and with OS results for different level of integrations.

In the beginning of the study, we tested unintegrated solutions and we solved the problem according to SIRO-RDM combinations. Later we integrated WSPT scheduling with process plan selection but due dates are still determined randomly. We solved the problem for WSPT-RDM combinations at this level. After that, we integrated WSLK due date assignment with process plan selection. Scheduling is performed in random order and we used SIRO dispatching rule. We tested here SIRO-WSLK combinations. Although WSLK due-date assignment is very useful, SIRO dispatching severely deteriorates the performance measure, that is why this is not as good as other integration levels. At the end, we integrated three functions which are process planning, scheduling and due date assignment. We used WSLK due date assignment rule at the first gene of the chromosomes and used WSPT dispatching rule at the second gene of the chromosomes and tried to find better routes at the remaining genes of the chromosomes. We used dominant genes at the chromosomes because first two genes are much more important compared to the routes of each job. We tested here WSPT-WSLK combinations. Full integration with GA and R-GA techniques are found the best in these four shop floors. Again full integration is always found better compared to the intermediate levels. GA and R-GA searches outperformed other searches and directed searches are always performed better compared to the undirected search. While GA search gave the best results for eight times, compared to R-GA search which gave the best solutions for seven times. ES search gave the best solution only once among all other search techniques.

As a conclusion, we can see that integration level is very important and highest integration level gives the best results. According to the results, we can also say that weighted scheduling and weighted due date assignment also improves global performance which is the sum of weighted tardiness, earliness and due date costs.

Traditionally three functions that we integrated are performed separately which leads to poor global performance and greatly affects the performance measure. In this competitive environment, we should utilize every way that makes us more competitive, reduces our costs and increases our profits. The performance measure is greatly reduced by higher integration level, with weighted scheduling, weighted due date assignment and with a better search technique.

If these three functions are performed sequentially, they give poor inputs for other functions. For example, independently prepared process plans can be poor input for scheduling and can cause unbalanced machine loads and can reduce shop floor performance. Furthermore, these plans may not be followed at the shop floor level at all as they are not realistic. Independently prepared scheduling without considering due dates may cause much more cost at performance measure. Independently given due dates can cause worse performance measure and poorly given due dates makes it harder to keep our promises.

To sum up, this study has shown that higher integration gives better performance measure and we should use highest integration level. WSPT is a strong dispatching rule that takes into account of weights of each customer, and WSLK is a strong due date assignment rule that considers the importance of each customer. Alternative process plans help us to improve scheduling and due date assignment performance so we get much better global performance. Thus it is very useful to implement WSLK rule while assigning due dates and we should give closer due dates for more important customers and relatively far due dates for relatively less important customers. We should also schedule jobs that have both shorter processing times and belong to important customers earlier than the other jobs.

In terms of solution techniques, directed searches are always better than undirected search. GA and R-GA techniques were found best and hybrid techniques are found promising.

Appendix A: Due-date assignment rules

WSLK (Weighted Slack) → Due = TPT + q_x * k (According to weights) q_x = q1,q2 or q3

q1=0.5*Pav, q2=Pav, q3=1.5*Pav

RDM (Random due-date assignment) → Due = N ~ $(3*P_{av},(P_{avg})^2)$

TPT: *Total processing time*

P_{avg}: *Mean processing time of all job waiting*

Appendix B: Dispatching rules

WSPT: *Weighted shortest processing time first*

SIRO (Service in Random order): *A job among waiting jobs is selected randomly to be processed.*

References

[1] Zhang, H.C., & Mallur, S. (1994). An integrated model of process planning and production scheduling. *International Journal of Computer Integrated Manufacturing*, 7, 356–364.

[2] Pinedo, M. and Chao, X. (1999). *Operations scheduling with applications in manufacturing and services*. McGraw-Hill Companies.

[3] Gordon, V., Proth, J.M., & Chu, C. (2002). A survey of the state-of-the-art of common due date assignment and scheduling research. *European Journal of Operational Research*, 139, 1–25.

[4] Demir, H.I. and Taskin, H. (2005). *Integrated Process Planning, Scheduling and Due-Date Assignment*. PhD Thesis.

[5] Ceven, E. and Demir, H.I. (2007). *Benefits of Integrating Due-Date Assignment with Process Planning and Scheduling*. Master of Science Thesis.

[6] Demir, H.I., Cakar, T., Ipek, M., Uygun, O., & Sari, M. (2015). Process Planning and Due-date Assignment with ATC Dispatching where Earliness, Tardiness and Due-dates are Punished. *Journal of Industrial and Intelligent Information*, 3, 197–204.

[7] Demir, H.I., Uygun, O., Cil, I., Ipek, M., & Sari, M. (2015). Process Planning and Scheduling with SLK Due-Date Assignment where Earliness, Tardiness and Due-Dates are Punished. *Journal of Industrial and Intelligent Information*, 3, 173–180.

[8] Zhu, C. (2012). Applying Genetic Local Search Algorithm to Solve the Job-Shop Scheduling Problem. *International Journal of Industrial Engineering: Theory, Applications and Practice*, 19.

[9] Wang, W., & Li, T. (2011). Improved cultural algorithms for job shop scheduling problem. *International Journal of Industrial Engineering: Theory, Applications and Practice*, 18.

[10] Nasr, N., & Elsayed, E. (1990). A. Job shop scheduling with alternative machines. *International Journal of Production Research*, 28, 1595–1609.

[11] Hutchison, J., Leong, K., Snyder, D., & Ward, P. (1991). Scheduling approaches for random job shop flexible manufacturing systems. *International Journal of Production Research*, 29, 1053–1067.

[12] Chen, Q.M., & Khoshnevis, B. (1993). Scheduling with flexible process plans. *Production Planning & Control*, 4, 333–343.

[13] Brandimarte, P. (1999). Exploiting process plan flexibility in production scheduling: A multi-objective approach. *European Journal of Operational Research*, 114, 59–71.

[14] Morad, N., & Zalzala, A. (1999). Genetic algorithms in integrated process planning and scheduling. *Journal of Intelligent Manufacturing*, 10, 169–179.

[15] Ming, X.G., & Mak, K.L. (2000). A hybrid Hopfield network-genetic algorithm approach to optimal process plan selection. *International Journal of Production Research*, 38, 1823–1839.

[16] Tan, W., & Khoshnevis, B. (2000). Integration of process planning and scheduling - a review. *Journal of Intelligent Manufacturing*, 11, 51–63.

[17] Kim, Y.K., Park, K., & Ko, J. (2003). A symbiotic evolutionary algorithm for the integration of process planning and job shop scheduling. *Computers & Operations Research*, 30, 1151–1171.

[18] Kumar, M., & Rajotia, S. (2003). Integration of scheduling with computer aided process planning. *Journal of Materials Processing Technology*, 138, 297–300.

[19] Lim, M.K., & Zhang, D.Z. (2004). An integrated agent-based approach for responsive control of manufacturing resources. *Computers & Industrial Engineering*, 46, 221–232.

[20] Tan, W., & Khoshnevis, B. (2004). A linearized polynomial mixed integer programming model for the integration of process planning and scheduling. *Journal of Intelligent Manufacturing*, 15, 593–605.

[21] Kumar, M., & Rajotia, S. (2005). Integration of process planning and scheduling in a job shop environment. *The International Journal of Advanced Manufacturing Technology*, 28, 109–116.

[22] Moon, C., Lee, Y.H., Jeong, C.S., & Yun, Y. (2008). Integrated process planning and scheduling in a supply chain. *Computers & Industrial Engineering*, 54, 1048–1061.

[23] Guo, Y.W., Li, W.D., Mileham, A.R., & Owen, G. W. (2009). Applications of particle swarm optimization in integrated process planning and scheduling. *Robotics and Computer-Integrated Manufacturing*, 25, 280–288.

[24] Leung, C.W., Wong, T.N., Mak, K.L., & Fung, R.Y.K. (2010). Integrated process planning and scheduling by an agent-based ant colony optimization. *Computers & Industrial Engineering*, 59, 166–180.

[25] Phanden, R.K., Jain, A., & Verma, R. (2011). Integration of process planning and scheduling: a state of the art review. *International Journal of Computer Integrated Manufacturing*, 24, 517–534.

[26] Petrović, M., Vuković, N., Mitić, M., & Miljković, Z. (2016). Integration of process planning and scheduling using chaotic particle swarm optimization algorithm. *Expert Systems with Applications*, 64, 569–588.

[27] Zhang, Z., Tang, R., Peng, T., Tao, L., & Jia, S. (2016). A method for minimizing the energy consumption of machining system: integration of process planning and scheduling. *Journal of Cleaner Production*, 137, 1647–1662.

[28] Zhang, L., & Wong, T.N. (2016). Solving integrated process planning and scheduling problem with

constructive meta-heuristics. *Information Sciences*, 340–341, 1–16.

[29] Panwalkar, S.S., Smith, M.L., & Seidmann, A. (1982). Common Due Date Assignment to Minimize Total Penalty for the One Machine Scheduling Problem. *Operations Research*, 30, 391–399.

[30] Biskup, D., & Jahnke, H. (2001). Common due date assignment for scheduling on a single machine with jointly reducible processing times. *International Journal of Production Economics*, 69, 317–322.

[31] Cheng, T.C.E., Chen, Z.L., & Shakhlevich, N.V. (2002). Common due date assignment and scheduling with ready times. *Computers & Operations Research*, 29, 1957–1967.

[32] Cheng, T.C.E., Kang, L.Y., & Ng, C.T. (2005). Single machine due-date scheduling of jobs with decreasing start-time dependent processing times. *International Transactions in Operational Research*, 12, 355–366.

[33] Lin, S.W., Chou, S.Y., & Chen, S.C. (2006). Meta-heuristic approaches for minimizing total earliness and tardiness penalties of single-machine scheduling with a common due date. *Journal of Heuristics*, 13, 151–165.

[34] Ying, K.C. (2008). Minimizing earliness–tardiness penalties for common due date single-machine scheduling problems by a recovering beam search algorithm. *Computers & Industrial Engineering*, 55, 494–502.

[35] Xia, Y., Chen, B., & Yue, J. (2008). Job sequencing and due date assignment in a single machine shop with uncertain processing times. *European Journal of Operational Research*, 184, 63–75.

[36] Gordon, V.S., & Strusevich, V.A. (2009). Single machine scheduling and due date assignment with positionally dependent processing times. *European Journal of Operational Research*, 198, 57–62.

[37] Li, J., Yuan, X., Lee, E.S., & Xu, D. (2011a). Setting due dates to minimize the total weighted possibilistic mean value of the weighted earliness–tardiness costs on a single machine. *Computers & Mathematics with Applications*, 62, 4126–4139.

[38] Adamopoulos, G.I., & Pappis, C.P. (1998). Scheduling under a common due-data on parallel unrelated machines. *European Journal of Operational Research*, 105, 494–501.

[39] Cheng, T.C.E., & Kovalyov, M.Y. (1999). Complexity of parallel machine scheduling with processing-plus-wait due dates to minimize maximum absolute lateness. *European Journal of Operational Research*, 114, 403–410.

[40] Birman, M., & Mosheiov, G. (2004). A note on a due-date assignment on a two-machine flow-shop. *Computers & Operations Research*, 31, 473–480.

[41] Lauff, V., & Werner, F. (2004). Scheduling with common due date, earliness and tardiness penalties for multimachine problems: A survey. *Mathematical and Computer Modelling*, 40, 637–655.

[42] Allaoua, H., & Osmane, I. (2010). Variable Parameters Lengths Genetic Algorithm for Minimizing Earliness-Tardiness Penalties of Single Machine Scheduling With a Common Due Date. *Electronic Notes in Discrete Mathematics*, 36, 471–478.

[43] Yang, S.J., Yang, D.L., & Cheng, T.C.E. (2010). Single-machine due-window assignment and scheduling with job-dependent aging effects and deteriorating maintenance. *Computers & Operations Research*, 37, 1510–1514.

[44] Huynh Tuong, N., & Soukhal, A. (2010). Due dates assignment and JIT scheduling with equal-size jobs. *European Journal of Operational Research*, 205, 280–289.

[45] Li, X., Gao, L., Zhang, C., & Shao, X. (2010). A review on Integrated Process Planning and Scheduling. *International Journal of Manufacturing Research*, 5, 161–180.

[46] Vinod, V., & Sridharan, R. (2011). Simulation modeling and analysis of due-date assignment methods and scheduling decision rules in a dynamic job shop production system. *International Journal of Production Economics*, 129, 127–146.

[47] Li, S., Ng, C.T., & Yuan, J. (2011b). Group scheduling and due date assignment on a single machine. *International Journal of Production Economics*, 130, 230–235.

[48] Zhang, R., & Wu, C. (2012). A hybrid local search algorithm for scheduling real-world job shops with batch-wise pending due dates. *Engineering Applications of Artificial Intelligence*, 25, 209–221.

[49] Yin, Y., Cheng, S.R., Cheng, T.C.E., Wu, C.C., & Wu, W.H. (2012). Two-agent single-machine scheduling with assignable due dates. *Applied Mathematics and Computation*, 219, 1674–1685.

[50] Iranpoor, M., Fatemi Ghomi, S.M.T., & Zandieh, M. (2013). Due-date assignment and machine scheduling in a low machine-rate situation with stochastic processing times. *Computers & Operations Research*, 40, 1100–1108.

[51] Yin, Y., Cheng, T.C.E., Cheng, S.R., & Wu, C.C. (2013). Single-machine batch delivery scheduling with an assignable common due date and controllable processing times. *Computers & Industrial Engineering*, 65, 652–662.

[52] Shabtay, D. (2016). Optimal restricted due date assignment in scheduling. *European Journal of Operational Research*, 252, 79–89.

[53] Koulamas, C. (2017). Common due date assignment with generalized earliness / tardiness penalties. *Computers & Industrial Engineering*, 109, 79–83.

[54] Mosheiov, G., & Sarig, A. (2010). Scheduling with a common due-window: Polynomially solvable cases. *Information Sciences*, 180, 1492–1505.

[55] Cheng, T.C.E., Yang, S.J., & Yang, D.L. (2012). Common due-window assignment and scheduling of linear time-dependent deteriorating jobs and a deteriorating maintenance activity. *International Journal of Production Economics*, 135, 154–161.

[56] Zhao, C., & Tang, H. (2012). A note to due-window assignment and single machine scheduling with deteriorating jobs and a rate-modifying activity. *Computers & Operations Research*, 39, 1300–1303.

[57] Janiak, A., Janiak, W., Kovalyov, M.Y., Kozan, E., & Pesch, E. (2013). Parallel machine scheduling and common due window assignment with job independent earliness and tardiness costs. *Information Sciences*, 224, 109–117.

[58] Wang, J.B., Liu, L., & Wang, C. (2013). Single machine SLK/DIF due window assignment problem with learning effect and deteriorating jobs. *Applied Mathematical Modelling*, 37, 8394–8400.

[59] Ji, M., Ge, J., Chen, K., & Cheng, T.C.E. (2013). Single-machine due-window assignment and scheduling with resource allocation, aging effect, and a deteriorating rate-modifying activity. *Computers & Industrial Engineering*, 66, 952–961.

[60] Ji, M., Chen, K., Ge, J., & Cheng, T.C.E. (2014). Group scheduling and job-dependent due window assignment based on a common flow allowance. *Computers & Industrial Engineering*, 68, 35–41.

[61] Yang, D.L., Lai, C.J., & Yang, S.J. (2014). Scheduling problems with multiple due windows assignment and controllable processing times on a single machine. *International Journal of Production Economics*, 150, 96–103.

[62] Liu, L., Wang, J.J. , Liu, F., & Liu, M. (2017). Single machine due window assignment and resource allocation scheduling problems with learning and general positional effects. *Journal of Manufacturing Systems*, 43, Part 1, 1–14.

[63] Rechenberg, I. (1965). *Cybernetic Solution Path of an Experimental Problem. Ministry of Aviation*, Royal Aircraft Establishment Library Translation No: 1122.

[64] Schwefel, H.P. (1981). *Numerical Optimization of Computer Models*. John Wiley & Sons, Inc.: New York, NY.

Halil Ibrahim Demir *was born in Sivas, Turkey in 1971. In 1988 he received a full scholarship and entered Bilkent University, Ankara, Turkey, to study in the Industrial Engineering Department. He got his Bachelor of Science degree in Industrial Engineering in 1993. In 1994, he received a full scholarship for graduate study in the USA from the Ministry of Education of Turkey. In 1997 he received a Master of Science degree in Industrial Engineering from Lehigh University, Bethlehem, Pennsylvania, USA. He was then accepted to Northeastern University, Boston, Massachusetts for Ph.D. study. He finished his Ph.D. courses at Northeastern and completed a Ph.D. thesis at Sakarya University, Turkey in 2005 for a Ph.D. in Industrial Engineering. He obtained an academic position at Sakarya University as an Assistant Professor. His Research Areas are Production Planning, Scheduling, Application of OR, Simulation, Artificial intelligence techniques, Genetic algorithms, Artificial Neural Networks, Fuzzy Logic and Decision making.*

Onur Canpolat *was born in Elazığ, Turkey in 1990. He is currently a Research Assistant in the Department of Industrial Engineering at Sakarya University, Sakarya, Turkey. He received BSc. and MSc. degree in Industrial Engineering from Sakarya University, Sakarya, Turkey in 2012 and 2016, respectively. He is currently a PhD student at the same university. His areas of interest include multi-criteria decision making, operations research and optimization.*

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957eISSN: 2146-5703
Vol.8, No.1, pp.84-91 (2018)
https://doi.org/10.11121/ijocta.01.2018.00468

RESEARCH ARTICLE

Design of an optimal state derivative feedback LQR controller and its application to an offshore steel jacket platform

Hakan Yazici [a] * and Mert Sever [b]

[a]*Department of Mechanical Engineering, Yildiz TechnicalUniversity, Turkey*
[b]*Department of Mechatronics Engineering, Yildiz Technical University, Turkey*
hyazici@yildiz.edu.tr, msever@yildiz.edu.tr

ARTICLE INFO

Article history:
Received: 27 April 2017
Accepted: 24 October 2017
Available Online: 8 November 2017

Keywords:
Optimal control
State derivative feedback
Convex optimization
Offshore steel jacket platform

AMS Classification 2010:
49J15, 90C22

ABSTRACT

This paper concerns with the optimal state derivative feedback LQR controller design for vibration control of an offshore steel jacket platform having active tuned mass damper against the wave induced disturbances. Considering that the state derivative signals such as acceleration and velocity are easier to measure rather than the state variables such as displacement, state derivative feedback control strategy is proposed to obtain practically applicable and easily realizable synthesis method. On the basis of convex optimization approach, state derivative feedback LQR controller design is formulated in Linear Matrix Inequalities (LMIs) form to get an optimal feasible solution set. Finally, an offshore steel jacket platform subject to nonlinear self excited wave force is used to illustrate the effectiveness of the proposed approach through simulations. The results show that proposed state derivative LQR controller is very effective in reducing vibration amplitudes of each floor of modeled offshore steel jacket platform and achieves compitable control performance to classical LQR controller design.

1. Introduction

The offshore steel jacket platforms play an important role in the oil exploration and drilling operation in the oceans [1]. As it is well known that offshore steel jacket paltforms are exposed to the various disturbances such as strong winds, ocean waves, earthquakes which are caused the structural vibrations and makes them very vulnerable and unsafe [2], [3]. Therefore, vibration control of offshore steel jacket platforms subject to environmental disturbances and working conditions has been receving a great deal of interest for the last two decades, and a lot of research effort has been devoted to the development of advanced control algorithms.

Abdel-Rohman has modeled a realistic offshore steal jacket platform using finite elements method to control the structural vibrations against wave loads [1]. Multi-loop feedback controller design has been developed in Terro [2] et al., for nonlinear wave excited steel jacket platforms. Wu et al. have dealt with the non-fragile state feedback H_∞ control problem to attenuate vibrations of steel-jacket platform subject to regular wave disturbance [4]. Mei et al. have proposed the design of a fuzzy H_∞ controller for active vibration control of an offshore platform with parameter uncertainties [5]. Sliding mode H_∞ vibration control problem has been considered in Zhang et al [6], for offshore steal jacket platform having nonlinear self excited wave forces and external disturbances. Li et al., have applied state feedback H_2 controller in reducing the effect of wave laoding on offshore platform [7]. State feedback stabilization control problem for offshore steel jacket platforms having actuator delay has been considered in Zhang et al [8]. They have assumed that all the state variables of the offshore platform are available for measurement. Zhang and Han have applied network based modelling and active vibration control for offshore steel jacket platform having tuned mass damper.

As it can be observed from the summarised literature, papers that address the state feedback vibration control problem of offshore steel jacket platforms are quite a few. As it is well known that state feedback controller assumes that all the state variables are available for measurement which are displacement and velocities in active structural control problem. However, displacement signals are not possible to be obtained accurately by integration, since the accelerometers are noisy and contain dc offset in low frequency region.

*Corresponding author

Therefore, a pure integrator is not practical and should be combined with a high-pass-filter (HPF) to remove the integrator drift. An HPF with sufficiently large time constant, results in a phase error. It is apparently seen that considerable amount of effort is required to accurately integrate acceleration signals. In addition, computation accuracy will gradually decrease by getting displacement from acceleration with double step integration process [10]. In the light of aforementioned considerations, state derivative feedback controller design strategy is a promising active structural controller approach in the following aspects. Firstly, only a single step integration is needed to obtain velocity for accelerometer outputs. Note that accelerometers are one of the most common sensors in active structural control problems. Secondly, closed loop system order has not been increased, since the state derivative feedback controllers are static and memoryless with no additional state variables. This situation motivates us with the fact that there exist still more room obtaining practically applicable optimal state derivative feedback LQR controller synthesis to attenuate the vibrations of offshore platforms. Moreover, to the best of authors' knowledge it can be seen that there is no result has been given in the literature on the active vibration control of offshore steel jacket platform by the use of state derivative feedback approach, so far.

In this study, because of the state derivative signals are easier to measure, an LMI based optimal state derivative LQR controller is developed to control of the offshore steel jacket platform having nonlinear wave disturbances. In controller design, first, stability and solvability conditions of an optimal state derivative feedback LQR controller is presented in LMI form and minimization of quadratic cost function is ensured by the use of convex optimization techniques. Then, in order to compare the proposed method, well known classical LQR controller is designed. Last, numerical simulations studies have been conducted to illustrate the effectiveness of the propsed control staregy. The main importance of this study is to develop an easily realizable synthesis method to obtain practically applicable optimal state derivative LQR controller achieves comparable performance improvement with the conventional state feedback LQR controller.

Rest of papers organized as follows. Mathematical model of the realistic offshore steel jacket platform and formulations of nonlinear Morison Equations based wave force are given in Section 2. The design of proposed state derivative feedback LQR controller are presented in Section 3. Simulation resulst with discussions are given in Section 4. Finally, Section 5 concludes the paper.

Notation: The notation to be used in the paper is fairly standard. $\Re$ stands for the set of real numbers, $\Re^{n\times n}$ is the set of $n\times n$ dimensional real matrices. 'diag' denotes the diagonal matrices. The identity and null matrices are denoted by I and 0, respectively. $X>0\,(\geq,\leq 0)$ denotes that X is a positive definite (positive semi-definite, negative definite) matrix. The notation '*' denotes off-diagonal block completion of a symmetric matrix. Finally, $diag\{M_1,\ldots,M_n\}$ stands for a diagonal matrix with elements $M_1,\ldots,M_n$ appearing on its diagonal.

2. Mathematical modeling of offshore steel jacket platform

In this section, a realistic offshore steel jacket platform model that includes an Active Tuned Mass Damper (ATMD) is used for controller design as shown in Figure 1 [2]. In this model, ATMD is used as an active control mechanism to supress structural vibrations, which is installed on the top floor of the offshore steel jacket platform.

The equations of the motion of the considered offshore steel jacket platform have been formulated in [2], [8], by the use of first two dominant vibration modes of the system as,

$$\begin{aligned} z_1(t) = &-2\xi_1\omega_1 z_1(t) \\ &-\phi_1 C_T\left[\phi_1 z_1(t)+\phi_2 z_2(t)\right] \\ &+\phi_1 C_T z_T(t)-\omega_1^2 z_1(t) \\ &-\phi_1 K_T\left[\phi_1 z_1(t)+\phi_2 z_2(t)\right] \\ &+\phi_1 K_T z_T(t)-\phi_1 u(t) \\ &+f_1\big(z_1(t),z_2(t),t)\big) \\ &+f_2(z_1(t),z_2(t),t) \end{aligned} \tag{1}$$

$$\begin{aligned} z_2(t) = &-2\xi_2\omega_2 z_2(t) \\ &-\phi_2 C_T\left[\phi_1 z_1(t)+\phi_2 z_2(t)\right] \\ &+\phi_2 C_T z_T(t)-\omega_2^2 z_2(t) \\ &-\phi_2 K_T\left[\phi_1 z_1(t)+\phi_2 z_2(t)\right] \\ &+\phi_2 K_T z_T(t)-\phi_2 u(t) \\ &+f_3\big(z_1(t),z_2(t),t\big) \\ &+f_4\big(z_1(t),z_2(t),t\big) \end{aligned} \tag{2}$$

$$\begin{aligned} z_T(t) = &-2\xi_T\omega_T z_T(t) \\ &+2\xi_T\omega_T\left[\phi_1 z_1(t)+\phi_2 z_2(t)\right] \\ &+\frac{1}{m_T}u(t) \\ &+\omega_T^2\left[\phi_1 z_1(t)+\phi_2 z_2(t)\right] \\ &-\omega_T^2 z_T(t). \end{aligned} \tag{3}$$

Here, $z_1(t)$ and $z_2(t)$ are the generalized coordinates, which represent the first and second vibration modes of modeled offshore steel jacket platform, respectively. ω_1 and ω_2 are the natural frequencies of the first and second modes of vibration, repectively. ξ_1 and ξ_2 are damping ratios of the first and second modes of vibration, respectively; ϕ_1 and ϕ_2 are the first and second mode shapes, respectively. $z_T(t)$ represents the horizontal displacement of ATMD, ξ_T and ω_T are the damping ratio and natural frequency of ATMD, respectively. m_T,

C_T and K_T are mass, damping and stiffness of ATMD, respectively. f_1, f_2, f_3 and f_4 are the nonlinear wave force and $u(t)$ represents the active control force.

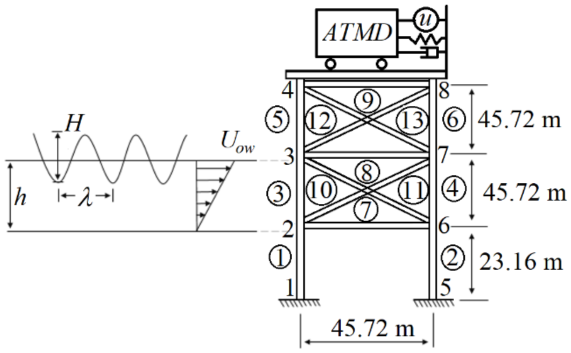


Figure 1. A simplified offshore steel jacket platform having ATMD [2].

The state variables and exogenous input can be defined as follow:

$$x(t) := [x_1(t) \quad x_2(t) \quad x_3(t) \quad x_4(t) \quad x_5(t) \quad x_6(t)]^T, \tag{4}$$

$$w(x,t) := \begin{bmatrix} f_1(x_1,x_3,t) + f_2(x_1,x_3,t) \\ f_3(x_1,x_3,t) + f_4(x_1,x_3,t) \end{bmatrix} \tag{5}$$

where,

$$x_1(t) := z_1(t), \quad x_2(t) := \dot{z}_1(t),$$
$$x_3(t) := z_2(t), \quad x_4(t) := \dot{z}_2(t),$$
$$x_5(t) := z_T(t), \quad x_6(t) := \dot{z}_T(t).$$

By the use of (4) and (5), the equaitionsc of motion (1), (2) and (3) can be rewritten into state space form as,

$$\dot{x}(t) = Ax(t) + B_u u(t) + B_w w(x,t) \tag{6}$$

where $A \in \Re^{n\times n}$ is the state matrix which is given by,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & 0 & 0 & 0 & 1 \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix} \tag{7}$$

with,

$$a_{21} = -\omega_1^2 - K_T\phi_1^2, a_{22} = -2\xi_1\omega_1 - C_T\phi_1^2,$$
$$a_{23} = -K_T\phi_1\phi_2, a_{24} = -C_T\phi_1\phi_2,$$
$$a_{25} = \phi_1 K_T, a_{26} = \phi_1 C_T \ ,$$
$$a_{41} = -K_T\phi_1\phi_2, a_{42} = -C_T\phi_1\phi_2,$$
$$a_{43} = -\omega_2^2 - K_T\phi_1^2, a_{44} = -2\xi_2\omega_2 - C_T\phi_2^2,$$
$$a_{45} = \phi_2 K_T, a_{46} = \phi_2 C_T \ ,$$
$$a_{61} = \omega_T^2\phi_1, a_{62} = 2\xi_T\omega_T\phi_1,$$
$$a_{63} = \omega_T^2\phi_1, a_{64} = 2\xi_T\omega_T\phi_2,$$
$$a_{65} = -\omega_T^2, a_{66} = -2\xi_T\omega_T.$$

$B_u \in \Re^{n\times m}$ is the control input matrix which is given by

$$B_u = \begin{bmatrix} 0 & -\phi_1 & 0 & -\phi_2 & 0 & \dfrac{1}{m_T} \end{bmatrix}^T \tag{8}$$

and $B_w \in \Re^{n\times p}$ is the disturbance input matrix which is given by

$$B_w = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T. \tag{9}$$

2.1. Disturbance model

As it is well known that offshore platforms are exposed to the nonlinear self-excited wave forces. In this study, a nonlinear Morison equation is used to calculate horizontal wave force [2]. Let us consider a joint point p on the platform. Unidirectional plane wave forces exerted on this point can be obtained as

$$F_p = \frac{1}{2}\rho C_D A_p |U'_{px}| U'_{px} + \rho C_I B_p a_{px} - \rho |C_I - 1| B_p \ddot{U}_{px}. \tag{10}$$

Here, F_p is the wave force vector, A_p is the lumped area at point p, U_{px} is horizontal velocity of water, U'_{px} is horizontal velocity of point p, the difference $U'_{px}=U_{px}-\dot{U}_{px}$ is the relative velocity of water with respect to point p. C_D and C_I are drag and inertia coefficients, relatively, a_{px} is horizontal acceleration of wave on point p, B_p is lumped volume on p, ρ water density, and $\ddot{U}_{px}$ is horizantal acceleration of point p. Velocity and acceleration of a horizontal advancing wave are related to wave characteristics and properties of motion area of the wave.

At point p, the horizontal velocity is given as,

$$U'_{px} = E_p \cos(kx_p - \Omega t) + U_{ow}\frac{Y_p}{h} - \dot{U}_{px}. \tag{11}$$

The horizontal acceleration is expressed as,

$$a_{px} = E_p \Omega \sin(kx_p - \Omega t) \tag{12}$$

where,

$$E_p = \frac{\Omega H \cosh(kY_p)}{2\sinh(kh)} \tag{13}$$

x_p and Y_p are the location of point p with respect to a fixed coordinate axes reference, hence Y_p is the height of point p from seabed, h is water depth, Ω is the frequency of the water, H is the height of water, λ is the wavelength, $k=2\pi/\lambda$ is wave number and U_{ow} is the current velocity of the water surface. The aforementioned self excited hydrodynamic forces f_1, f_2, f_3 and f_4 can be computed by the use of Equations (10)-(13).

3. Optimal state derivative feedback LQR controller design

In the last decade, the state derivative feedback has been extensively studied. Design of a full state derivative feedback control with state derivative estimator for acceleration feedback has been presented by Kwak et al. [9]. Abdelaziz and Valasek have developed a procedure to design state derivative feedback controller for pole placement of single input single output linear systems [10]. Then, design of robust pole placement and optimal regulator problems with state derivative feedback has been presented in [11], [12]. Linear Matrix Inequalities (LMIs) based solvability conditions for state derivative feedback has been firstly designed by Assunçao et al. [13]. Faria et al. have extended the problem with regional pole placement [14]. L_2 gain state derivative feedback controller design has been formulated via LMIs by Sever and Yazici [15].The proposed L_2 gain state derivative feedback control approach has been extended with robustness against polytopic type uncertainties [16]. Despite the fact that LMI based solutions of classical LQR approach is widely used in the literature [17], [18], [19], design of a state derivative feedback LQR via LMIs has not been considered so far as provided in this paper.

In this section, first an optimal state derivative feedback LQR controller synthesis is presented. Then, to compare the proposed state derivative feedback LQR controller, design of a classical LQR controller synthesis is provided.

Consider the linear time-invariant system described by

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{14}$$

where $x(t) \in \Re^n$ is the state vector and $u(t) \in \Re^m$ is thecontrol input vector. Our goal is to find an optimal state derivative feedback control in the form of

$$u(t) = -K\dot{x}(t) \tag{15}$$

where $K \in \Re^{m\times n}$ is a controller gain matrix. The closed-loop system is written in the reciprocal state space framework [9] by replacing the (15) into (14) as follows.

$$x(t) = A^{-1}(I + BK)\dot{x}(t) \tag{16}$$

The quadratic cost function is given by

$$J = \int_0^\infty \dot{x}(t)^{\mathrm{T}} Q\dot{x}(t) + u(t)^{\mathrm{T}} Ru(t)dt. \tag{17}$$

here, $Q \in \Re^{n\times n}$ and $R \in \Re^{m\times m}$ are the performance weight matrices. The following theorem presents a LMI based method to design optimal state derivative feedback LQR controller.

Theorem. *For a given values of Q and R, asymptotic stability of the reciprocal state space closed-loop system (16) is ensured with a minimum value of the quadratic cost function (17), if there exists a solution for the following optimization problem*

min tr(M), s.t. (18) and (19)

$$\begin{bmatrix} A^{-1}S + A^{-1}BW + SA^{-\mathrm{T}} + W^{\mathrm{T}}B^{\mathrm{T}}A^{-\mathrm{T}} & S & W^{\mathrm{T}} \\ * & -Q^{-1} & 0 \\ * & * & R^{-1} \end{bmatrix} < 0 \tag{18}$$

$$\begin{bmatrix} M & I \\ * & S \end{bmatrix} > 0. \tag{19}$$

Then, the optimal control law can be calculated as $u(t) = -K\dot{x}(t) = -WS^{-1}\dot{x}(t)$.

Proof. By substituting the (15) into (17), the cost function is turned into

$$J = \int_0^\infty \dot{x}(t)^T \left(Q + K^T RK\right)\dot{x}(t)dt. \tag{20}$$

Suppose that a positive definite matrix P exists which satisfies the equation (21) [20]

$$\dot{x}(t)^T \left(Q + K^T RK\right)\dot{x}(t) = -\frac{d}{dt}\left(\dot{x}^T(t)P\dot{x}(t)\right). \tag{21}$$

By integrating the (21)

$$\begin{aligned} J &= -\dot{x}^T(t)P\dot{x}(t)\Big|_0^\infty = \\ &- \dot{x}^{\mathrm{T}}(\infty)P\dot{x}(\infty) + \dot{x}^T(0)P\dot{x}(0) \end{aligned} \tag{22}$$

is obtained. Under the assumption of the closed-loop system (16) is asymptotically stable, the cost function converges to the

$$J = \dot{x}(0)^T P\dot{x}(0). \tag{23}$$

Hence, the equation (21) can be rewritten as follows

$$\begin{aligned} &\dot{x}(t)^T \left(Q + K^T RK\right)\dot{x}(t) = \\ &-\left(\ddot{x}^T(t)P\dot{x}(t) + \dot{x}^T(t)P\ddot{x}(t)\right). \end{aligned} \tag{24}$$

By using the closed-loop system (16) in reciprocal state space framework, (24) is converted to

$$\begin{aligned} &\dot{x}(t)^T \left(Q + K^T RK\right)\dot{x}(t) = \\ &-\dot{x}^T(t)\begin{pmatrix} PA^{-1}(I + BK) + \\ (I + BK)^T A^{-T}P \end{pmatrix}\dot{x}(t). \end{aligned} \tag{25}$$

Design of an optimal state derivative feedback controller problem can be cast to the matrix inequality constraint problem by change of variables. Let us define a new variable $Y=Y^T>P$. Then, substituting Y into (25) allows us to write

$$\begin{aligned} &YA^{-1} + A^{-\mathrm{T}}Y + YA^{-1}BK \\ &+ K^{\mathrm{T}}B^{\mathrm{T}}A^{-\mathrm{T}}Y + K^{\mathrm{T}}RK + Q < 0. \end{aligned} \tag{26}$$

By applying the Schur complement formula [21], (26) is congruent to

$$\begin{bmatrix} YA^{-1}+YA^{-1}BK+ \\ A^{-\mathrm{T}}Y+K^{\mathrm{T}}B^{\mathrm{T}}A^{-\mathrm{T}}Y & I & K^{\mathrm{T}} \\ * & -Q^{-1} & 0 \\ * & * & R^{-1} \end{bmatrix} < 0. \quad (27)$$

(27) is not in the LMI form yet due to the multiplication of decision variables *Y* and *K*. Pre and post multiply the (27) by $diag(S,I,I)$ where $S=S^T=Y^{-1}$ and (18) is obtained. Here, *W=KS* is a modest variable change operation. Recall that the quadratic cost function (23) has to be minimized by optimal state derivative feedback control law (15). Then, a new decision variable $M \in \Re^{c\times c}$ is introduced to set an upper bound on the cost as follows:

$$M > Y \leftrightarrow \begin{bmatrix} M & I \\ * & S \end{bmatrix} > 0. \quad (28)$$

In the light of the results obtained above, the proof is completed. □

3.1. Classical LQR controller design

In order to compare the effectiveness of the proposed state derivative feedback LQR controller, a classical LQR controller has been designed in this subsection.

As it is well known that LQR control problem is to find an optimal state feedback control law that minimized the quadratic cost function with the solution of following Algebraic Riccati Equation [22],

$$SA + A^{\mathrm{T}}S + Q + SBR^{-1}B^{\mathrm{T}}S = 0\,. \quad (29)$$

Then, the classical state feedback control law can be obtained as

$$u_c(t) = K_c x(t) = -R^{-1}B^{\mathrm{T}}Sx(t)\,. \quad (30)$$

4. Numerical examples

In this section, extensive numbers of simulations are carried out to verify the effectiveness and applicability of the proposed controller to a offshore steel jacket platform subject to nonlinear wave disturbance. The parameters of the considered offshore steel jacket platform having ATMD are taken from [2], [23] and listed in Table 1. In addition, non-linear self excited wave force $w(x,t)$ have been computed as Appendix A in [2]. All the simulations and computations are employed using Matlab with Simulink.

The matrices *A*and B_u of the modeled offshore steel jacket platform can be written by the use of these system parameters as

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -3.3235 & -0.0212 & 0.0184 & 0.0030 & -5.3449 & -0.8819 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0.0184 & 0.0030 & -118.135 & -0.1118 & 5.3465 & 0.8822 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -0.0114 & -0.0019 & 0.0114 & 0.0019 & -3.3051 & -0.5454 \end{bmatrix}$$

$$B_u = \begin{bmatrix} 0 & 0.003445 & 0 & -0.00344628 & 0 & 0.00213 \end{bmatrix}^{\mathrm{T}}.$$

Table 1. Parameter values of considered offshore steel jacket platform.

Parameter	Value
Wave Height (*H*)	12.19 m
Wave Length (λ)	182.88 m
Depth of Water (*h*)	76.2 m
Current Velocity (U_{ow})	0.122m/s
Wave Frequency (ω)	1.8 rad/s
Natural Frequency (ω_1)	1.818 rad/s
Natural Frequency (ω_2)	10.8683 rad/s
Damping Ratio (ξ_1)	0.005
First Mode Shape(ϕ_1)	-0.003445
Second Mode Shape (ϕ_2)	0.00344628
Damping Ratio (ξ_2)	0.005
Natural Frequency of ATMD (ω_T)	1.8180 rad/s
Mass of ATMD (m_T)	469.4836 kg
Stiffness of ATMD (K_T)	1551.5
Damping of ATMD (C_T)	256
Damping Ratio (ξ_T)	0.15

The performance weight matrices which are used in controller design are

$$Q = diag(500,500,500,500,500,500), \quad R = 0.001. \quad (31)$$

In the light of Theorem, in order to minimize the quadratic cost given by (31), proposed controller is designed. For the solution of the resulting LMIs,Yalmip parser and Sedumi solver are used [24], [25].Thus, the optimal state derivative feedback LQR control law is computed as

$$u(t) = -Kx(t) = -10^3 \times [1.1282 \quad -0.3589 \quad -0.1116 \quad -0.0110 \quad 2.0639 \quad 1.4837]x(t). \quad (32)$$

For brevity from this point onwards we will henceforth denote this controller as SDFLQR.

In addition to compare the performance of the proposed controller, the classical LQR controller has been designed for given system (14) with the performance weighting matrices (31). The resulted classical LQR control law can be obtained as

$$\begin{aligned} u_c(t) &= K_c x(t) \\ &= 10^3 \times [-1.3936 \quad -0.2053 \quad 2.9085 \\ &\quad 0.6227 \quad -0.1532 \quad -1.5248]\, x(t). \end{aligned} \tag{33}$$

For brevity from this point onwards we will henceforth denote this controller as LQR. The displacement responses of the first, second and third floors of the offshore steel jacket platform are shown in Figure 2, Figure 3 and Figure 4, repectively for the controlled and uncontrolled cases against the nonlinear wave forces.

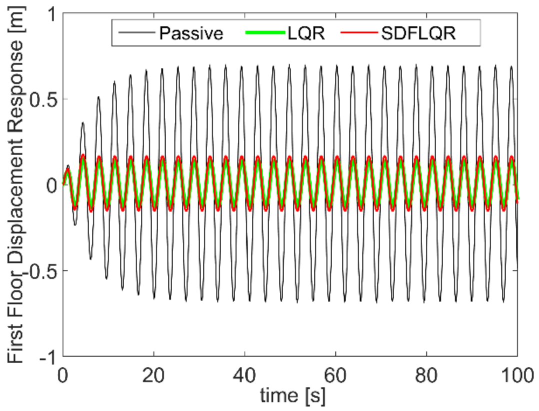


Figure 2. Controlled and uncontrolled displacemet time responses of first floor of offshore steel jacket platform.

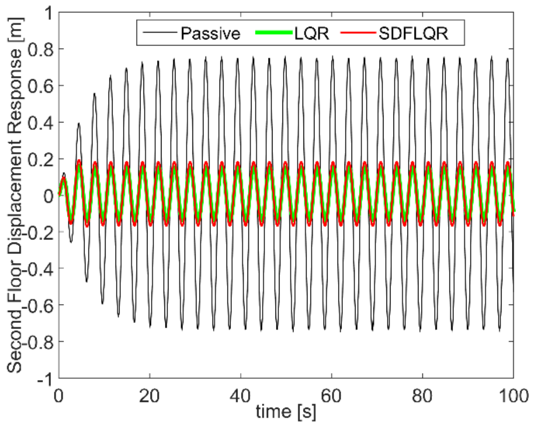


Figure 3. Controlled and uncontrolled displacemet time responses of second floor of offshore steel jacket platform.

As shown in Figure 2, Figure3 and Figure 4, vibration amplitudes of each floor of the offshore steel jacket platform are suppressed successfully by the use of SDFLQR and classical LQR. On the other hand Figure 5 demonstrates the cahange in control inputs for SDFLQR and classical LQR.

When the response plots of the offshore steel jacket platform with uncontrolled and controlled cases are compared,SDFLQR and classical LQR have very close vibration suppression performance. On the other hand, by taking into account that the state derivative signals are much available to obtain good accuracy, proposed SDFLQR is very promising solution for active vibration control of offshore steel jacket platform having nonlinear self excited wave induced disturbances.

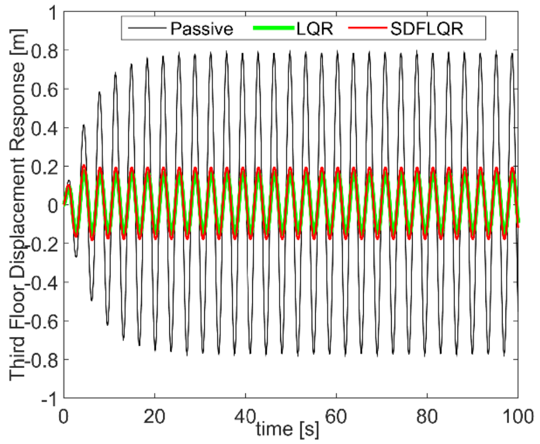


Figure 4. Controlled and uncontrolled displacemet time responses of the third floor of offshore steel jacket platform.

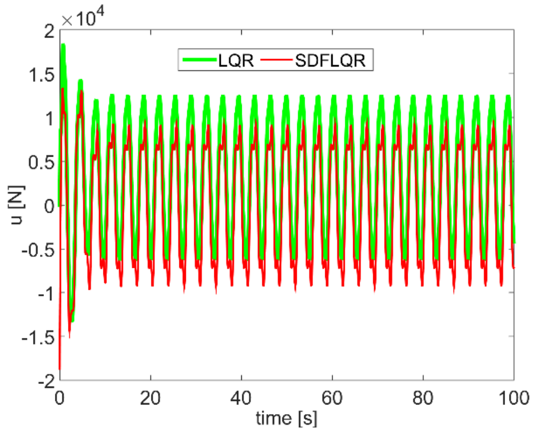


Figure 5. Time history of the applied control force for SDFLQR and LQR.

In this section, the root mean square (RMS) value, which is statistic measure of the magnitude of varying quantity, is employed to investigate the active vibration control performance. RMS analysis method is very useful to evaluate active control performance when the variants are positive and negative [26]. The corresponding RMS values of displacement responses of each floor of the considered offshore steel jacket platform and applied control forces are compared for the both controlled and uncontrolled cases in Table 2 for nonlinear wave disturbance input.

Remark: As can be observed from Table 2, proposed SDFLQR achieves compitable control performance to classical LQR control method. Note that the system response and the control effort is not equally invloved in the quadratic cost functions of both LQR and SDFLQR. The control input is weighted with a state derivative vector as $\dot{x}(t)$ for SDFLQR and weighted with a state vector as $x(t)$ for LQR. Therefore, similar choice of Q and R, results dissimilar performance objectives for SDFLQR and LQR. In the light of aforementioned discussions, it is natural to have slighlty different vibration attenuation levels. It is noteworthy that previously applied control approaches in the literature are applied state feedback to actively control the steel jacket platform vibrations [4], [5], [6], [7] and [8]. Accessing the displacement information from accelerometer outputs by a double integration, is required to realize state feedback control law. Considering that state derivative signals are much available to obtain good accuracy, proposed SDFLQR design process provides more practically applicable and easily realizable synthesis method and has a great potential for active vibration control of offshore steel jacket platform.

Table 2. Comparison of RMS values of displacement responses of each floor of the offshore steel jacket platform and applied control forces for the both controlled and uncontrolled cases.

Displacements and Applied Control Force	RMS Values		
	Passive	LQR	SDFLQR
First Floor (m)	0.4589	0.0967	0.1147
Second Floor (m)	0.4969	0.1056	0.1249
Third Floor (m)	0.5212	0.1122	0.1322
Control Force (N)	–	7.6782×10^3	7.1554×10^3

5. Conclusion

This paper presents an approach for designing state derivative feedback LQR controller to attenuate the vibration occurred in offshore steel jacket platform against the nonlinear wave forces. In controller design, the solvability conditions of the proposed control strategy is presented as LMI constraints on the basis of convex optimization approach. The main importance of this study is to devoloped an easily realizable synthesis method to obtain practically applicable optimal state derivative LQR controller which provides satisfactory control performance. In order to demonstare the effectiveness of the approach, performance of the proposed controller is examined in disturbance attenuation of nonlinear wave force excitations, in an offshore steel jacket paltform having ATMD. Simulation results indicate that the proposed control technique is all effective in reducing vibration amplitudes of each floor and guarantees the closed-loop stability. Vibration attenuation performance of the proposed controller can be improved by employing pole location constraints via LMI regions and H_2/H_∞ norm conditions. Finally, to cope with the practical problems such as parametric uncertainties and actuator imperfections, expanding the proposed method with the robustness against actuator delay and uncertain parameters might be a significant direction for future work.

Acknowledgments

The authors would like to thank the Associate Editor and anonymous reviewers for their valuable comments.

References

[1] Abdel-Rohman, M.(1996). Structural Control of Steel Jacket Platform. *Structural Engineering and Mechanics*, 4(2), 125-138.

[2] Terro, M.J., Mahmoud, M.S., Abdel-Rohman, M. (1999). Multi-loop Feedback Control of Offshore Steel Jacket Platforms. *Computers and Structures*, 70(2), 188-202.

[3] Zribi, M., Almutari, N., Abdel-Rohman, M., Terro, M. (2004). Nonlinear and Robust Control Schemes for offshore platforms. *Nonlinear Dynamics*, 35(1), 61-80.

[4] Wu, H.Z., Bao-lin, Z., Hui, M.A. (2012). Active Non-Fragile Control for Offshore Steel Jacket Platforms. Proceedings of the 31st Chinese Control Conference, 7481-7486, July 25-27, Hefei, China.

[5] Mei , X.Y., Bao-Lin, Z., Jian-Cun, W., Xiefu, J. (2016). Fuzzy H_∞ Control for Steel Jacket Platforms with Parameter Uncertainties, Proceedings of the 35st Chinese Control Conference, 3749-3754, July 27-29, Chengdu, China.

[6] Zhang B.L, Ma L., Han Q.L. (2013). Sliding mode H_∞ Control for offshore steel jacket platforms subject to nonlinear self excited wave force and external disturbance. *Nonlinear Analysis: Real World Applications*, 14(1), 163-178.

[7] Li, H.J., Lan S., Hu, J., Jakubiak, C. (2003). H_2 active vibration control for offshore platform subject to wave loading. *Journal of Sound and Vibration*, 263(4), 709-724.

[8] Zhang B.L., Hu Y.H., Tang, G.Y. (2012). Stabilization control for offshore steel jacket platforms with actuator time-delays. *Nonlinear Dynamics*, 70(2), 1593-1603.

[9] Kwak S.K., Washington G., Yedavalli R.K. (2002). Acceleration-based vibration control of distributed parameter systems using the “reciprocal state-space framework”. *Journal of Sound and Vibration*, 251(3), 543-557.

[10] Abdelaziz T.H.S. and Valasek M. (2004). Pole placement for SISO linear systems by state derivative feedback. *IEE Proceedings-Control Theory and Applications*, 151(4), 377-385.

[11] Abdelaziz T.H.S. (2009). Robust pole assignment for linear time-invariant systems using state-derivative feedback. *Journal of Systems and*

Control Engineering, 223(2), 187-199.

[12] Abdelaziz T.H.S. (2010). Optimal control using derivative feedback for linear systems. *Journal of Systems and Control Engineering* 224(2), 187-202.

[13] Assunçao E, Teixeira MCM, Faria FA et al. (2007). Robust state derivative feedback LMI-based designs for multivariable linear systems. *International Journal of Control*, 80(8), 1260-1270.

[14] Faria FA, Assunçao E, Teixeira MCM et al. (2009). Robust state derivative pole placement LMI based designs for linear systems. *International Journal of Control*, 82(1), 1-12.

[15] Sever M., Yazici H. (2017) Active Control of Vehicle Suspension System Having Driver Model via L_2 Gain State Derivative Feedback Controller. 2017 4th International Conference on Electrical and Electronics Engineering (ICEEE 2017), 215-222, April 8-10,Ankara, Turkey,

[16] Yazici H, Sever M. (2017). L_2 gain state derivative feedback control of uncertain vehicle suspension systems. *Journal of Vibration and Control*,In Press.

[17] Aktas A, Sever M, Yazici H. (2016). Gain scheduling LQR control of linear parameter varying overhead crane. In: IEEE 2016 National Conference on Electrical, Electronics and Biomedical Engineering,232-236, Dec 1–3; Bursa, Turkey.

[18] Sever M, Kaya EE, Arslan MS, Yazici H. (2016). Active trailer braking system design with linear matrix inequalities based multi objective robust LQR controller for vehicle-trailer systems. In: IEEE 2016 Intelligent Vehicles Symposium,726-731, June 19–22, Gothenburg, Sweeden.

[19] Yazici H, Sever M. (2017). Active control of a non-linear landing gear system having oleo pneumatic shock absorber using robust linear quadratic regulator approach. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*,In Press.

[20] Abdelaziz THS, Valasek M. (2005). State derivative feedback by LQR for linear time invariant systems. *IFAC Proceeding Volumes*, 38(1), 435-440.

[21] Boyd, S., Ghaoui, L.E., Feron, E., Balakrishnan, V. (1994)*Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia.

[22] Lin, F. (2007).*Robust Control Design an Optimal Control Approach*. John Willey&Sons, Hartfordshire.

[23] Bao-Lin, Z., Han, Q.L. (2014). Network-based Modelling and Active Control for Offshore Steel Jacket Platform with TMD Mechanisms. *Journal of Sound and Vibrations*, 333(25), 6796-6814.

[24] Löfberg, J. (2004) Yalmip: A Toolbox for Modeling and Optimization in MATLAB. Proceedings of the CACSD Conference, Taipei, Taiwan.

[25] Strum, J.F. (1999). Using SeDuMi 1.02 a Matlab for optimization over symmetric cones. Optimization Methods and Software, 11(2), 625-653.

[26] Zhao, Y., Sun, W., Gao, H. (2010). Robust control synthesis for seat suspension systems with actuator saturation with time varying input delay. *Journal of Sound and Vibration*, 329(21), 4335-4353.

Hakan Yazici *received his M.Sc and Ph.D. from Yildiz Technical University, all in Mechanical Engineering, in 2006 and 2011, respectively. He conducted post-doctoral research at Southern Illinois University Edwardsville, USA, from 2011 to 2012. His current research interest includes robust and optimal control of time-delay systems with actuator saturation, stability of functional differential equations with time delays, and active vibration control and its applications..*

Mert Sever *was born in Istanbul, Turkiye, in 1990. He received his B.Sc from Yildiz Technical University, in Mechanical Engineering, in 2014 and is a Msc. Candidate in Machine Theory and Control Program, in the same university, respectively. He is currently a Research Assistant at Yildiz Technical University, Department of Mechatronics Engineering. His current research interest includes LMI based robust and optimal control, active vibration control and its applications.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.1, pp.92-101 (2018)
http://doi.org/10.11121/ijocta.01.2018.00512

RESEARCH ARTICLE

The structure of one weight linear and cyclic codes over $\mathbb{Z}_2^r \times (\mathbb{Z}_2 + u\mathbb{Z}_2)^s$

İsmail Aydoğdu*

Department of Mathematics, Faculty of Arts and Sciences, Yıldız Technical University, İstanbul, Turkey
iaydogdu@yildiz.edu.tr

ARTICLE INFO

Article History:
Received 04 July 2017
Accepted 20 December 2017
Available 26 December 2017

Keywords:
One weight codes
$\mathbb{Z}_2\mathbb{Z}_2[u]$*-linear codes*
Duality

AMS Classification 2010:
94B05, 94B60

ABSTRACT

Inspired by the $\mathbb{Z}_2\mathbb{Z}_4$-additive codes, linear codes over $\mathbb{Z}_2^r \times (\mathbb{Z}_2 + u\mathbb{Z}_2)^s$ have been introduced by Aydogdu et al. more recently. Although these family of codes are similar to each other, linear codes over $\mathbb{Z}_2^r \times (\mathbb{Z}_2 + u\mathbb{Z}_2)^s$ have some advantages compared to $\mathbb{Z}_2\mathbb{Z}_4$-additive codes. A code is called constant weight (one weight) if all the nonzero codewords have the same weight. It is well known that constant weight or one weight codes have many important applications. In this paper, we study the structure of one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear and cyclic codes. We classify one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes and also give some illustrative examples.

1. Introduction

In algebraic coding theory, the most important class of codes is the family of linear codes. A linear code of length n is a subspace $\mathcal{C}$ of a vector space F_q^n where F_q is a finite field of size q. When $q = 2$ then we have linear codes over F_2 which are called binary codes. Binary linear codes have very special and important place all among the finite field codes because of their easy implementations and applications. Beginning with a remarkable paper by Hammons et al. [1], interest of codes over variety of rings have been increased. Such studies motivate the researchers to work on different rings even over other structural algebras such as groups or modules. A $\mathbb{Z}_4$-submodule of $\mathbb{Z}_4^n$ is called a quaternary linear code. The structure of binary linear codes and quaternary linear codes have been studied in details for the last two decades. The reader can see some of them in [2–4]. In 2010, Borges et al. introduced a new class of error correcting codes over the ring $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$ called additive codes that generalizes the class of binary linear codes and the class of quaternary linear codes in [5]. A $\mathbb{Z}_2\mathbb{Z}_4$-additive code $\mathcal{C}$ is defined to be a subgroup of $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$ where $\alpha + 2\beta = n$. If $\beta = 0$ then $\mathbb{Z}_2\mathbb{Z}_4$-additive codes are just binary linear codes, and if $\alpha = 0$, then $\mathbb{Z}_2\mathbb{Z}_4$-additive codes are the quaternary linear codes over $\mathbb{Z}_4$. $\mathbb{Z}_2\mathbb{Z}_4$-additive codes have been generalized to $\mathbb{Z}_2\mathbb{Z}_{2^s}$-additive codes in 2013 by Aydogdu and Siap in [6], and recently this generalization has been extended to $\mathbb{Z}_{p^r}\mathbb{Z}_{p^s}$-additive codes, for a prime p, by the same authors in [7]. Later, cyclic codes over $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$ have been introduced in [8] in 2014 and more recently, in [9], one weight codes over such a mixed alphabet have been studied. A code $\mathcal{C}$ is said to be one weight code if all the nonzero codewords in $\mathcal{C}$ have the same Hamming weight where the Hamming weight of any string is the number of symbols that are different from the zero symbol of the alphabet used. In [10], Carlet determined one weight linear codes over $\mathbb{Z}_4$ and in [11], Wood studied linear one weight codes over $\mathbb{Z}_m$. Constant weight codes are very useful in a variety of applications such as data storage, fault-tolerant circuit design and computing, pattern generation for circuit testing, identification coding, and optical overlay networks [12].

*Corresponding Author

Moreover, the reader can find the other applications of constant weight codes; determining the zero error decision feedback capacity of discrete memoryless channels in [13], multiple access communications and spherical codes for modulation in [14,15], DNA codes in [16,17], powerline communications and frequency hopping in [18].

Another important ring of four elements other than the ring $\mathbb{Z}_4$, is the ring $\mathbb{Z}_2 + u\mathbb{Z}_2 = R = \{0, 1, u, 1+u\}$ where $u^2 = 0$. For some of the works done in this direction we refer the reader to [19–21]. It has been shown that linear and cyclic codes over this ring have advantages compared to the ring $\mathbb{Z}_4$. For an example; the finite field $GF(2)$ is a subring of the ring R. So factorization over $GF(2)$ is still valid over the ring R. The Gray image of any linear code over R is always a binary linear code which is not always the case for $\mathbb{Z}_4$.

In this work, we are interested in studying one weight codes over $\mathbb{Z}_2^r \times (\mathbb{Z}_2 + u\mathbb{Z}_2)^s = \mathbb{Z}_2^r \times R^s$. This family of codes are special subsets of $\mathbb{Z}_2^r \times R^s$ which their all nonzero codewords have the same weight. Since the structure of one weight binary linear codes were well classified by Bonisoli [22], we conclude some results that coincides with the results in [22] for $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear codes, and we classify cyclic codes over $\mathbb{Z}_2^r \times R^s$ and also we give some one weight linear and cyclic code examples. Furthermore, we look at the Gray (binary) images of one weight cyclic codes over $\mathbb{Z}_2^r \times R^s$ and we determine their parameters.

2. Preliminaries

Let $R = \mathbb{Z}_2 + u\mathbb{Z}_2 = \{0, 1, u, 1+u\}$ be the four-element ring with $u^2 = 0$. It is easily seen that the ring $\mathbb{Z}_2$ is a subring of the ring R. Then let us define the set

$$\mathbb{Z}_2\mathbb{Z}_2[u] = \{(a, b) \mid a \in \mathbb{Z}_2 \text{ and } b \in R\}.$$

But we have a problem here, because the set $\mathbb{Z}_2\mathbb{Z}_2[u]$ is not well-defined with respect to the usual multiplication by $u \in R$. So, we must define a new method of multiplication on $\mathbb{Z}_2\mathbb{Z}_2[u]$ to make this set as an R-module. Now define the mapping

$$\begin{aligned} \eta &: R \to \mathbb{Z}_2 \\ \eta(p + uq) &= p. \end{aligned}$$

which means; $\eta(0) = 0$, $\eta(1) = 1$, $\eta(u) = 0$ and $\eta(1+u) = 1$. It can be easily shown that η is a ring homomorphism. Furthermore, for any element $e \in R$, we can also define a scalar multiplication on $\mathbb{Z}_2\mathbb{Z}_2[u]$ as follows.

$$e\,(a, b) = (\eta(e)a, eb).$$

This multiplication can be extended to $\mathbb{Z}_2^r \times R^s$ for $e \in R$ and $v = (a_0, a_1, ..., a_{r-1}, b_0, b_1, ..., b_{s-1}) \in \mathbb{Z}_2^r \times R^s$ as,

$$ev = (\eta(e)a_0, \eta(e)a_1, ..., \eta(e)a_{r-1}, eb_0, eb_1, ..., eb_{s-1}).$$

Lemma 1. *$\mathbb{Z}_2^r \times R^s$ is an $R-$module under the multiplication defined above.*

Definition 1. *A non-empty subset $\mathcal{C}$ of $\mathbb{Z}_2^r \times R^s$ is called a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code if it is an R-submodule of $\mathbb{Z}_2^r \times R^s$.*

Now, take any element $a \in R$, then there exist unique $p, q \in \mathbb{Z}_2$ such that $a = p + uq$. Also note that the ring R is isomorphic to $\mathbb{Z}_2^2$ as an additive group. Therefore, any $\mathbb{Z}_2\mathbb{Z}_2[u]-$linear code $\mathcal{C}$ is isomorphic to an abelian group of the form $\mathbb{Z}_2^{k_0+k_2} \times \mathbb{Z}_2^{2k_1}$, where k_0, k_2 and k_1 are nonnegative integers. Now define the following sets.

$$\mathcal{C}_s^F = \langle \{(a, b) \in \mathbb{Z}_2^r \times R^s \mid b \text{ free over } R^s\} \rangle$$

where if $\langle b \rangle = R^s$ then b is called free over R^s.

$$\begin{aligned} \mathcal{C}_0 &= \langle \{(a, ub) \in \mathbb{Z}_2^r \times R^s \mid a \neq 0\} \rangle \subseteq \mathcal{C} \backslash \mathcal{C}_s^F \\ \mathcal{C}_1 &= \langle \{(a, ub) \in \mathbb{Z}_2^r \times R^s \mid a = 0\} \rangle \subseteq \mathcal{C} \backslash \mathcal{C}_s^F. \end{aligned}$$

Therefore, denote the dimension of $\mathcal{C}_0$, $\mathcal{C}_1$ and $\mathcal{C}_s^F$ as k_0, k_2 and k_1 respectively. Under these parameters, we say that such a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code $\mathcal{C}$ is of type $(r, s; k_0, k_1, k_2)$. $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear codes can be considered as binary codes under a special Gray map. For $(x, y) \in \mathbb{Z}_2^r \times R^s$, where $(x, y) = (x_0, x_1, \ldots, x_{r-1}, y_0, y_1, \ldots, y_{s-1})$ and $y_i = p_i + uq_i$ the Gray map is defined as follows.

$$\begin{aligned} &\Phi : \mathbb{Z}_2^r \times R^s \to \mathbb{Z}_2^n \\ &\Phi(x_0, \ldots x_{r-1}, p_0 + uq_0, \ldots p_{s-1} + uq_{s-1}) \\ &= (x_0, \ldots x_{r-1}, q_0, \ldots, q_{s-1}, p_0 + q_0, \ldots, p_{s-1} + q_{s-1}), \end{aligned} \tag{1}$$

where $n = r + 2s$.

The Hamming distance between two strings x and y of the same length over a finite alphabet Σ denoted by $d(x, y)$ is defined as the number of positions at which these two strings differ. The Hamming weight of a string x over an alphabet Σ is

defined as the number of its nonzero symbols in the string. More formally, the Hamming weight of a string is $wt(x) = |\{i : x_i \neq 0\}|$. Also note that $wt(x-y) = d(x,y)$.
The minimum distance of a linear code $\mathcal{C}$, denoted by $d(\mathcal{C})$ is defined by

$$d(\mathcal{C}) = min\{d(c_1,c_2) : c_1,\ c_2 \in \mathcal{C}, c_1 \neq c_2\}.$$

The Lee distance for the codes over R is the Lee weight of their differences where the Lee weights of the elements of R are defined as $wt_L(0) = 0,\ wt_L(1) = 1,\ wt_L(u) = 2$ and $wt_L(1+u) = 1$. The Gray map defined above is a distance preserving map which transforms the Lee distance in $\mathbb{Z}_2^r \times R^s$ to the Hamming distance in $\mathbb{Z}_2^n$. Furthermore, for any $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code $\mathcal{C}$, we have that $\Phi(\mathcal{C})$ is a binary linear code as well. This property is not valid for the $\mathbb{Z}_2\mathbb{Z}_4$-additive codes. And also, we define

$$wt(v) = wt_H(v_1) + wt_L(v_2),$$

where $v = (v_1, v_2)$, $wt_H(v_1)$ is the Hamming of weight of v_1 and $wt_L(v_2)$ is the Lee weight of v_2. If $\mathcal{C}$ is a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code of type $(r,s;k_0,k_1,k_2)$ then the binary image $C = \Phi(\mathcal{C})$ is a binary linear code of length $n = r+2s$ and size 2^n. It is also called a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code. Now, let $v = (a_0,\ldots,a_{r-1},b_0,\ldots,b_{s-1})$, $w = (d_0,\ldots,d_{r-1},e_0,\ldots,e_{s-1}) \in \mathbb{Z}_2^r \times R^s$ be any two elements. Then we can define the inner product as

$$\langle v,w\rangle = \left(u\sum_{i=0}^{r-1} a_i d_i + \sum_{j=0}^{s-1} b_j e_j\right) \in \mathbb{Z}_2 + u\mathbb{Z}_2.$$

According to this inner product, the dual linear code $\mathcal{C}^\perp$ of a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code $\mathcal{C}$ is also defined in a usual way,

$$\mathcal{C}^\perp = \{w \in \mathbb{Z}_2^r \times R^s |\ \langle v,w\rangle = 0 \text{ for all } v \in \mathcal{C}\}.$$

Hence, if $\mathcal{C}$ is a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code, then $\mathcal{C}^\perp$ is also a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code.
The standard forms of generator and parity-check matrices of a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code $\mathcal{C}$ are given as follows.

Theorem 1. *[23] Let $\mathcal{C}$ be a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code of type $(r,s;k_0,k_1,k_2)$. Then the standard forms of the generator and the parity-check matrices of $\mathcal{C}$ are:*

$$G = \left[\begin{array}{cc|ccc} I_{k_0} & A_1 & 0 & 0 & uT \\ 0 & S & I_{k_1} & A & B_1 + uB_2 \\ 0 & 0 & 0 & uI_{k_2} & uD \end{array}\right]$$

$$H = \left[\begin{array}{cc|ccc} -A_1^t & I_{r-k_0} & -uS^t & 0 & 0 \\ -T^t & 0 & -(B_1+uB_2)^t + D^tA^t & -D^t & I_{s-k_1-k_2} \\ 0 & 0 & -uA^t & uI_{k_2} & 0 \end{array}\right]$$

where $A,\ A_1,\ B_1,\ B_2,\ D,\ S$ and T are matrices over $\mathbb{Z}_2$.

Therefore, we can conclude the following corollary.

Corollary 1. *If $\mathcal{C}$ is a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code of type $(r,s;k_0,k_1,k_2)$ then $\mathcal{C}^\perp$ is of type $(r,s;r-k_0,s-k_1-k_2,k_2)$.*

The weight enumerator of any $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code $\mathcal{C}$ of type $(r,s;k_0,k_1,k_2)$ is defined as

$$W_{\mathcal{C}}(x,y) = \sum_{c\in\mathcal{C}} x^{n-wt(c)} y^{wt(c)}$$

where, $n = r+2s$. Moreover, the MacWilliams relations for codes over $\mathbb{Z}_2\mathbb{Z}_2[u]$ can be given as follows.

Theorem 2. *[23] Let $\mathcal{C}$ be a $\mathbb{Z}_2\mathbb{Z}_2[u]-$linear code. The relation between the weight enumerators of $\mathcal{C}$ and its dual is*

$$W_{\mathcal{C}^\perp}(x,y) = \frac{1}{|\mathcal{C}|} W_{\mathcal{C}}(x+y, x-y).$$

We have given some information about the general concept of codes over $\mathbb{Z}_2^r \times (\mathbb{Z}_2 + u\mathbb{Z}_2)^s$. To make reader understanding the paper easily we give the following example.

Example 1. *Let $\mathcal{C}$ be a linear code over $\mathbb{Z}_2^3 \times (\mathbb{Z}_2 + u\mathbb{Z}_2)^4$ with the following generator matrix.*

$$\left[\begin{array}{ccc|cccc} 1 & 1 & 0 & 0 & u & u & u \\ 0 & 1 & 1 & 1 & 1+u & u & 0 \\ 0 & 1 & 0 & u & u & u & 0 \end{array}\right].$$

We will find the standard form of the generator matrix of $\mathcal{C}$ and then using this standard form, we find the generator matrix of the linear dual code $\mathcal{C}^\perp$ and also we determine the types of both $\mathcal{C}$ and its dual.
Now, applying elementary row operations to above generator matrix, we have the standard form as follows.

$$G = \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 0 & u & 0 & u \\ 0 & 1 & 0 & 0 & 0 & u & 0 \\ 0 & 0 & 1 & 1 & 1+u & 0 & 0 \end{array}\right].$$

Since, G is in the standard form we can write this matrix as

$$G = \left[\begin{array}{ccc||cccc} 1 & 0 & 0 & 0 & u & 0 & u \\ 0 & 1 & 0 & 0 & 0 & u & 0 \\ \hline 0 & 0 & 1 & 1 & 1+u & 0 & 0 \end{array}\right]$$

$$= \left[\begin{array}{cc||cc} I_{k_0} & A_1 & 0 & uT \\ \hline 0 & S & I_{k_1} & B_1 + uB_2 \end{array}\right].$$

Hence, with the help of Theorem 1 the parity-check matrix of $\mathcal{C}$ *is*

$$H = \left[\begin{array}{ccc|cccc} 0 & 0 & 1 & u & 0 & 0 & 0 \\ 1 & 0 & 0 & 1+u & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

Therefore,

- $\mathcal{C}$ *is of type* $(3,4;2,1,0)$ *and has* $2^2 4^1 = 16$ *codewords.*
- $\mathcal{C}^\perp$ *is of type* $(3,4;1,3,0)$ *and has* $2^1 4^3 = 128$ *codewords.*
- $\mathcal{C} = \{(0,0,0,|0,0,0,0), (1,0,0,|0,u,0,u),$ $(0,1,0,|0,0,u,0), (0,0,1,|1,\bar{u},0,0),$ $(0,0,0,|u,u,0,0), (0,0,1,|\bar{u},1,0,0),$ $(1,1,0,|0,u,u,u), (1,0,1,|1,1,0,u),$ $(0,1,1,|1,\bar{u},u,0), (1,1,1,|1,1,u,u),$ $(1,0,0,|u,0,0,u), (0,1,0,|u,u,u,0),$ $(1,1,0,|u,0,u,u), (1,0,1,|\bar{u},\bar{u},0,u),$ $(0,1,1,|\bar{u},1,u,0), (1,1,1,|\bar{u},\bar{u},u,u)\}$, *where* $\bar{u} = 1+u$.
- $W_{\mathcal{C}}(x,y) = x^{11} + 3x^8y^3 + x^7y^4 + 2x^6y^5 + 4x^5y^6 + x^4y^7 + 2x^3y^8 + 2x^2y^9$.
- $W_{\mathcal{C}^\perp}(x,y) = \frac{1}{|\mathcal{C}|} W_{\mathcal{C}}(x+y, x-y) = x^{11} + 6x^9y^2 + 8x^8y^3 + 16x^7y^4 + 32x^6y^5 + 26x^5y^6 + 24x^4y^7 + 15x^3y^8$.
- *The Gray image* $\Phi(\mathcal{C})$ *of* $\mathcal{C}$ *is a* $[11,4,3]$ *binary linear code.*
- $\Phi(\mathcal{C}^\perp)$ *is a* $[11,7,2]$ *binary linear code.*

3. The Structure of One Weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear Codes

In this part of the paper, we study the structure of one weight codes over $\mathbb{Z}_2^r \times R^s$. Since the binary(Gray) images of $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear codes are always linear, our results about the one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear codes will coincide with the results of the paper [22]. So, in this section of the paper we will prepare for Section 4 and also we give some fundamental definitions and illustrative examples of one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear codes.

Definition 2. *Let* $\mathcal{C}$ *be a* $\mathbb{Z}_2\mathbb{Z}_2[u]$*-linear code.* $\mathcal{C}$ *is called a one (constant) weight code if all of its nonzero codewords have the same weight. Furthermore, if such weight is* m *then* $\mathcal{C}$ *is called a code with weight* m.

Definition 3. *Let* c_1, c_2, e_1, e_2 *be any four distinct codewords of a* $\mathbb{Z}_2\mathbb{Z}_2[u]$*-linear code* $\mathcal{C}$*. If the distance between* c_1 *and* e_1 *is equal to the distance between* c_2 *and* e_2*, that is,* $d(c_1,e_1) = d(c_2,e_2)$*, then* $\mathcal{C}$ *is said to be equidistant.*

Theorem 3. *[22] Let* C *be a* $[n,k]$ *linear code over* $\mathbb{F}_q$ *with all nonzero codewords of the same weight. Assume that* C *is nonzero and no column of a generator matrix is identically zero. Then* C *is equivalent to the* λ*-fold replication of a simplex (i.e., dual of the Hamming) code.*

Corollary 2. *Let* $\mathcal{C}$ *be an equidistant* $\mathbb{Z}_2\mathbb{Z}_2[u]$*-linear code with distance* m*. Then* $\mathcal{C}$ *is a one weight code with weight* m*. Moreover, the binary image* $\Phi(\mathcal{C})$ *of* $\mathcal{C}$ *is also a one weight code with weight* m.

Example 2. *It is worth to note that the dual of a one weight code is not necessarily a one weight code. Let* $\mathcal{C}$ *be a* $\mathbb{Z}_2\mathbb{Z}_2[u]$*-linear code of type* $(2,2;0,1,0)$ *with* $\mathcal{C} = \langle(1,1|1+u,1+u)\rangle$*. Then* $\mathcal{C} = \{(0,0|0,0), (1,1|1+u,1+u), (1,1|1,1), (0,0|u,u)\}$ *and* $\mathcal{C}$ *is a one weight code with weight* $m = 4$*. On the other hand, the dual code* $\mathcal{C}^\perp$ *is generated by* $\langle(1,0|u,0),(0,1|u,0),(0,0|1,1)\rangle$ *and of type* $(2,2;2,1,0)$*. But* $d(\mathcal{C}^\perp) = 2$ *and* $\mathcal{C}^\perp$ *is not a one weight code.*

Remark 1. *The dual code for length greater than 3 is never a one weight code.*

Example 3. *Let* $\mathcal{C}$ *be a* $\mathbb{Z}_2\mathbb{Z}_2[u]$*-linear code with the standard form of the generator matrix* $\left[\begin{array}{ccc|cc} 1 & 0 & 1 & 0 & u \\ \hline 0 & 1 & 1 & 1 & 1+u \end{array}\right]$*, then* $\mathcal{C}$ *is of type* $(3,2;1,1,0)$ *and one weight code with weight* 4*. Furthermore,* $\Phi(\mathcal{C})$ *is a binary linear code with parameters* $[7,3,4]$*. Here, note that the binary image of* $\mathcal{C}$ *is the binary simplex code of length* 7*, which is the dual of the* $[7,4,3]$ *Hamming code.*

Now, we give a theorem which gives a construction of one weight codes over $\mathbb{Z}_2^r \times R^s$.

Corollary 3. *Let* $\mathcal{C}$ *be a one weight* $\mathbb{Z}_2\mathbb{Z}_2[u]$*-linear code of type* $(r,s;k_0,k_1,k_2)$ *and weight* m*. Then, a one weight code of type* $(\gamma r, \gamma s; k_0, k_1, k_2)$ *with weight* γm *exists, where* γ *is a positive integer.*

Definition 4. *Let* $\mathcal{C}$ *be a* $\mathbb{Z}_2\mathbb{Z}_2[u]$*-linear code. Let* $\mathcal{C}_r$ *(respectively* $\mathcal{C}_s$*) be the punctured code of* $\mathcal{C}$ *by* *deleting the coordinates outside r (respectively s). If $\mathcal{C} = \mathcal{C}_r \times \mathcal{C}_s$ then $\mathcal{C}$ is called separable.*

Corollary 4. *There do not exist separable one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear codes.*

Proof. Since $\Phi(\mathcal{C}_r \times \mathcal{C}_s) = \Phi(\mathcal{C}_r) \times \Phi(\mathcal{C}_s)$, the proof is obvious. □

Corollary 5. *If $\mathcal{C}$ is a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code of type $(r, s; k_0, k_1, k_2)$ with no all zero columns in the generator matrix of $\mathcal{C}$. Then the sum of the weights of all codewords of $\mathcal{C}$ is equal to $\frac{|\mathcal{C}|}{2}(r+2s)$.*

Proof. From [22], since the sums of the weights of a binary linear code $[n, k]$ is $n2^{k-1}$, the sum of the all codewords of $\mathcal{C}$ is

$$\sum_{c\in\mathcal{C}} wt(c) \;=\; r\frac{|\mathcal{C}|}{2} + s|\mathcal{C}| = \frac{|\mathcal{C}|}{2}(r + 2s).$$

□

Corollary 6. *Let $\mathcal{C}$ be a one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code of type $(r, s; k_0, k_1, k_2)$ and weight $\mathtt{m}$. If there is no zero columns in the generator matrix of $\mathcal{C}$, then;*

i) *$\mathtt{m} = \alpha\ 2^{(k_0+2k_1+k_2)-1}$ where α is a positive integer satisfying $(r + 2s) = \alpha\left(2^{k_0+2k_1+k_2} - 1\right)$. In addition, if $\mathtt{m}$ is an odd integer, then r is also odd and $\mathcal{C} = \langle(\underbrace{1\cdots 1}_{r \text{ times}} | \underbrace{u\cdots u}_{s \text{ times}})\rangle$.*

ii) *$d(\mathcal{C}^\perp) \geq 2$. Also, $d(\mathcal{C}^\perp) \geq 3$ if and only if $\alpha = 1$.*

iii) *for $\alpha = 1$, if $|\mathcal{C}| \geq 4$ then $d(\mathcal{C}^\perp) = 3$.*

We have known from the above corollary that if $\mathcal{C}$ is a one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code of type $(r, s; k_0, k_1, k_2)$ and weight $\mathtt{m}$ then there is a positive integer α such that $\mathtt{m} = \alpha\ 2^{(k_0+2k_1+k_2)-1}$, so the minimum distance for a one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code must be even. In the following, we characterize the structure of $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear codes.

Theorem 4. *Let $\mathcal{C}$ be a one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code over $\mathbb{Z}_2^r \times R^s$ with generator matrix G and weight $\mathtt{m}$.*

i) *If $v = (a|b)$ is an any row of G, where $a = (a_0, \ldots, a_{r-1}) \in \mathbb{Z}_2^r$ and $b = (b_0, \ldots, b_{s-1}) \in R^s$, then the number of units(1 or $1 + u$) in b is either zero or $\frac{\mathtt{m}}{2}$.*

ii) *If $v = (a|b)$ and $w = (c|d)$ are two distinct rows of G, where b and d are free over R^s, then the coordinate positions where b has units (1 or $1 + u$) are the same that the coordinate positions where d has units.*

iii) *If $v = (a|b)$ and $w = (c|d)$ are two distinct rows of G, where b and d are free over R^s, then $|\{j : b_j = d_j = 1 \text{ or } 1 + u\}| = |\{j : b_j = 1, d_j = 1+u \text{ or } b_j = 1+u, d_j = 1\}| = \frac{\mathtt{m}}{4}$.*

Proof. i) The weight of $v = (a|b)$ is $wt(v) = wt_H(a) + wt_L(b) = \mathtt{m}$. Since $\mathcal{C}$ is linear $uv = (0|ub)$ is also in $\mathcal{C}$ then, if $ub = 0$ then b does not contain units. If $ub \neq 0$, then $wt(v) = \mathtt{m} = 0 + wt_L(ub)$ and therefore, $wt_L(ub) = 2|\{j : b_j = 1 \text{ or } 1 + u\}| = \mathtt{m}$. Hence, the number of units in b is $\frac{\mathtt{m}}{2}$.

ii) Multiplying v and w by u we have, $uv = (0|ub)$ and $uw = (0|ud)$. If v and w have units in the same coordinate positions then we get $uv + uw = 0$. So, assume that they have some units in different coordinates. Since $\mathcal{C}$ is a one weight code with weight $\mathtt{m}$, if $uv + uw \neq 0$ then the number of coordinates where b and d have units in different places must be $\frac{\mathtt{m}}{2}$. To obtain this, the number of coordinates where $\{b_j = 1 = d_j\}$ and $\{b_j = 1 + u = d_j\}$ has to be $\frac{\mathtt{m}}{2}$, and in all other coordinates where $\{b_j = 1 \text{ or } 1 + u\}$ we need $\{d_j = 0 \text{ or } u\}$, and also in all other coordinates where $\{b_j = 0 \text{ or } u\}$ we need $\{d_j = 1 \text{ or } 1+u\}$. Hence, consider the vector $v+(1+u)w$. This vector has the same weight as $v + w$ in the first r coordinates but for the last s coordinates, it has $u's$ in the coordinates where $\{b_j = 1 = d_j\}$ and $\{b_j = 1 + u = d_j\}$, so its weight is greater than $\mathtt{m}$. This contradiction gives the result.

iii) Let $x = v+w$ and $y = v+(1+u)w$ be two vectors in $\mathcal{C}$. The binary parts of these two vectors are the same, and for the coordinates over R^s we know from ii) that v and w have units in the same coordinate positions, and for the all other coordinates in R^s, the values of x and y are the same. Therefore, the sum of the weights of the units in v must be same in x and y. So, they also have the same number of coordinates with u. But this is only possible if $|\{j : b_j = d_j = 1 \text{ or } 1 + u\}| = |\{j : b_j = 1, d_j = 1 + u \text{ or } b_j = 1 + u, d_j = 1\}|$. We also know from i) that the number of units in v is $\frac{\mathtt{m}}{2}$, so we have the result.

□

Theorem 5. *Let $\mathcal{C}$ be a one weight code of type $(r, s; k_0, k_1, k_2)$. Then $k_1 \leq 1$ and $\mathcal{C}$ has the following standard form of the generator matrices. If $k_1 = 0$ then*

$$G = \left[\begin{array}{cc|cc} I_{k_0} & A_1 & 0 & uT \\ 0 & 0 & uI_{k_2} & uD \end{array}\right].$$

If $k_1 = 1$ then

$$G = \left[\begin{array}{cc|ccc} I_{k_0} & A_1 & 0 & 0 & uT \\ 0 & s & 1 & a & b_1 + ub_2 \\ 0 & 0 & 0 & uI_{k_2} & uD \end{array}\right]$$

where s, a, b_1, b_2 are vectors over $\mathbb{Z}_2$.

Proof. From Theorem 4 i), we know that any two distinct free vectors have their units in the same coordinate positions. So, if we add the first free row of the generator matrix to the other rows, we have only one free row in the generator matrix. Hence, $k_1 \leq 1$ and considering this and using the standard form of the generator matrix for a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code $\mathcal{C}$ given in Theorem 1, we have the result. $\square$

4. One Weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic Codes

In this section, we study the structure of one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes. At the beginning, we give some fundamental definitions and theorems about $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes. This information about $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes was given in [24], with details.

Definition 5. *An R-submodule $\mathcal{C}$ of $\mathbb{Z}_2^r \times R^s$ is called a $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code if for any codeword $v = (a_0, a_1, \ldots, a_{r-1}, b_0, b_1, \ldots, b_{s-1}) \in \mathcal{C}$, its cyclic shift*

$$T(v) = (a_{r-1}, a_0, \ldots, a_{r-2}, b_{s-1}, b_0, \ldots, b_{s-2})$$

is also in $\mathcal{C}$.

Any codeword $c = (a_0, a_1, \ldots, a_{r-1}, b_0, b_1, \ldots, b_{s-1}) \in \mathbb{Z}_2^r \times R^s$ can be identified with a module element such that

$$\begin{aligned} c(x) &= (a_0 + a_1 x + \ldots + a_{r-1}x^{r-1}, b_0 + b_1 x \\ &+ \ldots + b_{s-1}x^{s-1}) \\ &= (a(x), b(x)) \end{aligned}$$

in $R_{r,s} = \mathbb{Z}_2[x]/\left(x^r - 1\right) \times R[x]/\left(x^s - 1\right)$. This identification gives a one-to-one correspondence between elements in $\mathbb{Z}_2^r \times R^s$ and elements in $R_{r,s}$.

Theorem 6. *[24] Let $\mathcal{C}$ be a $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code in $R_{r,s}$. Then we can identify $\mathcal{C}$ uniquely as $\mathcal{C} = \langle (f(x), 0), (l(x), g(x) + ua(x)) \rangle$, where $f(x) | \left(x^r - 1\right) (\bmod 2)$, and $a(x) | g(x) | \left(x^s - 1\right) (\bmod 2)$, and $l(x)$ is a binary polynomial satisfying $\deg(l(x)) < \deg(f(x))$, $f(x) | \left(\frac{x^s-1}{a(x)}\right) l(x) (\bmod 2)$ and $f(x) \neq \left(\frac{x^s-1}{a(x)}\right) l(x) (\bmod 2)$.*

Considering the theorem above, the type of $\mathcal{C} = \langle (f(x), 0), (l(x), g(x) + ua(x)) \rangle$ can be written in terms of the degrees of the polynomials $f(x)$, $a(x)$ and $g(x)$. Let $t_1 = \deg f(x)$, $t_2 = \deg g(x)$ and $t_3 = \deg a(x)$. Then $\mathcal{C}$ is of type ([24])

$$(r, s; r - t_4, s - t_2, t_2 + t_4 - t_1 - t_3)$$

where $d_1(x) = \gcd\left(f(x), \frac{x^s - 1}{g(x)} l(x)\right)$ and $t_4 = \deg d_1(x)$.

Corollary 7. *If $\mathcal{C}$ is a one weight cyclic code generated by $(l(x), g(x) + ua(x)) \in R_{r,s}$ with weight m then $m = 2s$.*

Proof. We know from Theorem 5 that if $\mathcal{C}$ is a one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code then k_1, which generates the free part of the code, is less than or equal to 1. So, in the case where $\mathcal{C}$ is cyclic, it means that $s - t_2 \leq 1$, where $t_2 = \deg g(x)$. Therefore we have $\deg g(x) = s - 1$ and the polynomial $g(x) + ua(x)$ generates the vector with all unit entries and length s. If we multiply the whole vector (length$= r+s$) by u, then we have a vector with all entries 0 in the first r coordinates and all coordinates u in the last s coordinates. So the weight of this vector is $2s$. Hence the weight of $\mathcal{C}$ must be $2s$. $\square$

Theorem 7. *[24] Let $\mathcal{C} = \langle (f(x), 0), (l(x), g(x) + ua(x)) \rangle$ be a cyclic code in $R_{r,s}$ where $f(x)$, $l(x)$, $g(x)$ and $a(x)$ are as in Theorem 6 and $f(x)h_f(x) = x^r - 1$, $g(x)h_g(x) = x^s - 1$, $g(x) = a(x)b(x)$.*
Let

$$S_1 = \bigcup_{i=0}^{deg(h_f)-1} \left\{x^i * (f(x), 0)\right\},$$

$$S_2 = \bigcup_{i=0}^{deg(h_g)-1} \left\{x^i * (l(x), g(x) + ua(x))\right\}$$

and

$$S_3 = \bigcup_{i=0}^{deg(b)-1} \left\{x^i * (h_g(x)l(x), uh_g(x)a(x))\right\}.$$

Then $S = S_1 \cup S_2 \cup S_3$ forms a minimal spanning set for $\mathcal{C}$ as an R-module.

Let $\mathcal{C} = \langle (f(x), 0), (l(x), g(x) + ua(x)) \rangle$ be a one weight cyclic code in $R_{r,s}$. Consider the codewords $(v, 0) \in \langle (f(x), 0) \rangle$ and $(w_1, w_2) \in \langle (l(x), g(x) + ua(x)) \rangle$. Since $\mathcal{C}$ is a one weight code, $wt(v, 0) = wt(w_1, w_2)$. Further, since $\mathcal{C}$ is an R-submodule, $u(w_1, w_2) = (0, uw_2) \in \mathcal{C}$ and $wt(v, 0) = wt(0, uw_2)$. Moreover, $(v, uw_2) \in \mathcal{C}$ because of the linearity of $\mathcal{C}$. But it is clear that $wt(v, uw_2) \neq wt(v, 0)$ and $wt(v, uw_2) \neq wt(0, uw_2)$. Hence, $\langle (f(x), 0) \rangle$ can not generate a one weight code.

Now, let us suppose that $\mathcal{C} = \langle (l(x), g(x) + ua(x)) \rangle$ is a one weight cyclic code in $R_{r,s}$. We know from Corollary 7 that $\deg g(x) = s-1$, $\mathtt{m} = 2s$ and $g(x)$ generates a vector of length s with all unit entries. Therefore, $l(x)$ also must generate a vector over $\mathbb{Z}_2$ with weight s. Hence, to generate such a cyclic one weight code we have two different cases; $r = s$ and $r > s$.

If $r = s$ then, to generate a vector with weight s, the degree of $l(x)$ must be $s - 1$. So, $(l(x), g(x) + ua(x))$ generates the codeword $(\underbrace{1 \cdots 1}_{\text{length } s} \,|\, \underbrace{unit \cdots unit}_{\text{length } s})$.
Further, if we multiply $(l(x), g(x) + ua(x))$ by $h_g(x)$ we get $(h_g(x)l(x), uh_g(x)a(x))$ and it generates codewords of order 2. Since $r = s$ and the degrees of the polynomials $l(x)$ and $g(x)$ are $s-1$ we have $h_g = x + 1$ and $h_g(x)l(x) = 0$. Hence, $uh_g(x)a(x)$ must generate a vector with weight $2s$, i.e, $h_g(x)a(x)$ must generate a vector of length s with all unit entries. This means that

$$\begin{aligned} h_g(x)a(x) &= \frac{x^s - 1}{(x+1)} \\ \Rightarrow (x+1)a(x) &= \frac{x^s - 1}{(x+1)} \\ a(x) &= \frac{x^s - 1}{(x+1)^2}. \end{aligned}$$

Hence we get $a(x) = \frac{x^s - 1}{(x+1)^2}$. But, since we always assume that s is an odd integer, $a(x)$ is not a factor of $(x^s - 1)$ and this contradicts with the assumption $a(x) | (x^s - 1)$. So, we can not allow $ua(x)h_g(x)$ to generate a vector, i.e, we must always choose $a(x) = g(x)$ to obtain $ua(x)h_g(x) = 0$. So in the case where $\mathcal{C}$ is a one weight cyclic code generated by $(l(x), g(x) + ua(x))$ in $R_{r=s,s}$, we only have $\mathcal{C}$ is a $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code of type $(s, s; 0, 1, 0)$ with weight $\mathtt{m} = 2s$.

In the second case we have $r > s$. We know that $\mathcal{C}$ is a one weight cyclic code with weight $\mathtt{m} = 2s$ and $g(x) = \frac{x^s - 1}{x+1}$ generates a vector with exactly s nonzero and all unit entries. Let $v = (v_1, v_2)$ be a codeword of $\mathcal{C}$ such that $v_1 =< l(x) >$ and $v_2 =< g(x) + ua(x) >$. We can write v as

$$(\underbrace{a_0 a_1 \cdots a_{k-1} a_k}_{s \text{ nonzero entries}} \,|\, \underbrace{unit \cdots unit}_{\text{length } s})$$

where $a_i \in \mathbb{Z}_2, k \in \mathbb{Z}$. Since $\mathcal{C}$ is an R-submodule we can multiply v by u, then we have

$$(\underbrace{00 \cdots 0}_{\text{length } r} \,|\, \underbrace{u \cdots u}_{\text{length } s}).$$

Let $w = (w_1, w_2)$ be another codeword of $\mathcal{C}$ generated by $(h_g(x)l(x), ua(x)h_g(x))$. Since $\mathcal{C}$ is a one weight code of weight $2s$, we can write $w = (\underbrace{b_0 b_1 b_2 \cdots b_{t-1} b_t}_{2s-2p \text{ nonzero entries}} \,|\, \underbrace{u0uu0 \cdots uu0u}_{p \text{ nonzero entries}})$, $b_i \in \mathbb{Z}_2, t \in \mathbb{Z}$. Since $w + uv$ must be a codeword in $\mathcal{C}$, we have

$$w + uv = (\underbrace{b_0 b_1 b_2 \cdots b_{t-1} b_t}_{2s-2p \text{ nonzero entries}} \,|\, \underbrace{0u00u \cdots 00u0}_{s-p \text{ nonzero entries}}).$$

Therefore, $wt(w + uv) = 2s - 2p + 2s - 2p = 4s - 4p$ and since $\mathcal{C}$ is a one weight code with $\mathtt{m} = 2s$,

$$4s - 4p = 2s \Longrightarrow 2s = 4p \Longrightarrow s = 2p.$$

But this contradicts with our assumption, that is, s is an odd integer. Consequently, for $r > s$ and $g(x) \neq 0$ there is no one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code. Under the light of all this discussion, we can give the following proved theorem.

Theorem 8. *Let $\mathcal{C}$ be a $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code in $R_{r=s,s}$ generated by $(l(x), g(x) + ua(x))$ with $\deg l(x) = \deg a(x) = \deg g(x) = s - 1$. Then $\mathcal{C}$ is a one weight cyclic code of type $(r, s; 0, 1, 0)$ with weight $\mathtt{m} = 2s$. Furthermore, there do not exist any other one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code with $g(x) \neq 0$.*

Example 4. *Let $\mathcal{C} = \langle (l(x), g(x) + ua(x)) \rangle$ be a cyclic code in $R_{7,7}$ with $l(x) = g(x) = a(x) = (1 + x + x^3)(1 + x^2 + x^3) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6$. Hence, $\mathcal{C}$ is a one weight code with weight $\mathtt{m} = 14$ and the following generator matrix,*

$$\left(\begin{array}{ccccccc|ccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1+u & 1+u & 1+u & 1+u & 1+u & 1+u & 1+u \end{array} \right).$$

Furthermore, the dual cyclic code $\mathcal{C}^{\perp}$ has the following generator matrix

$$\left(\begin{array}{ccccccc|ccccccc}
0&0&0&0&0&1&1&0&0&0&0&0&0&0\\
1&0&0&0&0&0&1&0&0&0&0&0&0&0\\
1&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&1&1+u\\
0&0&0&0&0&0&1&1+u&0&0&0&0&0&1\\
1&0&0&0&0&0&0&1&1+u&0&0&0&0&0\\
0&1&0&0&0&0&0&0&1&1+u&0&0&0&0\\
0&0&1&0&0&0&0&0&0&1&1+u&0&0&0\\
0&0&0&1&0&0&0&0&0&0&1&1+u&0&0\\
1&1&1&1&1&1&1&u&u&u&u&u&u&u
\end{array}\right).$$

It is obvious from this matrix that $\mathcal{C}^\perp$ is not a one weight code. However, it is a $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code of type $(7,7;7,6,0)$ *and its image under the Gray map is a binary cyclic code with the parameters* $[21,19,2]$.

5. Examples of One Weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic Codes

In this part of the paper, we give some examples of one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes. Furthermore, we look at their binary images under the Gray map that we defined in (1). Actually, according to the results of [22], any binary linear one (constant) weight code with no zero column is equivalent to a λ-fold replication of a simplex code. Hence, the examples of one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes that will be given in this section are all λ-fold replication of simplex code S_k. Therefore, any such code has length $n = \lambda 2^k - 1$, dimension k and weight (or minimum distance) $d = \lambda 2^{k-1}$. It is also well-known that a binary simplex code is cyclic in the usual sense.

If the minimum distance of any code $\mathcal{C}$ get the possible maximum value according to its length and dimension, then $\mathcal{C}$ is called optimal (distance-optimal) or good parameter code. For an example, the binary image of a dual code in Example 4 has the parameters $[21,19,2]$ which are optimal. Let $\mathcal{C}$ be a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code with minimum distance $d = 2t+1$, then we say $\mathcal{C}$ is a t-error correcting code. Since, the Gray map preserves the distances, $\Phi(\mathcal{C})$ is also a t-error correcting code of length $r+2s$ over $\mathbb{Z}_2$. Since, $|\Phi(\mathcal{C})| = |\mathcal{C}|$, we can write a sphere packing bound for a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code $\mathcal{C}$. With the help of usual sphere packing bound in $\mathbb{Z}_2$,

$$|\Phi(\mathcal{C})| \sum_{j=0}^{t} \binom{r+2s}{j} \le |2^{r+2s}|,$$

we have

$$|\mathcal{C}| \sum_{j=0}^{t} \binom{r+2s}{j} \le |2^{r+2s}| = |\mathbb{Z}_2^r \times R^s|.$$

If $\mathcal{C}$ attains the sphere packing bound above then it is called a *perfect code*. Let $\mathcal{C}$ be a $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code of type $(3,2;2,1,0)$ with standard form of the generator matrix

$$G = \left(\begin{array}{ccc|cc} 1&0&1&0&u\\ 0&1&0&0&u\\ \hline 0&0&1&1&1+u \end{array}\right).$$

It is easy to check that $\mathcal{C}$ attains the sphere packing bound, so $\mathcal{C}$ is a perfect code. Moreover, the dual code $\mathcal{C}^\perp$ of $\mathcal{C}$ is generated by the matrix

$$H = \left(\begin{array}{ccc|cc} 1&0&1&u&0\\ \hline 1&1&0&1+u&1 \end{array}\right) \tag{2}$$

and $\mathcal{C}^\perp$ is a one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-linear code with weight $\mathtt{m} = 4$.

Plotkin bound for a code over F_q^n with the minimum distance d is given by,

1. If $d = \left(1 - \frac{1}{q}\right) n$, then $|\mathcal{C}| \le 2qn$.
2. If $d > \left(1 - \frac{1}{q}\right) n$, then $|\mathcal{C}| \le \frac{qd}{qd-(q-1)n}$.

If $\mathcal{C} \subseteq F_q^n$ attains the Plotkin bound then $\mathcal{C}$ is also an equidistant code [25]. Since any one weight binary linear code is a λ-fold replication of a simplex code and have the parameters $[\lambda(2^k-1), k, \lambda(2^{k-1})]$, a binary image of any one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code always meet the Plotkin bound.

Finally, we will give the following examples of one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes. We also determine the parameters of the binary images of these one weight cyclic codes. Further we list some of them in Table 1.

Example 5. *Let $\mathcal{C}$ be a $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code in $R_{15,15}$ generated by $(l(x), g(x)+ua(x))$ where*

$$\begin{aligned}
l(x) &= 1 + x^3 + x^4 + x^6 + x^8 + x^9 + x^{10} + x^{11},\\
g(x) &= x^{15} - 1,\\
a(x) &= 1 + x^3 + x^4 + x^6 + x^8 + x^9 + x^{10} + x^{11}.
\end{aligned}$$

Then $\mathcal{C}$ is a one weight code with weight $\mathtt{m} = 24$ and following generator matrix

$$\left(\begin{array}{ccccccccccccccc|ccccccccccccccc}
1&0&0&1&1&0&1&0&1&1&1&1&0&0&0&u&0&0&u&u&0&u&0&u&u&u&u&0&0&0\\
0&1&0&0&1&1&0&1&0&1&1&1&1&0&0&0&u&0&0&u&u&0&u&0&u&u&u&u&0&0\\
0&0&1&0&0&1&1&0&1&0&1&1&1&1&0&0&0&u&0&0&u&u&0&u&0&u&u&u&u&0\\
0&0&0&1&0&0&1&1&0&1&0&1&1&1&1&0&0&0&u&0&0&u&u&0&u&0&u&u&u&u
\end{array}\right).$$

Furthermore, the binary image $\Phi(\mathcal{C})$ of $\mathcal{C}$ is a $[45, 4, 24]$ code, which is a binary optimal code [26]. Also, it is important to note that $\Phi(C)$ is a 3-fold replication of the simplex code S_4 of length 15.

Example 6. *The $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic code $\mathcal{C} = \langle(l(x), g(x) + ua(x))\rangle$ in $R_{9,9}$ is a one weight code with $\mathtt{m} = 18$, where $l(x) = g(x) = a(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8$. $\mathcal{C}$ has the generator matrix of the form,*

$$\left(\begin{array}{ccccccccc|ccccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \bar{u} & \bar{u} & \bar{u} & \bar{u} & \bar{u} & \bar{u} & \bar{u} & \bar{u} & \bar{u} \end{array}\right)$$

where $\bar{u} = 1 + u$. The Gray image of $\mathcal{C}$ is a 9-fold replication of the simplex code S_2 of length 3 with the optimal parameters $[27, 2, 18]$.

Example 7. *Let $\mathcal{C} = \langle(l(x), g(x) + ua(x))\rangle$, $l(x) = a(x) = 1 + x + x^2 + x^4$, $g(x) = x^7 - 1$, be a cyclic code in $R_{7,7}$. Then the generator matrix of $\mathcal{C}$ is*

$$\left(\begin{array}{ccccccc|ccccccc} 1 & 1 & 1 & 0 & 1 & 0 & 0 & u & u & u & 0 & u & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & u & u & u & 0 & u & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & u & u & u & 0 & u \end{array}\right).$$

$\mathcal{C}$ is a one weight code with $\mathtt{m} = 12$ and $\Phi(\mathcal{C})$ is a 3-fold replication of the simplex code S_3 of length 7 with the parameters $[21, 3, 12]$.

6. Conclusion

In this paper, we study the one weight linear and cyclic codes over $\mathbb{Z}_2^r \times (\mathbb{Z}_2 + u\mathbb{Z}_2)^s$ where $u^2 = 0$. We also classify one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes and present some illustrative examples. We further list some binary linear codes with their parameters which are derived from the Gray images of one weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic codes.

Acknowledgement

The author would like to thank the anonymous reviewers for their careful checking of the paper and the valuable comments and suggestions.

References

[1] Hammons, A. R., Kumar, V., Calderbank, A. R., Sloane, N.J.A. and Solé, P. (1994). The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals, and related codes. *IEEE Trans. Inf. Theory,* 40, 301-319.

[2] Calderbank, A.R. and Sloane, N.J.A. (1995). Modular and p-adic cyclic codes. *Designs, Codes and Cryptography*, 6, 21-35.

[3] Greferath, M. and Schmidt, S. E. (1999). Gray isometries for finite chain rings. *IEEE Trans. Info. Theory*, 45(7), 2522-2524.

[4] Honold, T. and Landjev, I. (1998). *Linear codes over finite chain rings.* In Optimal Codes and Related Topics, 116-126, Sozopol, Bulgaria.

[5] Borges, J., Fernández-Córdoba, C., Pujol, J., Rifà, J. and Villanueva, M. (2010). $\mathbb{Z}_2\mathbb{Z}_4$-linear codes: Generator Matrices and Duality. *Designs, Codes and Cryptography*, 54(2), 167-179.

[6] Aydogdu, I. and Siap, I. (2013). The structure of $\mathbb{Z}_2\mathbb{Z}_{2^s}-$Additive codes: bounds on the minimum distance. *Applied Mathematics and Information Sciences(AMIS)*, 7(6), 2271-2278.

[7] Aydogdu, I. and Siap, I. (2015). On $\mathbb{Z}_{p^r}Z_{p^s}$-additive codes. *Linear and Multilinear Algebra*, , 63(10), 2089-2102.

[8] Abualrub, T., Siap, I. and Aydin, N. (2014). $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes. *IEEE Trans. Inf. Theory*, 60(3), 1508-1514.

[9] Dougherty, S.T., Liu, H. and Yu, L. (2016). One Weight $\mathbb{Z}_2\mathbb{Z}_4$ additive codes. *Applicable Algebra in Engineering, Communication and Computing*, 27, 123-138.

[10] Carlet, C. (2000). *One-weight $\mathbb{Z}_4$-linear codes*, In: Buchmann, J., Høholdt, T., Stichtenoth, H., Tapia-Recillas,H. (eds.) Coding Theory, Cryptography and Related Areas. 57-72. Springer, Berlin.

[11] Wood, J.A.(2002) *The structure of linear codes of constant weight.* Trans. Am. Math. Soc. 354, 1007-1026.

[12] Skachek, V. and Schouhamer Immink, K.A. (2014). *Constant weight codes: An approach based on Knuth's balancing method. IEEE Journal on Selected Areas in Communications*, 32(5), 909-918.

[13] Telatar, I.E. and Gallager, R.G (1990). *Zero error decision feedback capacity of discrete memoryless channels.* in BILCON-90: Proceedings of 1990 Bilkent International Conference on New Trends in Communication, Control and Signal Processing, E. Arikan, Ed. Elsevier, 228-233.

[14] Dyachkov, A.G. (1984). Random constant composition codes for multiple access channels. *Problems Control Inform. Theory/Problemy Upravlen. Teor. Inform.*, 13(6), 357-369.

[15] Ericson, T. and Zinoviev, V. (1995). Spherical codes generated by binary partitions of symmetric point sets. *IEEE Trans. Inform. Theory*, 41(1), 107-129.

[16] King, O.D. (2003). *Bounds for DNA codes with constant GC-content. Electron. J. Combin.*, 10(1), Research Paper 33, (electronic).

[17] Milenkovic, O. and Kashyap, N. (2006). *On the design of codes for DNA computing.* Ser. Lecture Notes in Computer Science, vol. 3969. Berlin: Springer-Verlag, 100-119.

[18] Colbourn, C. J., Kløve, T. and Ling, A. C. H. (2004). Permutation arrays for powerline communication and mutually orthogonal Latin squares. *IEEE Trans. Inform. Theory*, 50(6), 1289-1291.

[19] Abualrub, T. and Siap, I. (2007). Cyclic codes over the rings $\mathbb{Z}_2 + u\mathbb{Z}_2$ and $\mathbb{Z}_2 + u\mathbb{Z}_2 + u^2\mathbb{Z}_2$. *Designs Codes and Cryptography*, 42(3), 273-287.

[20] Al-Ashker, M. and Hamoudeh, M. (2011). Cyclic codes over $\mathbb{Z}_2 + u\mathbb{Z}_2 + u^2\mathbb{Z}_2 + \cdots + u^{k-1}\mathbb{Z}_2$. *Turk. J. Math.*, 35, 37-749.

[21] Dinh, H. Q. (2010). Constacyclic codes of length p^s over $F_{p^m} + uF_{p^m}$. *Journal of Algebra*, 324, 940-950.

[22] Bonisoli, A. (1984). Every equidistant linear code is a sequence of dual Ham- ming codes. *Ars Combin.*, 18, 181-186.

Table 1. Some Examples of One Weight $\mathbb{Z}_2\mathbb{Z}_2[u]$-cyclic Codes

Generators	$\mathbb{Z}_2\mathbb{Z}_2[u]-$type	Binary Image
$l(x) = 1 + x + x^2 + x^4,\ g(x) = x^{21} - 1,\ a(x) = 1 + x + x^2 + x^4 + x^7 + x^8 + x^9 + x^{11} + x^{14} + x^{15} + x^{16} + x^{18}$	$[7, 21; 0; 0, 3]$	$[49, 3, 28]$
$l(x) = a(x) = 1 + x^2 + x^4 + x^5 + x^6 + x^8 + x^9 + x^{13} + x^{14} + x^{15} + x^{16} + x^{17} + x^{20} + x^{21} + x^{23} + x^{26},\ g(x) = x^{31} - 1$	$[31, 31; 0; 0, 5]$	$[93, 5, 48]$
$l(x) = 1 + x + x^3 + x^4 + x^6 + x^7 + x^9 + x^{10} + x^{12} + x^{13} + x^{15} + x^{16} + x^{18} + x^{19} + x^{21} + x^{22} + x^{24} + x^{25}\ , g(x) = x^{15} - 1,\ a(x) = 1 + x + x^3 + x^4 + x^6 + x^7 + x^9 + x^{10} + x^{12} + x^{13}$	$[27, 15; 0; 0, 2]$	$[57, 2, 38]$
$l(x) = 1 + x^2 + x^3 + x^4 + x^7 + x^9 + x^{10} + x^{11} + x^{14} + x^{16} + x^{17} + x^{18} + x^{21} + x^{23} + x^{24} + x^{25} + x^{28} + x^{30} + x^{31} + x^{32},\ g(x) = x^{21} - 1,\ a(x) = 1 + x^2 + x^3 + x^4 + x^7 + x^9 + x^{10} + x^{11} + x^{14} + x^{16} + x^{17} + x^{18}$	$[35, 21; 0, 0, 3]$	$[77, 3, 44]$

[23] Aydogdu, I., Abualrub, T. and Siap, I. (2015). On $\mathbb{Z}_2\mathbb{Z}_2[u]-$additive codes. *International Journal of Computer Mathematics*, 92(9), 1806-1814.

[24] Aydogdu, I., Abualrub, T. and Siap, I. (2017). $\mathbb{Z}_2\mathbb{Z}_2[u]$-Cyclic and constacyclic codes. *IEEE Trans. Inf. Theory*, 63(8), 4883-4893.

[25] Van Lint, J.H. (1992). *Introduction to Coding Theory.* Springer-Verlag, New York.

[26] Grassl, M., *Code tables: Bounds on the parameters of various types of codes.* Online database. Available at http://www.codetables.de/

İsmail Aydoğdu *is a research assistant at the Yildiz Technical University. He received his Ph.D. degree in Mathematics from Yildiz Technical University in 2014. His research interest is mainly in the area of Algebra and Its Applications, especially Algebraic Coding Theory. He has more than 20 presentations in distinguished international conferences and also he has published a couple of research articles in distinguished journals including Finite Fields and Their Applications, and IEEE Transactions on Information Theory.*

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.1, pp.102-116 (2018)
http://doi.org/10.11121/ijocta.01.2018.00464

RESEARCH ARTICLE

On discrete time infinite horizon optimal growth problem

Ayşegül Yıldız Ulus*

Department of Mathematics, Galatasaray University, İstanbul, Turkey
aulus@gsu.edu.tr

ARTICLE INFO

Article History:
Received 30 March 2017
Accepted 10 December 2017
Available 28 December 2017

Keywords:
Optimal growth
Infinite horizon optimal control
Dynamic programming
Lagrange multiplier
Pontryagin's principle

AMS Classification 2010:
49J21, 65K05, 39A22, 91B55

ABSTRACT

Optimal growth problem is an important optimization problem in the theory of economic dynamics. This paper provides an overview of the main approaches used in the existing literature in solving infinite horizon discrete time optimal growth problem and includes very recent developments.

1. Introduction

Dynamic economic theory has been developed via the use of optimal control problems, especially in the context of optimal growth models with infinite planning horizons. On one hand, from the pure economic perspective, optimal growth models serve as one of the best tools in explaining the capital accumulation. On the other hand, from the mathematical viewpoint, optimal growth problem itself can be identified as an interesting dynamic optimization problem. Therefore, while the assumptions of the model describe and shape the economic framework, they also determine the layers of the mathematical difficulty of the problem. Several approaches have been developed in the literature to solve the optimal growth problem. In some of these approaches, one needs to make several strong assumptions in order to address seemingly technically difficult problems. In other approaches, in order to understand the economic implications if a certain assumption fails to hold, one looks for a new mathematical framework.

This paper provides a comprehensive review of four distinct approaches in solving discrete time infinite horizon optimal growth problem: (i) passing to the limit approach; (ii) dynamic programming, (iii) Lagrange multiplier method for infinite horizon and (iv) Pontryagin's approach. It is important to note that these distinct approaches involve different mathematical arguments. In each approach covered in this paper, we attempt to provide the difficulties in obtaining the solution and outline the possible ways to avoid these difficulties. We also provide a comparative discussion about the assumptions of the optimal growth model. Furthermore, we review the different techniques through some relevant examples.

In this paper, we consider an economy that faces a resource allocation problem. The main elements of the given economic model are initial endowment, production function and the preferences. In this economy, we suppose that there are infinite periods and there exists a single household (or consumer) who consumes a single good at each period. A simple production function is assumed where the good is produced from one input, that

*Corresponding Author

is capital. The output is either consumed or saved as capital to the next period. The consumption or saving decision with respect to budget constraint is the only allocation decision that the economy must make. The output is consumed with respect to the preferences of the consumer which are represented by a utility function. The intertemporal utility is defined as the discounted sum of the single period utilities where the discount factor is between 0 and 1 which reflects the property of *additive separability*. We then suppose that the discrete time infinite horizon additively separable optimal growth model involves a benevolent social planner who maximizes the intertemporal utility subject to the constraints of production possibilities and consumption-saving activity.

Based upon the earlier literature, the approach of passing to the limit has been the first one that is utilized in solving the above mentioned problem. It is natural to start with the finite horizon leading to a finite dimensional constrained optimization problem. Here, we should first address the following question: is the limit of the finite horizon problem the unique solution to the infinite horizon problem? To this end, one should note in such a case that we typically face the difficulty in establishing the legitimacy of interchanging the maximum operator and the limit operator. Therefore, for the most of the relevant cases, the answer is negative to the above question.

Dynamic programming has been another important approach that is widely used in solving this type of economic optimization problem. It basically reformulates the actual problem by breaking into sub-decision problems. In doing this, optimum decisions are derived sequentially which leads to a sequence of value functions. This well known method was first studied by R. Bellman in 1957, in [1]. Later, this technique has been applied to dynamic models in economics with a principal reference being Stokey, Lucas and Prescott (1989) ([2]). Dating from Lucas and Stokey (1984) ([3]) and [2], important contributions have been made in the literature to apply dynamic programming techniques to analyze infinite horizon optimal growth problems in different models generating more general results. Le Van and Morhaim (2002) ([4]) provides a unified approach covering bounded and unbounded returns, and Kamihigashi (2014) ([5]) is a resource for a summary of the results in the literature for dealing with unbounded cases, to be a generalization of [2] without making topological assumptions in the additive separable case. For a generalization to models with non-additive and recursive preferences via aggregating function (aggregator) (that include additive separable models), one can refer to [2], for dynamics to [3] and for recent general settings and results dealing with bounded and unbounded returns to Bich et al.(2017) ([6]).

Although, dynamic programming is a very efficient way in order to solve the infinite horizon optimal growth problem, there has been a tendency in the literature to return back to Lagrange multipliers method. However, under such a case, Lagrange multipliers would belong to an infinite dimensional decision space. Thus, the question becomes whether it is possible to derive the sufficient conditions for a Karush-Kuhn-Tucker type theorem to hold in the infinite case. This question has been studied in the literature since Bewley (1972) ([7]) for the general case. Dechert (1982) ([8]) provides an explanation of the structure of the problem in details for the Banach spaces in general. To this end, he uses the functional analysis not only to tackle this problem, but also to demonstrate the sources of the difficulty in switching from a finite problem to an infinite dimensional problem. Multiplier sequence has a nice representation if the space is *reflexive* and the generalization can be done without facing any problem. The question becomes: what if the space is *non-reflexive* such as ℓ^1? In fact, multipliers lie in ℓ^1 in the optimal growth problem.[1] [8] shows that the existence of these multipliers is guaranteed only by the *Axiom of Choice*. There may be no other constructive way to calculate these multipliers. Le Van and Saglam (2004) ([9]) extends the work [8] to the set-up where objective and constraint functions do not need to be real-valued in order to cover the cases where Inada type conditions are assumed. [9] also discusses some other interesting applications of this method in economics.

In the classical optimal growth model, when written as an equivalent minimization program, the objective and the constraint functions are scalar valued and supposed to be convex. There have been some works in the existing literature relaxing the assumption of convexity in order to obtain results in non-convex cases. As an example, Rustichini (1998) ([10]) studied the general optimization problem using non-convex models. The questions of whether the separating vectors do exist and they can be represented by a sequence of real numbers in infinite dimensional spaces have

[1] Here, we denote by ℓ^1 the space of real sequences $a = (a_t)_t$ such that $\sum_{t=0}^{\infty} |a_t|$ is convergent in $\mathbb{R}$. Note that endowed with the norm $||a||_1 = \sum_{t=0}^{\infty} |a_t|$, ℓ^1 is Banach but not reflexive since $(\ell^1)' = \ell^\infty$ but $(\ell^\infty)' \neq \ell^1$. Here, we denote by ℓ^∞ the space of bounded sequences $a = (a_t)$ such that $\sup_t |a_t| \leq \infty$.

been addressed in [10]. Moreover, vector optimization problems on Banach spaces without convexity assumptions have also been considered in Dutta and Tammer (2006) ([11]). Here, it is important to note that an additional assumption stating that the objective function is locally Lipschitz was necessary in [11]. However, in this context, using the approach of Pontryagin's principle, Blot and Chebbi (2000) ([12]), Blot and Hayek (2008) ([13]), Blot and Hayek (2014) ([14]) and Blot et al. (2015) ([?]) give useful results without restrictive assumptions. [13], [14] and [?] consider dynamic systems which are governed by difference equations and difference inequations respectively. In all of these cited works, a vector valued problem is considered, that is, the states and the controls are vector valued. Moreover, these works use weaker convexity assumptions than the usual ones to obtain strong Pontryagin's principles and they provide weak Pontryagin's principles without convexity conditions. The solution approach in [12] is based on using reductions to finite horizon problems. However in [13], [14] and [?], the authors consider the problem in the space of bounded sequences, which allows them to use functional analyctic approach which is based on the use of abstract results of the optimization theory and optimal control problems in ordered Banach spaces. In the spirit of the Karush-Kuhn-Tucker theorem, they establish the necessary and sufficient conditions in the form of weak Pontryagin's principles. This result can be used for different kinds of optimal control problems that can be found in economics, optimal management of renewable resources, sustainable development theory and game theory.

In this paper, besides studying the two classical approaches (passing to limit approach and dynamic programming) including very recent extensions and developments, we aim to apply the most recent two functional analytic approaches for solving optimal growth problem: Lagrange multiplier method for infinite horizon and the approach of weak Pontryagin's principle. Lagrange multiplier method for infinite horizon is due to [8] and based on extending Lagrange multiplier method to infinite dimensional space. In some sense, this approach can be seen as an extension of the passing to the limit approach and also serves as an alternative method to the dynamic programming. We give sufficient conditions on the objective and constraint functions under which the Lagrange multiplier can be represented by a ℓ^1 sequence. We assume Inada conditions as in [9]. In economics, Lagrange multiplier method has been the key tool for obtaining the solution in optimization problems in economics and Lagrange multipliers provide meaningful insights in the economic models. Therefore, it is useful not only for providing solution to the problem but also analyzing the nature of the solution. The idea of the approach of weak Pontryagin's principle is to transform the optimal control problem into a dynamical system. A solution to the discrete time optimal growth problem is given as a special case of the results obtained in [14]. The result is useful as the assumptions are easy to check. To compare these two functional analytic approaches, we have to note that in Lagrange multiplier method we need the concavity assumptions of one period utility and the production functions but one can avoid convexity conditions to obtain weak Pontryagin's principle. Furthermore, in the approach of Pontryagin's principle vector states and vector controls are used hence encompasses in this sense also the Lagrange multiplier method.

The rest of the paper is organized as follows. In Section 2, we describe the set-up of the optimal growth problem. Section 3 gives the mathematical background of the classical approaches for the optimal growth problem together with the recent developments. Then, in Section 4, the functional analytic approach is studied. Section 5 concludes.

2. One-sector optimal growth model: set-up

This section presents the set-up of deterministic discrete time infinite horizon one-sector optimal growth model. We consider an economy as a problem of resource allocation. The primitives of the model are initial endowment, production function and the preferences.

We consider an economy $\mathcal{E}$ of infinite periods from time $t = 0$ to $t = \infty$. We suppose that there is 1 unit of time each period. There is a single household[2] who consumes a single good at each period. This good (output) is produced from one input, capital. At time $t = 0$, the amount of capital is supposed to be k_0 units. The output is produced from capital by a production function $f : \mathbb{R}_+ \to \mathbb{R}_+$.

In each period t, we suppose the single good (output) with quantity $y_t \in \mathbb{R}_+$ which is produced from one input: k_t by a production function f where $y_t = f(k_t)$. The output is either consumed as $c_t \geq 0$ or saved as capital to the next period

[2]The model assumes a *representative* household. It is justified if for example all the households are supposed to be identical in the economy $\mathcal{E}$.

as k_{t+1} satisfying the following process being repeated until infinity:

$$c_t + k_{t+1} \leq y_t = f(k_t) \text{ with } k_t \geq 0$$

The consumption level is determined according to the unique consumer's preferences which is defined by a one period non-decreasing utility (reward) function $u : \mathbb{R}_+ \to \mathbb{R}$. The intertemporal utility is then defined as follows:

$$\sum_{t=0}^{\infty} \beta^t u(c_t)$$

where $0 < \beta < 1$ is the discount factor.

2.1. Social planner's problem

We first give some definitions in order to describe the problem.

Definition 1. *For any $k_0 > 0$, when $k = (k_1, k_2, \ldots k_t, \ldots)$ is such that $0 \leq k_{t+1} \leq f(k_t)$ for all t, we say that it is a feasible accumulation path from k_0.*

Definition 2. *A consumption sequence $c = (c_0, c_1, c_2, \ldots c_t, \ldots)$ is feasible from $k_0 > 0$ if there exists a sequence $k \in \Pi(k_0)$ that satisfies $0 \leq c_t \leq f(k_t) - k_{t+1}$ for all t.*

Definition 3. *The set of feasible allocation from k_0 is denoted by $\Pi(k_0)$. That is, $\Pi(k_0) := \{(k, c) = (k_t, c_t)_{t=0}^{\infty} : c_t + k_{t+1} - f(k_t) \leq 0 \text{ for all } t = 0 \ldots\}$.*

The objective of a benevolent social planner is to maximize the utility of household by choosing the feasible allocation (k, c), that is, subject to the feasibility constraints with a given positive initial capital. The problem can be written as follows:

$$(P) \begin{cases} \max \sum_{t=0}^{\infty} \beta^t u(c_t) \\ \text{s.t.} \\ c_t + k_{t+1} - f(k_t) \leq 0, \forall t \geq 0 \\ c_t \geq 0, \forall t \geq 0 \\ k_t \geq 0, \forall t \geq 1 \\ k_0 > 0, \text{given} \end{cases}$$

The objective function states that social planner must only decide the consumption level at each period in order to maximize the utility. The constraints reflect that non-consumed, i.e., saved amount of output will be added to the capital of the next period and hence will determine the future production levels. Furthermore, since the temporal utility function u_t is non-decreasing, at optimum, output will not be wasted so that the consumption at t will be equal to the difference of output and quantity saved, that is $c_t = f(k_t) - k_{t+1}$. Eliminating c_t from the problem (P) gives us a new formulation $(\widetilde{P})$:

$$(\widetilde{P}) \begin{cases} \max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u[f(k_t) - k_{t+1}] \\ \text{s.t.} \\ 0 \leq k_{t+1} \leq f(k_t), \forall t \geq 0 \\ k_0 > 0, \text{given} \end{cases}$$

3. Classical approaches of solution

In this section, we discuss the two classical approaches of solution to the optimal growth problem, namely passing to the limit and dynamic programming. We give their mathematical arguments with respect to the assumptions of the model and provide some examples.

3.1. Assumptions

In the following, we give a list of assumptions of the model. In Section 3.2, the entire list will prove to be useful, in Section 3.3, one may assume the weaker versions of the ones in this list.

(EA) (Endowment Assumption) $k_0 > 0$, given.

(Prod) (Production Assumption)

1. *f is stricly concave in $\mathbb{R}_+$,*
2. *f is continuously differentiable in $\mathbb{R}_+$,*
3. *f is strictly increasing,*
4. *$f(0) = 0$, $\lim_{k\to 0} f'(k) = +\infty$ and $f'(\infty) < 1$ (Inada conditions).*

(Pref) (Preferences Assumption)

1. *u is bounded,*
2. *u is stricly concave in $\mathbb{R}_+$,*
3. *u is continuously differentiable in $\mathbb{R}_+$,*
4. *u is strictly increasing,*
5. *$\lim_{c\to 0} u'(c) = +\infty$ and $\lim_{c\to+\infty} u'(c) = 0$ (Inada conditions)*
6. *$u(c_0, c_1, \ldots) = \sum_{t=0}^{\infty} \beta^t u(c_t)$ where $0 < \beta < 1$.*

Remark 1. (1) *Assumption (EA) is quite standard. We assume that, at the beginning, we have some positive capital.*
(2) *By the assumption Prod (1-3), we suppose that the production function is strictly concave, continuously differentiable in $\mathbb{R}_+$ and strictly increasing. These assumptions can be weakened to the degree that one can overcome the mathematical difficulty. In assumption Prod (4), we assume that the production function satisfies the asymptotic conditions, called also*

Inada conditions, which will guarantee the existence of interior solutions for the optimization problem. Since $\lim_{k\to 0} f'(k) = +\infty$ *and* $f'(\infty) < 1$*, there is a maximum feasible level of capital, which we can call* k_{max}*. This condition is satisfied when for example the production function is of so-called Cobb-Douglas type, i.e.* $f(k) = Ak^{\alpha}$ *with* A *constant in* $\mathbb{R}_+$ *and* $0 < \alpha < 1$*.*

(3) *By assumption Pref(1), we assume that the one period utility function is bounded. Pref (2-4) are the analogous versions of Prod (2-4). According to Pref (5)-Inada conditions, the marginal utility of consumption for a starving agent would be so high and the marginal utility for a satiated consumer would be so low. We assume by Pref (6) that the preferences over intertemporal consumption sequence take the additively separable form.*

3.2. Passing to the limit

We are interested in the infinite horizon case. Nevertheless, it was logic to start with the finite horizon. The approach of passing to the limit has naturally been the first one for solving this problem. The problem was then a finite dimensional constrained optimization problem. In economics, the method of Lagrange has been widely applied for solving finite dimensional constrained optimization problems. That is, under the assumption of the model that we cited in Section 3.1, namely *(EA)*, *Prod(1-4)* and *Pref(1-6)*, there exist Lagrange multipliers so that the solution to the constrained maximization problem is also an extreme value of the objective function of the social planner without constraints.

The set of sequences $\{k_{t+1}\}_{t=0}^{T}$ satisfying the constraints of the problem is a closed, bounded[3] and convex subset of $\mathbb{R}^{T+1}$ and the objective function is continuous (as the sum of the continuous functions) and strictly concave by the assumptions *Pref(2)* and *Pref(3)*. Hence, there is exactly one solution which is characterized by Karush-Kuhn-Tucker conditions. By the Assumptions $f(0) = 0$ in *Prod(4)* and $u'(0) = \infty$ in *Pref(5)*, it is clear that the constraints do not bind except for $k_{T+1} = 0$. Thus, the solution satisfies the first order and the boundary conditions for all $t = 1, \ldots T$:

$$\beta f'(k_t)u'[f(k_t) - k_{t+1}] = u'[f(k_{t-1}) - k_t] \quad (1)$$

$$k_{T+1} = 0, k_0 \text{ given.} \quad (2)$$

These conditions give us a *2nd order difference equation* in k_t which has a 2-parameter family of solutions but the one which satisfies the boundary conditions is the unique solution.

Here, the question turns out to be whether the limit of the finite horizon problem is the unique solution to the infinite horizon problem. The answer is positive for some parametric examples in economics, for instance, as in the following example. However, this method involves in general one difficulty that to establish the legitimacy of interchanging the operators max with $\lim_{T\to\infty}$, to guarantee that

$$\max \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t u(c_t) = \lim_{T\to\infty} \max \sum_{t=0}^{T} \beta^t u(c_t)$$

This difficulty is overcome if the uniform convergence of the solution path is satisfied. However, this will bring restrictive assumptions on the model. Instead, different approaches are developed by which not only the problem is solved but also with weaker assumptions of the model.

Example 1. *Consider a logarithmic utility function* $u(c_t) = \ln c_t$ *and Cobb-Douglas production function:* $f(k_t) = (k_t)^{\alpha}$ *with* $0 < \alpha < 1$*. Thus, the optimal growth problem will be:*

$$(\widetilde{P}) \begin{cases} \max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln[(k_t)^{\alpha} - k_{t+1}] \\ s.t. \\ 0 \le k_{t+1} \le (k_t)^{\alpha}, \forall t \ge 0 \\ k_0 > 0, given \end{cases}$$

By the help of the above equations (1) and (2), one can check that the unique solution to the corresponding problem $(\widetilde{P})$ *is:*

$$k_{t+1} = \alpha\beta \frac{1 - (\alpha\beta)^{T-t}}{1 - (\alpha\beta)^{T-t+1}} k_t^{\alpha} \text{ for all } t = 0, 1 \ldots T$$

Passing to the limit, we find that $k_{t+1} = \alpha\beta k_t^{\alpha}$ *is the unique solution for the infinite horizon problem.*

Remark 2. *Note that the assumption of boundedness of the utility function is not satisfied in the previous example. Boundedness is needed in order to guarantee the existence of optimal solution though a solution can exist without it as in the previous example.*

[3]Showing that it is closed is straightforward as $c_t \in [0, f(k_t)]$. To show that it is bounded, we note that by the assumption, *Prod(4)*, there is a maximum feasible level of capital k_{max}.

3.3. Dynamic programming

Dynamic programming has been another useful approach for solving the optimal growth problem. [2] is the principal reference for the use of this method in optimal growth problem. In this section, after giving the idea of the approach and the Principle of Optimality, we will give an overview of the results in the literature according to the assumptions of the model. The first three subsections deal with the additively separable optimal growth problem. Section 3.3.4 discusses non-additive model.

The idea of dynamic programming is to divide the problem up into separate sub-problems. The first step is to define and solve the problem of the initial period and then to proceed forward.

The problem at the initial period that the social planner faces is to choose current period's consumption c_0 and capital to begin with for the next period k_1. If we knew the preferences of planner over (k_1, c_0), we could simply maximize the *appropriate* function of (k_1, c_0) over the opportunity set defined by the constraint:

$$c_t + k_{t+1} - f(k_t) \leq 0, \text{ for all } t \geq 0.$$

Suppose that the above problem is solved for all possible values of k_0. Then, we could define a function $v : \mathbb{R}_+ \to \mathbb{R}$ by taking $v(k_0)$ to be the value of the maximized objective function, for each $k_0 \geq 0$:

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u[f(k_t) - k_{t+1}] \qquad (3)$$

such that

$$0 \leq k_{t+1} \leq f(k_t), \forall t \geq 0$$

$$k_0 > 0, \text{given}$$

A function of this type is called a *value function.*

With v so defined, $v(k_1)$ would give the utility from period 1 and that could be obtained with k_1. $\beta v(k_1)$ would be then the value of this utility discounted back to period 0.

In terms of this value function v, the planner's problem in period $t = 0$ would be the following optimal growth program:

$$\begin{cases} \max_{k_1, c_0}[u(c_0) + \beta v(k_1)] \\ \text{ s.t.} \\ c_0 + k_1 \leq f(k_0), \\ c_0, k_1 \geq 0, k_0 > 0 \text{ given.} \end{cases}$$

v is unknown at this point. Thus, solving the above program provides also v. That is, v must satisfy:

$$v(k_0) = \max_{0 \leq k_1 \leq f(k_0)} \{u[f(k_0) - k_1] + \beta v(k_1)\}$$

Irrespective of the date, we can rewrite the problem of planner with current capital stock denoted by $z, y \in \mathbb{R}_+$ as a *functional equation* (equation in the unknown function of v):

$$v(z) = \max_{0 \leq y \leq f(z)} \{u[f(z) - y] + \beta v(y)\} \qquad (4)$$

The study of *dynamic optimization* problems through the analysis of such functional equation is called *dynamic programming.*

We can view the above equation (4) (called also Bellman equation) through a functional operator (*Bellman operator*) :

$$(Tw)(z) = \max_{0 \leq y \leq f(z)} \{u[f(k) - y] + \beta w(y)\}$$

solutions of (4) being fixed points of T.

The idea is then to study the link between the value function of the optimal growth program with the solutions of Bellman equation. That is, to study the link between the value function of the optimal growth program with the fixed points of the Bellman operator. Thus, one has to verify the following issues:

(i) (*Existence*) Existence of a fixed point of Bellman operator is obtained as the value function of the optimal growth program is a fixed point of T. $v(z)$ (which is the unknown of the Bellman equation) satisfies $(Tv)(z) = v(z)$. Existence is guaranteed by some sufficient conditions via a *Banach-type Fixed Point Theorem* and *Berge's Maximum Theorem.*

(ii) (*Uniqueness*) Studying a fixed point of T allows us to reach the value function of the optimal growth program, if uniqueness of such a fixed point is obtained, then the (unique) fixed point is the value function.

(iii) (*Reachability*) Bellman operator gives an algorithm to reach (under appropriate conditions) the value function of the optimal growth program. As in some problems the suitable starting points to reach the value function must be restricted. By means of iterating on the Bellman operator will provide the convergence to this value function from any "initial suitable feasible guess".

The following theorem gives the details of the Bellman's Principle of Optimality whose idea is given above and show that the dynamic programming technique allows to recover the value function of the optimal growth problem.

Theorem 1. *(Principle of Optimality) The solution v to the Bellman equation (4) evaluated at $z = k_0$ gives the maximum in optimal growth program (3) when the initial state is k_0. Moreover a sequence $\{k_{t+1}\}_{t=0}^{\infty}$ attains the maximum in (3) if and only if it satisfies for all $t \geq 0$:*

$$v(k_t) = u[f(k_t) - k_{t+1}] + \beta v(k_{t+1}) \tag{5}$$

The Principle of Optimality is verified under a series of topological assumptions for the bounded case as well as for two important particular cases: with bounded returns and with unbounded returns (see Chapter 4 of [2]). The following sections give the versions of these results for our setting.

3.3.1. Optimal growth with bounded utility

In this section, we consider the optimal growth problem under the assumptions of the model given in Section 3.1.

Theorem 2. *Under the assumptions (EA), Prod$(1-4)$ and Pref$(1-6)$,*

1. *solutions to the functional equation (3) and sequence plans (4) coincide exactly,*
2. *the Bellman operator has a unique fixed point in the space of bounded contionous functions and this fixed point is the value function v,*
3. *value iteration converges uniformly to the value function starting from any bounded continuous function.*

Proof. *(3)* As u is supposed to be bounded by the assumption *Pref(1)* and $0 < \beta < 1$ by *Pref(6)*, then $\Pi(k_0) \neq \emptyset$ and $\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t u[f(k_t) - k_{t+1}]$ exists for all $k_0 \in \mathbb{R}_+$. The maximum function v^* is then bounded and satisfies:

$$v^*(k_0) = \max_{k_0 \in \Pi(k_0)} \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t u[f(k_t) - k_{t+1}]$$

Thus, $v^*(k_0)$ is the maximum in (3). It is natural to seek the solutions to (4) in bounded continuous functions. Any bounded continuous solution to (4) satisfies $\lim_{n\to\infty} \beta^n v(k_n) = 0$ then $v = v^*$. Moreover, given a solution to (4), for any k_0, a sequence $\{k_t^*\}$ attains the maximum in (3) if and only if it is generated by the following mechanism where $0 \leq k_{t+1} \leq f(k_t)$:

$$v(k_t) = u(f(k_t) - k_{t+1}) + \beta v(k_t)$$

(2) If we define,

$$(Tv)(z) = \max_{0 \leq y \leq f(z)} \{u[f(z) - y] + \beta v(y)\},$$

instead of (4), we can write $v = Tv$. As the feasibility condition is given as a closed interval $[0, f(z)]$ together with the convexity of $\mathbb{R}_+$ and the boundedness and the continuity assumptions given in *Prod(1)* and *Pref(1-3)*, T has a unique fixed point in the space of bounded continuous functions. This fixed point is the value function v^*.

(3) By the assumptions *Prod(1-3)*, *Pref(2-4)*, v is stricly increasing, strictly concave and continuously differentiable. If $\{v_n\}$ is a sequence of approximations defined by $v_n = T^n v_0$ with an appropriate choice of bounded contionous starting function v_0, then this value iteration converges uniformly to the value function v^*.

□

3.3.2. Optimal growth with bounded returns

This section deals with the optimal growth problem with bounded returns under the following list of assumptions which is slightly weaker than the list given in Section 3.1:

(EA) (Endowment Assumption) $k_0 > 0$, given.

$(\widetilde{Prod})$ (Production Assumption)

1. *f is continuous,*
2. *f is concave in $\mathbb{R}_+$,*
3. *f is continuously differentiable in $\mathbb{R}_+$,*
4. *f is strictly increasing,*
5. *$f(0) = 0$, for some $\bar{k} > 0$: for all $0 \leq k \leq \bar{k}$ we have $k \leq f(k) \leq \bar{k}$ and for $k > \bar{k}$ we have $f(k) < \bar{k}$.*

$(\widetilde{Pref})$ (Preferences Assumption)

1. *u is continuous,*
2. *u is stricly concave in $\mathbb{R}_+$,*
3. *u is continuously differentiable in $\mathbb{R}_+$,*
4. *u is strictly increasing,*
5. *$\lim_{c\to 0} u'(c) = +\infty$ and $\lim_{c\to+\infty} u'(c) = 0$ (Inada conditions)*
6. *$u(c_0, c_1, \ldots) = \sum_{t=0}^{\infty} \beta^t u(c_t)$ where $0 < \beta < 1$.*

Remark 3. *We have to mention especially that u is not supposed to be bounded. However, note that under the assumptions $\widetilde{Pref}(1-3)$ and $\widetilde{Prod}(1-3)$ the function G which is defined as $G(k_t, k_{t+1}) := u[f(k_t) - k_{t+1}]$ is bounded. Thus, the case is called optimal growth with bounded returns.*

Theorem 3. *Under the assumptions (EA), $\widetilde{Prod}(1-4)$ and $\widetilde{Pref}(1-6)$,*

1. *solutions to the functional equation (3) and sequence plans (4) coincide exactly,*
2. *the Bellman operator has a unique fixed point in the space of bounded contionous functions and this fixed point is the value function v,*
3. *value iteration converges uniformly to the value function starting from any bounded continuous function.*

Proof. (1) By Remark 3, G is bounded. Thus, if B is a bound for $G(z,y)$, then the maximum function v^* satisfies $|v^*(z)| \leq \frac{B}{1-\beta}$ as

$$v^*(k_0) = \max_{k_0 \in \Pi(k_0)} \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t u[f(k_t) - k_{t+1}]$$

Thus, $v^*(k_0)$ is the maximum in (3). Any bounded continuous solution to (4) satisfies $\lim_{n\to\infty} \beta^n v(k_n) = 0$ then $v = v^*$. Moreover, given a solution to (4), for any k_0, a sequence $\{k_t^*\}$ attains the maximum in (3) if and only if it is generated by the following mechanism where $0 \leq k_{t+1} \leq f(k_t)$:

$$v(k_t) = u(f(k_t) - k_{t+1}) + \beta v(k_t)$$

(2) and (3) here are essentially analogous versions of (2) and (3) in Theorem 2. It suffices to remark that v is stricly increasing, strictly concave and continuously differentiable as $G(.,y)$ is so by means of the assumptions $\widetilde{Prod}(2-4)$ and $\widetilde{Pref}(2-4)$.

□

3.3.3. Optimal growth with unbounded returns

In economics, the utility function are often unbounded from above and/or below. In [2], this case is partly considered and called *optimal growth with unbounded returns.* That is, it is the case where the maximum function v^* satisfies the Bellman equation (4) but the following boundedness assumption is not satisfied:

If $\lim_{n\to\infty} \beta^n v(k_n) = 0$ for all $(k_0, k_1, \dots) \in \Pi(k_0)$ then $v = v^*$.

In this case, the problem is that the functional equation (4) would give many solutions. The sufficient conditions for a solution to equation (4) to be the maximum function v^* are given in Theorem 4.14 in [2]. The idea is to guess a solution to the equation (4) and start with an appropriate function $\hat{v}$ that is an upper bound for v^* and then iterarate down to the fixed point of T. We will discuss these sufficient conditions in the following two examples which are used quite often in economics. These examples will prove to be useful for our comparative study. Nevertheless, in the literature, there has been an extensive research in order to give a general setting for dealing with the unbounded case. One can refer to Le Van and Morhaim (2002) ([4]) which provides a unified approach covering bounded and unbounded utilities. The recent reference Kamihigashi (2014) ([5]) is intended to be a resource for a summary of the results in the literature for dealing with such unbounded cases, to be a generalization of [2] without making topological assumptions. Unlike the former ones, in [5], instead of a Banach-type Fixed Point Theorem, Knaster-Tarski Fixed Point Theorem is used to show the existence of a fixed point of the Bellman operator.

Example 2. *We consider the same problem of Example 1 and we solve it by dynamic programming. One can overcome the difficulty due to the unboundedness of the utility by choosing a specific functional form as an upper bound.*
The problem corresponding to(4) is then:

$$v(z) = \max_{0 \leq y \leq (z^\alpha)} \{\ln[(z^\alpha) - y] + \beta v(y)\}$$

The sufficient condition of having a unique solution is to find a bound function $\hat{v}(z)$ for the maximum function v^*:

$$v^*(z) \leq \frac{\alpha \ln k}{(1-\alpha\beta)}, \forall z > 0$$

We may take $\hat{v}(z) = \frac{\alpha \ln z}{(1-\alpha\beta)}$
With T defined as follows:

$$(Tw)(z) = \max_{0 \leq y \leq z^\alpha} \{\ln[f(z) - y] + \beta w(y)\}$$

By some calculations, one can show that the following $v(z)$ is the fixed point of T:

$$v(z) = \frac{1}{\beta}[\ln(1-\alpha\beta) + \frac{\alpha\beta}{1-\alpha\beta}\ln(\alpha\beta)] + \frac{\alpha\beta}{1-\alpha\beta}\ln z$$

so that the optimal sequence is generated as follows:

$$k_{t+1} = \alpha\beta k_t^\alpha \textit{ for all } t = 0, 1 \dots$$

Example 3. *(Cake Eating Problem) In this example, suppose that one consumer has a cake of a given initial size of k_0. In each period, the consumer eats some part of the cake with respect to its preferences and save the remainder satisfying*

$k_{t+1} = k_t - c_t$ for all $t = 0, 1 \ldots$. Suppose that the consumer's preferences are represented by the utility function $u(c_t) = \ln c_t$. Hence, finding the optimal path of consumption of the cake can be interpreted as solving the following optimal growth problem with linear production function $f(z) = z$ for all $z \in \mathbb{R}_+$:

$$\widetilde{(P)} \begin{cases} \max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln[k_t - k_{t+1}] \\ s.t. \\ 0 \leq k_{t+1} \leq k_t, \forall t \geq 0 \\ k_0 > 0, given \end{cases}$$

We can solve $\widetilde{(P)}$ by dynamic programming. We choose here again a specific functional form as an upper bound. We proceed as follows:

Since $\ln k_t \leq \ln k_0$ for all $t = 1, 2 \ldots$, we will have:

$$\ln[k_t - k_{t+1}] \leq \ln k_t \leq \ln k_0 \text{ and}$$

$$\sum_{t=0}^{\infty} \beta^t \ln[k_t - k_{t+1}] \leq \frac{1}{1-\beta} \ln k_0$$

Hence $v^*(z) \leq \frac{1}{1-\beta} \ln k_0$ where v^* is the supremum function. Define $\hat{v}(z) = \frac{1}{1-\beta} \ln k_0$.

With T defined by:

$$(Tw_n)(z) = \max_{0 \leq y \leq z} \{\ln[z - y] + \beta w_n(y)\}$$

one has:

$$T\hat{v}(z) = \max_{0 \leq y \leq z} \{\ln[z - y] + \frac{\beta}{1-\beta} \ln y\}$$

The first order conditions of the right hand side of the above equation gives us $y = \beta z$ and therefore we have:

$$T\hat{v}(z) = \frac{1}{1-\beta} \ln z + \ln(1-\beta) + [\frac{\beta}{1-\beta} \ln \beta]$$

By the iteration, we will have:

$$T^n\hat{v}(z) = \frac{1}{1-\beta} \ln z + [\ln(1-\beta) - \frac{\beta}{1-\beta} \ln \beta] \sum_{j=0}^{n} \beta^j$$

Defining $v(z) = \lim_n T^n \hat{v}(z)$, and taking the limit of above equation will give us the fixed point of T, that is:

$$v(z) = \frac{1}{1-\beta} \ln z + \frac{1}{1-\beta} [\ln(1-\beta) - \frac{\beta}{1-\beta} \ln \beta]$$

Since $Tv(z) = v(z) = \max_{0 \leq y \leq z} \{\ln[z - y] + \beta v(y)\}$, first order condition of the right hand side of this equation gives us the following optimal sequence:

$$k_{t+1} = \beta k_t \text{ for all } t = 0, 1 \ldots$$

3.3.4. Non-additive optimal growth problem

In this paper, we have so far considered an additively separable model which is in fact satisfied by the assumptions $Pref(6)$ and $\widetilde{Pref}(6)$. In this section we will consider the non-additive model via recursive preferences and aggregating functions which are due to Lucas and Stokey (1984) ([3]).

Definition 4. *The utility function u is recursive if $u(c) = u(c_0, c_1, \ldots, c_n, \ldots)$ is a function $A(c_0, u(c_1, \ldots, c_n, \ldots))$ of today's consumption c_0 and the intertemporal utility from tomorrow. The function A aggregates the today's consumption c_0 and future utility into the current utility and is called an aggregating function (aggregator).*

Definition 5. *The aggregating function $A : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ has the following properties:*

(1) *(AI) A is continuous,*
(2) *(AII) $A(0, 0) = 0$,*
(3) *(AIII) For any $z \in \mathbb{R}_+$, $A(., z)$ is bounded,*
(4) *(AIV) $|A(x, z) - A(x, z')| \leq \beta|z - z'|$ for $x, z, z' \in \mathbb{R}_+$ and $0 < \beta < 1$,*

The class of utility functions that are considered are then defined by $u_A(c) = A[c_0, u(c_1, c_2 \ldots)]$. The following theorem describes the source and the properties of this class according to the aggregating function. In such a model, dynamic programming approach can be applied with recursive preferences that have a contraction property.

Theorem 4. *Let S be the vector space of all bounded (with the norm $||u||_\infty = \sup_{c \in \ell_+^\infty} |u(c)|$) and continuous functions such that $u : \ell_+^\infty \to \mathbb{R}$. Let A satisfy AI, AII, $AIII$ and AIV and let T_A be an operator defined as $T_A : S \to S$ and $(T_A u)(c)) = A[c_0, u(c_1, c_2 \ldots)]$ where $c = (c_0, c_1, c_2 \ldots) \in \ell_+^\infty$. Then, T_A has a unique fixed point u_A in S. Moreover, if A is increasing and concave then u_A is increasing and concave.*

Proof. By the definition of T_A and by the property (AIV), T_A is a contraction. Hence existence of a unique fixed point holds by Banach Fixed Point Theorem as S is complete. Moreover, A is increasing as T_A takes increasing functions to increasing functions. T_A is a contraction then $T_A u$ is concave if $u \in S$ is concave. Thus, the unique fixed point u_A is concave. □

Remark 4. *The above setting encompasses the additive seperable model. That is, additive case is a special case if we consider the aggregating function* $A[c_0, u(c_1, c_2 \ldots)] = u(c_0) + \beta u(c_1, c_2, \ldots)$ *where* $0 < \beta < 1$.

Corollary 1. *(Additive Recursive Preferences) Let* T_A *be an operator defined by* $T : S \to S$ *and* $(Tu)(c) = A[c_0, u(c_1, c_2 \ldots)] = u(c_0) + \beta u(c_1, c_2, \ldots)$ *where* $c = (c_0, c_1, c_2 \ldots) \in \ell^\infty_+$. *Then,*

(1) *Any function* $u : \ell^\infty_+ \to \mathbb{R}$ *satisfying* $u(c) = \sum_{t=0}^\infty \beta^t u(c_t)$ *where* $0 < \beta < 1$ *is bounded and continuous if* $u : \mathbb{R}_+ \to \mathbb{R}$ *is bounded and continuous.*

(2) *The function* $u \in S$ *defined by* $u(c) = \sum_{t=0}^\infty \beta^t u(c_t)$ *is the unique fixed point of* T_A.

4. Functional analytic approach

In this section, we give two different functional analytic approaches to solve our particular problem defined in Section 2. In Section 4.1, we apply the main result of [8] tracking the lines of [9] which is indeed the Lagrange multiplier method for optimal growth. In Section 4.3 we apply the approach of weak Pontryagin's principle due to [14] to our problem. We then compare these results with respect to the assumptions of the model.

4.1. Lagrange multiplier method for infinite dimensional space

The aim of this section is to set the optimal growth problem (P) given in Section 2 as a minimization problem $(\tilde{\tilde{P}})$ and showing that all the conditions of the Main Theorem in [8] are fulfilled for the optimal growth problem $(\tilde{\tilde{P}})$.

Set $x = (k, c) \in \ell^\infty \times \ell^\infty$, $F : \ell^\infty \times \ell^\infty \to \mathbb{R} \cup \{+\infty\}$ and

$$F(x) = -\sum_{t=0}^{\infty} \beta^t u(c_t)$$

$\Phi_t = (\Phi_t^1, \Phi_t^2, \Phi_t^3)$ where

$$\begin{aligned}\Phi_t^1(x) &= c_t + k_{t+1} - f(k_t), \forall t \geq 0,\\ \Phi_t^2(x) &= -c_t, \forall t \geq 0,\\ \Phi_t^3(x) &= -k_{t+1}, \forall t \geq 0.\end{aligned}$$

together with $C = dom(F) = \ell^\infty_+ \times \ell^\infty$ and $\Gamma = dom(\Phi) = \ell^\infty \times \ell^\infty_+$ and $C \cap \Gamma = \ell^\infty_+ \times \ell^\infty_+$.

Then $(\tilde{\tilde{P}})$ will be:

$$(\tilde{\tilde{P}}) \begin{cases} \min F(x) \\ s.t. \\ \Phi(x) \leq 0 \\ x \in \ell^\infty \times \ell^\infty \end{cases}$$

Remark that with the above settings the problem $(\tilde{\tilde{P}})$ is equivalent to the optimal growth problem (P).

4.2. Assumptions

$(\widetilde{\widetilde{(EA)}})$ *(Endowment Assumption)*

(1) $k_0 > 0$, *given,*
(2) *The allocations are denoted by* x *and* $x := (k, c) = ((k_t)_t, (c_t)_t) \in \ell^\infty_+ \times \ell^\infty_+$.

$(\widetilde{\widetilde{Prod}})$ *(Production Assumption)*

(1) f *is concave in* $\mathbb{R}_+$,
(2) f *is differentiable in* $\mathbb{R}_+$,
(3) f *is strictly increasing,*
(4) $f(0) = 0, 1 < f'(0) \leq +\infty$, $f'(\infty) < 1$. *And* $f(k) = -\infty$ *if* $k < 0$.

$(\widetilde{\widetilde{Pref}})$ *(Preferences Assumption)*

(1) u *is concave in* $\mathbb{R}_+$,
(2) u *is strictly increasing in* $\mathbb{R}_+$,
(3) u *is differentiable in* $\mathbb{R}_+$,
(4) $u'(0) \leq +\infty$ *and* $u(c) = -\infty$ *if* $c < 0$,
(5) $u(c_0, c_1, \ldots) = \sum_{t=0}^\infty \beta^t u(c_t)$.

Remark 5. (1) *We suppose bounded sequences of allocations by* $(\widetilde{\widetilde{AE}})(2)$.

(2) *One can mention that the above list is indeed weaker than the list in Section 3.1: The boundedness of the utility function is dropped. Neither utility nor the production function are supposed to be stricly concave. Instead, concavity and differentiability will be adequate. However, we make an asymptotic assumption of production function which satisfies* $f'(0) > 1$. *This assumption will be essential for this technique of Lagrange multipliers method (see Example 5 in the following) while it was not essential in the approach of dynamic programming (see also Example 2).*

(3) *By* $(\widetilde{\widetilde{Pref}})(4)$ *we assume additive separable utility, however we refer to [9] for the extension to the recursive preferences.*

Proposition 1. *Under the Assumptions* $(\widetilde{\widetilde{EA}})$, $\widetilde{\widetilde{Prod}}$ *and* $\widetilde{\widetilde{Pref}}$, *if the sequence* $x = (k^*, c^*) \in \ell^\infty \times \ell^\infty$ *is optimal, then there exists* $\lambda \in \ell^1_+$ *such* *that the following conditions hold:*

$$\sum_{t=0}^{\infty}\beta^t u(c_t^*) - \sum_{t=0}^{\infty}\lambda_t^1(c_t^* + k_{t+1}^* - f(k_t^*)) + \sum_{t=0}^{\infty}\lambda_t^2 c_t^* + \sum_{t=0}^{\infty}\lambda_t^3 k_t^* \geq \sum_{t=0}^{\infty}\beta^t u(c_t) - \sum_{t=0}^{\infty}\lambda_t^1(c_t + k_{t+1} - f(k_t)) + \sum_{t=0}^{\infty}\lambda_t^2 c_t + \sum_{t=0}^{\infty}\lambda_t^3 k_t \quad (6)$$

$$\lambda_t^1(c_t^* + k_{t+1}^* - f(k_t^*)) = 0, \forall t. \quad (7)$$

$$\lambda_t^2 c_t^* = 0, \forall t. \quad (8)$$

$$\lambda_t^3 k_t^* = 0, \forall t. \quad (9)$$

Proof. Under the assumptions since u and f are concave then F and Φ are convex. Since $f'(0) > 1$, for any $k_0 > 0$, there exists k' such that $0 < k' + \epsilon < f(k_0)$ and $0 < k' + \epsilon < f(k')$ with $\epsilon > 0$. Let $k^0 = (k_0, k', k', \ldots)$, $c^0 = (\epsilon, \epsilon, \epsilon, \ldots)$ and $x^0 = (k^0, c^0)$.. Note that $\sup_t \Phi_t(c^0) < 0$. Thus *Slater's condition*[4] is satisfied. Under the assumptions made above, in order to be able to apply the result argued in [8] to the space ℓ^1 one needs a key result which is the following identification:

$$(\ell^\infty)' = \ell^1 \oplus \ell^s \text{ (Rudin (1973) in [16])}$$

For each $\lambda \in (\ell^\infty)'_+$ we adopt the notation $\lambda = \lambda^1 + \lambda^s$ where $\lambda^1 \in \ell^1_+$ and $\lambda^s \in \ell^s_+$. The sufficient conditions so that $\lambda^s = 0$ are given by two additional assumptions in [8]. These assumptions are verified with above setting under the assumptions $\widetilde{\widetilde{(EA)}}$, $\widetilde{\widetilde{Prod}}$ and $\widetilde{\widetilde{Pref}}$ for our problem (see [9]). Hence, the conditions of the Main Theorem in [8] are fulfilled for the optimal growth problem. There exists thus $\lambda \in \ell^1_+$ such that for all $x = (k, c) \in \ell^\infty \times \ell^\infty$, if $x^* = (k^*, c^*)$ is optimal, then

$$F(x) + \lambda\Phi(x) \geq F(x^*) + \lambda\Phi(x^*)$$

and

$$\lambda\Phi(x^*) = 0$$

This leads us to the final result with the above settings of Φ, F which establishes the extension of Lagrange Multiplier Method with Karush-Kuhn-Tucker conditions. □

Corollary 2. *The Lagrange multipliers sequence associated to this optimal growth problem is the sequence $\{\beta^t u'(c_t^*)\}$ and satisfies the so-called Euler equation:*

$$u'(c_t^*) = \beta u'(c_{t+1}^*) f'(k_{t+1}^*) \text{ for all } t = 0, 1, \ldots$$

Corollary 3. *Let the assumptions of the Proposition 1 be satisfied for an optimal growth problem. Moreover, suppose that u is strictly concave and continuously differentiable with $u'(0) = +\infty$. If $x^* = (c^*, k^*)$ is an optimal solution, then the sequence $\{\beta^t u'(c_t^*)\}$ is in $\ell^1_+/\{0\}$.*

Let us consider the optimal growth problem with logarithmic utility and Cobb-Douglas production solved in Example 1 and in Example 2 by two different methods. The following example will be the third way of having the solution and will directly generate the Lagrange multipliers:

Example 4. *The assumptions of the Corollary 3 are all satisfied, that is, $u(c_t) = \ln c_t$, therefore it is strictly increasing, continuously differentiable and $u'(0) = +\infty$, we obtain the sequence $\{\beta^t u'(c_t^*)\}$ in $\ell^1_+/\{0\}$:*
As $u'(0) = +\infty$, $c_t^* > 0$ and $k_t^* > 0$, by the equations (8) and (9), we have $\lambda_t^2 = \lambda_t^3 = 0$ for every t. Let us define, $c_t = c_t^*$ for every t, $k_t = k_t^*$ for every $t \neq T$ and $c_T = c_T^* + \epsilon$ such that $c_T^* + \epsilon > 0$. By means of equation (1), we will have:

$$\beta^T u(c_T^*) - \lambda_T^1(c_T^*) \geq \beta^T u(c_T^* + \epsilon) - \lambda_T^1(c_T^* + \epsilon)$$

For all ϵ sufficiently small, we have thus:

$$\beta^T u'(c_T^*) - \lambda_T^1 = 0$$

However, by Proposition 1, $\lambda \in \ell^1_+$ which implies $\{\beta^t u'(c_t^*)\} \in \ell^1_+/\{0\}$.
For the particular case of this example $\lambda_T^1 = \{\frac{\beta^T}{c_T^*}\} \in \ell^1_+$.

Remark 6. *An alternative proof of obtaining the sequence of $\{\beta^t u'(c_t^*)\}$ in $\ell^1_+/\{0\}$ is due to Dana and Le Van (2003) ([17]). Under the assumptions $\widetilde{\widetilde{(EA)}}$, $\widetilde{\widetilde{Prod}}$ and $\widetilde{\widetilde{Pref}}$, it is shown in [17] that there exists a unique optimal sequence $x^* = (c^*, k^*)$ verifying that $c^* > 0$ and $k^* > 0$. Moreover the sequence k^* is monotonic and $x^* = (c^*, k^*)$ satisfies Euler equation which is used to prove the existence of the sequence $\{\beta^t u'(c_t^*)\}$ in $\ell^1_+/\{0\}$. This sequence is interpreted as the prices*

[4]Slater's condition which is a specific example of a constraint qualification states that the feasible region must have an interior point.

p^* of the corresponding competitive equilibrium $(x^*, p^*) = (c^*, k^*, p^*) \in \ell^\infty_+ \times \ell^\infty_+ \times \ell^1_+/\{0\}$. The assumption of $f'(0) > 1$ (which is called as Interiority Assumption in [17]) is essential to have multipliers in ℓ^1_+. Without this assumption, that is if $f'(0) \leq 1$, the multipliers are not necessarily in ℓ^1_+ as in the following example.

Example 5. *Let us reconsider the Cake Eating Problem. Remember that we consider a linear production function $f(z) = z$ for all $z \in \mathbb{R}_+$. Hence $f'(0) = 1$ and the condition $f'(0) > 1$ is not satisfied. Suppose that we have multipliers in ℓ^1_+. By the help of the Inada conditions and Euler equation we will have:*

$$u'(c^*_t) = \beta u'(c^*_{t+1}) f'(k^*_{t+1}) \textit{ for all } t = 0, 1, \dots$$

equivalently

$$\lambda_t = \beta \lambda_{t+1} f'(k^*_{t+1}) \textit{ for all } t = 0, 1, \dots$$

Since $f'(k^*_{t+1}) = f'(0) = 1$, $\lambda_{t+1} > \lambda_t$ for every t which implies $\lambda_t > \lambda_0 = u'(c^*_0) > 0$ proving that $\lambda_t \notin \ell^1_+$. A contradiction.

Hence, a solution cannot be given to the Cake Eating Problem by means of this approach.

4.3. Approach of Pontryagin's principle

In this section, we apply *Theorem 3.1* and *Theorem 5.1* of [15] [5] to our optimal growth problem defined by scalar state and control variables. These theorems establish weak Pontryagin's principles as necessary and sufficient conditions of optimality. The idea of this approach is to transform the optimal growth problem to a dynamical system by the help of weak Pontryagin's principles. This approach is also functional analytic and based on the use of abstract results of optimization theory in the space ℓ^∞ in the spirit of the Karush-Kuhn-Tucker theorem.

The aim of this section is to set the optimal growth problem (P) as an optimal control problem $\widehat{(P)}$ and to show that the necessary conditions given by Theorem 3.1 of [15] and sufficient conditions given by Theorem 5.1 of [15] are fulfilled for $(\widehat{P})$.

Set $x = (k, c) \in \ell^\infty \times \ell^\infty$ and $g(k_t, c_t) := f(k_t) - c_t$ for all $t = 0, 1, \dots$ where $k_t \in \mathbb{R}_+$ is the scalar state variable and $c_t \in \mathbb{R}_+$ is the scalar control variable. The dynamic system is governed by the following difference inequation (DI):

$$(DI)\ k_{t+1} \leq g(k_t, c_t) \text{ for all } t = 0, 1, \dots$$

Then $\widehat{(P)}$ will be:

$$\widehat{(P)} \begin{cases} \max J(x) = J(k, c) := \sum_{t=0}^{\infty} \beta^t u(c_t) \\ s.t. \\ k_{t+1} \leq g(k_t, c_t) \\ k_0 > 0 \text{ given, } c_t \geq 0,\ k_t \geq 0 \end{cases}$$

Remark that with the above settings two problems $\widehat{(P)}$ and (P) are equivalent.

Note that the Pontryagin's Hamiltonian function associated to $\widehat{(P)}$ and the multipliers 1 and λ is defined by $H_t : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that

$$H_t(k_t, c_t, 1, \lambda) := \beta^t u(c_t) + \lambda g(k_t, c_t)$$

Proposition 2. *Let the following assumptions be satisfied:*

$(\widehat{Prod})$ (Production Assumption) $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable,

$(\widehat{Pref})$ (Preferences Assumption) $u : \mathbb{R} \to \mathbb{R}$ is continuously differentiable.

If the feasible accumulation sequence $x^* = (k^*, c^*)$ in $int\ell^\infty_+ \times int\ell^\infty_+$ is an optimal solution of $\widehat{(P)}$ then it is a solution of the following system:

$$u'(c_t) = \beta u'(c_{t+1}) f'(k_{t+1}) \textit{ for all } t = 1, 2 \dots \tag{10}$$

$$f(k_t) = c_t + k_{t+1} \textit{ for all } t = 0, 1, 2 \dots . \tag{11}$$

Conversly, under $\widehat{Prod}$ and $\widehat{Pref}$, let the above equations (10) and (11) be fulfilled for a feasible allocation $x^* = (k^*, c^*)$ in $int\ell^\infty_+ \times int\ell^\infty_+$. Let there exist $(\lambda^*_t) \in \ell^1_+$ such that the Pontryagin's Hamiltonian function, associated to $(\widehat{P})$ and the multipliers 1 and λ, $H_t(k_t, c_t, 1, \lambda)$ is concave with respect to (k_t, c_t) for all $t = 0, \dots$. Then $x^* = (k^*, c^*)$ is an optimal solution of $(\widehat{P})$.

Proof. Since u is independent of k_t and supposed to be continuously differentiable and since f is continuously differentiable then so is $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Under the assumptions $\widehat{Prod}$ and $\widehat{Pref}$, the assumptions[6] of *Theorem 3.1* in [15] are verifed, therefore, we can directly use its conclusion. There exists then a sequence of multipliers $\lambda^* \in \ell^1_+$ such that the following conditions, which

[5] These are also Theorem 3.3 and Theorem 3.8 of [14].

[6] Essentially the Assumption $(H1)$ in [15]. Note that Assumption $(H4)$ is always satisfied in our case since $\frac{\partial g}{\partial c}(k_t, c_t) = -1 \neq 0$ for all $t = 0, 1 \dots$

are so-called *Adjoint Equation* (AE) , *Weak Maximum Principle* (WMP) and *Complementary Slackness* (CS), hold:

$$(AE)\ \lambda_t^* = \nabla_k H_t(k_t^*, c_t^*, 1, \lambda_{t+1}^*)$$

$$(WMP)\ \nabla_c^* H_t(k_t^*, c_t^*, 1, \lambda_{t+1}^*) = 0$$

$$(CS)\ \lambda_{t+1}^*(g(k_t^*, c_t^*) - k_{t+1}^*) = 0$$

which imply respectively:

$$\lambda_t^* = \lambda_{t+1}^* \cdot \frac{\partial g}{\partial k_t^*}(k_t^*, c_t^*) + \beta^t.0 \text{ for all } t = 1, 2, \ldots \tag{12}$$

$$\lambda_{t+1}^* \cdot \frac{\partial g}{\partial c_t^*}(k_t^*, c_t^*) + \beta^t u'(c_t^*) = 0 \text{ for all } t = 0, 1, \ldots \tag{13}$$

$$\lambda_{t+1}^*(g(k_t^*, c_t^*) - k_{t+1}^*) = 0 \text{ for all } t = 0, 1, \ldots \tag{14}$$

that give us the following system:

$$\lambda_t^* = \lambda_{t+1}^* f'(k_t^*) \text{ for all } t = 1, 2, \ldots \tag{15}$$

$$\lambda_{t+1}^*(-1) + \beta^t u'(c_t^*) = 0 \text{ for all } t = 0, 1, \ldots \tag{16}$$

$$\lambda_{t+1}^*(f(k_t^*) - c_t^* - k_{t+1}^*) = 0 \text{ for all } t = 0, 1, \ldots \tag{17}$$

From the equations (15) and (16), the system reduces to:

$$u'(c_t^*) = \beta u'(c_{t+1}^*) f'(k_{t+1}^*) \text{ for all } t = 0, 1, \ldots \tag{18}$$

$$\lambda_{t+1}^*(f(k_t^*) - c_t^* - k_{t+1}^*) = 0 \text{ for all } t = 0, 1, \ldots \tag{19}$$

Remark that the multipliers associated to this problem are defined by $\lambda_{t+1}^* = \beta^t u'(c_t^*)$ and satisfy (18) which is *Euler equation* together with (19).

Conversly, if the equations (18) and (19) are satisfied, then setting $\lambda_{t+1}^* = \beta^t u'(c_t^*)$, the assumptions of *Theorem 5.1* of [15] are fulfilled. That is, as $\lambda_{t+1}^*(f(k_t^*) - c_t^* - k_{t+1}^*) = 0$ for all $t = 0, 1, \ldots$, $f(k_t^*) - c_t^* > k_{t+1}^*$ implies $\lambda_{t+1}^* = 0$. Moreover, since $(\lambda_t^*) \in \ell_+^1$ we have necessarily $\lim_{t\to\infty} \lambda_t^* = 0$ which is so-called *Transversality Condition* at infinity. Moreover, if the Pontryagin's Hamiltonian function $H_t(k, c, 1, \lambda_{t+1}) = \beta^t u(c_t) + \lambda_{t+1}(f(k_t) - c_t)$ is concave with respect to (k_t, c_t) then optimality holds.

□

Remark 7. (1) *Endowment Assumption and Inada conditions are fulfilled by the statement of the Proposition 2 as the sequence $x^* = (k^*, c^*)$ is supposed to be a feasible allocation sequence in $int\ell_+^\infty \times int\ell_+^\infty$.*
(2) *The result is useful as the assumptions are easy to check and one may avoid the concavity assumptions of u and f. However the concavity of the Hamiltonian is needed for the sufficient conditions of the optimality.*

Example 6. *A solution to the problem in Example 1 can be given by the approach of Pontryagin's principle. $u(c_t) = \ln c_t$, $f(k_t) = (k_t)^\alpha$ with $0 < \alpha < 1$ are continuously differentiable on $\mathbb{R}_+$. A solution $x^* = (k^*, c^*)$ to this problem is then equivalent to the solution of the following system which holds by (18) and (19):*

$$\frac{1}{c_t} = \beta\alpha \frac{1}{c_{t+1}}(k_{t+1})^{\alpha-1} \textit{ for all } t = 0, 1, \ldots$$

$$(k_t)^\alpha - c_t - k_{t+1} = 0 \textit{ for all } t = 0, 1, \ldots$$

which generates the optimal sequence: $k_{t+1}^* = \alpha\beta(k_t^*)^\alpha$ for all $t = 0, 1\ldots$ as in Example 1, Example 2 and Example 4.

5. Conclusion

The optimal growth problem and its solution require advanced dynamic optimization techniques. In this paper, we analyze four of them in a discrete time infinite horizon framework. Besides the two classical approaches, namely passing to the limit approach and dynamic programming, we study two functional analytic approaches. The first of them serves as the extension of Lagrangian method to infinite dimensional spaces by emphasizing the works [8] and [9]. The second one transforms the optimal growth problem to a dynamical system by the help of weak Pontryagin's principles. While studying each of these approaches, we discuss the potential difficulties in obtaining the solution and point out possible ways to avoid these difficulties. Under each case, we provide a discussion about the assumptions of the model and review the techniques through some relevant examples.

Optimal growth model typically involves several assumptions on both the production and consumption sides (mainly on preferences). In general, the analysis of the specific assumptions of a model in economic theory is crucial in order to encompass the most interesting cases in the applications of the theoretical models. Some of these assumptions are needed purely for the mathematical reasons, that is, in order to be able to solve the optimization problem. Specifically, we always need some restrictive assumptions on the objective and constraint functions such as concavity, differentiability, monotonocity, boundedness and asymptotic assumptions. Once these assumptions are made and the mathematical framework is established, the solution can be given. Then, from the economic viewpoint, additional efforts are put forward in weakening some of the restrictive assumptions.

This paper provides a comparative analysis of different mathematical approaches based on a specific list of assumptions within the given economic model. First, for the passing to the limit approach to work in optimal growth model, we point out that it is necessary to be able to interchange the limit and maximum operators. This is satisfied only if the solution path sequence is uniformly convergent. Therefore, its economic applicability is limited. Then, we study the dynamic programming technique in the same context and find that it leads to a solution that enables us to consider a larger set of economic examples. To make this point more clear, note that utility functions are often assumed to be unbounded in economics and thus the boundedness assumption needed in the passing to the limit approach is too restrictive while this assumption can be avoided in dynamic programming. We overview important contributions in the literature to apply dynamic programming techniques to analyze infinite horizon optimal growth problems with unbounded returns and with non-additive and recursive preferences via aggregating functions.

We finally show that a solution to the optimal growth problem can be obtained under weaker assumptions on production and preferences by the two functional analytic approaches relative to the previous two techniques. To be more specific, in Lagrange multiplier method, unlike the classical approaches, neither the utility nor the production function is supposed to be stricly concave and continuously differentiable. Instead, concavity and differentiability are adequate. Here, we should emphasize that an additional assumption is made on the asymptotic behavior of the production function which satisfies $f'(0) > 1$. We show that this assumption here is essential while it is not essential in the approach of dynamic programming. The approach of weak Pontryagin's principle is useful as the assumptions are fewer and easy to check. To compare these two functional analytic approaches, we have to note that in Lagrange multiplier method we need the concavity assumptions of u and f but in the approach of weak Pontryagin's principle we do not need. However, note that the concavity of the Hamiltonian is needed for the sufficient conditions of the optimality.

This paper, by its comparative set-up, can be seen as a source for the researchers who intend to use these approaches in similar types of accumulation and growth problems.

Acknowledgments

The author gratefully acknowledges the hospitality of Economics Department of Ryerson University, Toronto, Canada, where a part of the paper was written. This research was supported by the research project TÜBİTAK-2219-1059B191501755.

References

[1] Bellman, R. (1957). Dynamic Programming. Princeton University Press, Princeton.

[2] Stokey, N., Lucas Jr., R.E., & Prescott, E.C. (1989). Recursive Methods in Economic Dynamics. Harvard University Press, New-York.

[3] Lucas, R.E., & Stokey, N. (1984). Optimal growth with many consumers. Journal of Economic Theory, 32(1), 139-171.

[4] Le Van, C., & Morhaim, L. (2002). Optimal growth models with bounded or unbounded returns: a unifying approach. Journal of Economic Theory, 105(1), 158-187.

[5] Kamihigashi, T. (2014). Elementary results on solutions to the Bellman equation of dynamic programming: existence, uniqueness and convergence. Economic Theory, 56(2), 251-273.

[6] Bich P., Drugeon J.P. & Morhaim, L. (2017). On temporal aggregators and dynamic programming. Economic Theory, 1-31.

[7] Bewley, T.F. (1972). Existence of equilibria in economies with finitely many commodities. Journal of Economic Theory, Vol (4), 514540.

[8] Dechert, W.D. (1982). Lagrange multipliers in infinite horizon discrete time optimal control models. Journal of Mathematical Economics, 9(3), 285-302.

[9] Le Van, C., & Saglam, C. (2004). Optimal growth models and the Lagrange multiplier. Journal of Mathematical Economics, 40(3), 393-410.

[10] Rustichini, A. (1998). Lagrange multipliers in incentive-constrained problems. Journal of Mathematical Economics, 29(4), 365-380.

[11] Dutta, J., & Tammer, C. (2006). Lagrangian conditions for vector optimization in Banach spaces. Mathematical Methodes of Operational Research, 64(3), 521-540.

[12] Blot J., & Chebbi, H. (2000). Discrete Time Pontryagin Principles with Infinite Horizons. Journal of Mathematical Analysis and Applications, 246(1), 265- 279.

[13] Blot J., & Hayek, N. (2008). Infinite horizon discrete time control problems for bounded processes. Advances in Difference Equations, 2008(1), 654267.

[14] Blot J., & Hayek, N. (2014). Infinite-horizon optimal control in the discrete-time framework. Springer Briefs in Optimization, Springer, New York.

[15] Blot. J., Hayek N., Pekergin, F. & Pekergin, N. (2015). Pontryagin principles for bounded discrete-time processes. Optimization, 64(3), 505-520.

[16] Rudin, W.(1973). Functional Analysis. McGraw-Hill, New York.

[17] Dana, R.A., & Le Van, C. (2003). Dynamic Programming in Economics. Springer Science & Business Media.

Ayşegül Yıldız Ulus *has been an Assistant Professor in the Department of Mathematics in Galatasaray University. She has a BA in Mathematics (Middle East Technical University-Turkey), M.Sc. and PhD degree in Applied Mathematics (University of Paris I-Panthéon Sorbonne-France). Her research interest includes optimization, functional anlaysis, fixed point theory and mathematical economics.*

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.1, pp.117-126 (2018)
https://doi.org/10.11121/ijocta.01.2018.00420

RESEARCH ARTICLE

A hybrid QFD-AHP methodology and an application for heating systems in Turkey

Yavuz Ozdemir [a*], Pelin Alcan [b], Huseyin Basligil [a], Dursun Cakrak [a]

[a] *Department of Industrial Engineering, Yildiz Technical University, Turkey*
[b] *Department of Industrial Engineering, Istanbul Gelisim University, Turkey*
ozdemiry@yildiz.edu.tr, palcan@gelisim.edu.tr, basligil@yildiz.edu.tr, dursuncakrak@hotmail.com

ARTICLE INFO

Article history:
Received: 22 November 2016
Accepted: 29 November 2017
Available Online: 10 January 2018

Keywords:
Analytic hierarchy process
Heating system
Quality function deployment
Quality house

AMS Classification 2010:
90B50, 62P30

ABSTRACT

Selection of the heating system is based on many characteristics from the customer side. Operating cost, comfort, ease of use and aesthetic of the systems are some of the most important ones of these characteristics. In this article, data is collected primarily for the implementation of quality house. With these data, customer requirements are listed and defined in terms of degree of importance from the customer side. Then, the relationship between customer requirements and technical requirements are described. Also, column weights are calculated according to the defined relations. Finally, the results obtained using a quality house is integrated with Analytic Hierarchy Process (AHP) methodology for system selection. Then results are interpreted. The main contribution of this paper is to determine the best heating system selection using the relationship between customer and technical requirements. To the authors' knowledge, this will be the first study which uses the integrated QFD-AHP method for heating system selection.

1. Introduction and literature review

The residence is a spatiality that has economic value, changing value, aesthetical value and usage value. The residence is a building or a part of building which meets the necessities of people which provides a group of people to live separately from the others and which has a unique door by opening towards directly to the street.

During the all choices of MCDM (Multiple Criteria Decision Making) which aim to assist the decision maker in selection the best is implemented with the help of such methods as ELECTRE, TOPSIS, AHP, etc [1]. The Analytic Hierarchy Process is a methodology which is based on hierarchical structure of criteria, measurement and synthesis. AHP aims to help decision maker to get over the difficulties [2, 3]. Contrarily to other methods, AHP, given a number of functions, allows to specify the most desirable and objective value for each function. This occurs within a matrix of assessment in which the functions appear on both axes. The Quality Function Deployment (QFD)-AHP is a very flexible method, and allows analyzing customer's demands in an effective and objective manner. In particular, it permits to identify the customer's proper needs and to focus on the technical activity about output [4].

In the literature, based on the Analytic Hierarchy Process (AHP), the solar water heating system was the most inexpensive type heater in domestic use [5]. In conclusion, it was found that the solar water heating system was the most desirable system to be used in Jordan.

Nieminen and Huovila [6] described experiences of applying QFD in the decision making process in building design using the IEA (International Energy Agency) task with 23 criteria. Three case studies were shortly presented. The study [7] specified the fundamental requirements for a prioritization process. Where prioritization should take place during the requirements phase, and who should be involved in the prioritization process were studied. Current techniques such as AHP and QFD were analyzed to how well they satisfy the fundamental needs of a successful prioritization process. A framework was described that incorporates the many aspects of prioritizing requirements.

The thesis of Alanne [8] i) identified the need of decision support in the commercialization of

*Corresponding author

sustainable energy technologies in buildings, ii) characterized decision-making problems related to the above context, iii) developed and implemented a methodology to assess energy technologies for buildings, and iv) presented two fields of application where the above assessment is essential. Moreover, a multi-criteria portfolio model was applied to determine the most preferred retrofit measures in an apartment building. In their paper, Alanne et al. [9] considered the selection of a residential energy supply system as a multi-criteria decision-making (AHP) problem, which involved both financial and environmental issues. On the other hand, as an update of Huang et al.'s article, the study of Zhou et al. [10] gave the developments of DA (Decision Analysis) in E&E (Energy and Environmental modeling) in recent years. That survey showed the increased popularity of MCDM methods. Besides, the working paper of Nebel et al. [11] was an interim report from the Systems Research Work Plan - "Criteria Development and Embedding Systems". Two systems were selected from a prioritized list of residential building systems obtained through a series of workshops and project team discussion meetings by AHP method. The aim of the work of De Felice and Petrillo [4] was to propose a new methodological approach about defining customer's specifications through the instrument of an integrated QFD-AHD model. AHP was well designed for that because of its mathematically and rigorous process for prioritization and decision making.

With this study, the hopes of people from heating systems which are used in the houses and which will be used at the future and the differences between these heating systems are emphasized. Customer demands are emphasized with the Quality Function Deployment (QFD) application and the technical requirements are listed and the comparison are made. And the technical necessaries are listed. AHP application is made by taking the results from Quality house application for system choosing. The alternatives are radiator, fan-coil, air-condition and floor heating systems. When we investigate the researches about this topic, QFD-AHP, a study that comprises all these four heating systems was not encountered. The provided results can be a numerically guide for the CIBSE Best Practice guide.

In the sections that follow, we first present the heating systems in Section 2. We then define QFD and methodology in Section 3. A QFD development for heating systems is explained in Section 4. A description of AHP methodology is given in Section 5. In Section 6, heating system selection using Analytic Hierarchy Process (AHP) is given. Finally, Section 7 concludes the paper and points future work.

2. Heating systems and exports in Turkey

A heating system is a mechanism for maintaining temperatures at an acceptable level; by using thermal energy within a home, office, or other dwelling. While considering about efficient energy rating, some factors are taken into consideration, such as thermal irregularities in building envelope, energy efficiency of the boilers, the distribution system and the performance of the control system [12]. The floor heating system has constituted the rate of 50% of the heating system in the recent days at Europe. The rest of the rate has been including radiator, convector and the others.

Also, heating systems and equipments consist of burners, boilers, radiators, water heaters, dehumidifiers, electric and non-electric heaters, stoves and their equipments. In 2013, heating systems and equipments export of Turkey increased by 3,7% with respect to previous year and reached US$ 1,9 billion. According to data of 2013, in Turkey's heating systems and equipments export, Iraq, United Kingdom, Germany, Azerbaijan, and Turkmenistan are the top five countries (Figure 1) [13].

Rank	Country	2009	2010	2011	2012	2013	Change 12-13 %	2013 Share
1	Iraq	118.904	127.560	152.138	229.711	217.390	-5,36%	10,96%
2	United Kingdom	174.638	163.740	191.900	182.377	198.679	8,94%	10,01%
3	Germany	85.773	77.177	195.111	189.025	174.760	-7,55%	8,81%
4	Azerbaijan	49.441	52.466	86.700	98.154	140.604	43,25%	7,09%
5	Turkmenistan	76.206	84.619	106.752	90.061	102.686	14,02%	5,18%
6	Russian Federation	50.046	30.762	38.392	57.595	85.118	47,79%	4,29%
7	France	40.455	47.963	66.610	57.340	59.304	3,43%	2,99%
8	China	8.924	15.446	20.824	32.238	54.724	69,75%	2,76%
9	Romania	61.975	57.823	63.078	55.203	53.586	-2,93%	2,70%
10	Oman	2.345	2.025	2.785	37.892	51.618	36,22%	2,60%
11	Ukraine	39.638	42.701	56.690	54.892	49.215	-10,34%	2,48%
12	Italy	22.987	30.352	49.357	42.046	42.391	0,82%	2,14%
13	Saudi Arabia	35.379	31.501	18.148	39.132	41.202	5,29%	2,08%
14	Libya	62.073	53.256	12.574	31.393	35.157	11,99%	1,77%
15	Iran	12.273	29.246	51.366	49.975	34.660	-30,65%	1,75%
16	Spain	35.487	36.727	35.995	36.008	34.536	-4,09%	1,74%
17	Georgia	15.154	20.106	30.877	36.515	32.982	-9,68%	1,66%
18	Poland	21.935	23.322	32.019	26.967	30.373	12,63%	1,53%
19	Algeria	32.211	16.485	22.630	47.399	27.433	-42,12%	1,38%
20	Belgium	12.342	15.028	14.465	21.506	27.143	26,21%	1,37%
	Others	474.929	435.073	452.904	497.275	490.512	-1,36%	24,72%
	TOTAL	1.433.138	1.393.378	1.701.345	1.912.700	1.984.063	3,73%	100,00%

Figure 1. Turkey heating systems and equipments export by country (thousand $) [13].

2.1. Floor heating systems

Simulation model of floor heating system is mainly introduced as heat transfers in pipe to indoor and also this usage of it is approved as the basic shape for characterization and dimension. Different types of floor heating system have been investigated and at this point it is considered to being of finited element models with respect to thermal properties and dynamical behavior. The classification of the thermal output to indoor has been established with the purpose of being able to designed and dimensioned such as system in EN1264. Various kinds of control strategies are investigated not to loss indoor heat and consume the energy. Various floor covering materials have been found to impact temperatures, reaction time and energy consumption [14]. The heating floor elements such as, water, coils, electric cables are placed into concrete layer in the floor [15].

2.2. Radiator heating systems

As for radiator systems, the movement of the air heated by grazing the hot radiator surfaces towards the part of the room that is close to the ceiling and the presence of relatively cool air at the inferior half of the room which is the real usage capacity cannot be prevented. Because of this sufficiency of the heat diffusion at the horizontal and vertical sections in the room, the pleasant warmth on the floor surface and the thermal satisfaction of the person with the wall radiation effect in the floor heating, many practitioners confirm that the room temperatures anticipated in the planning of the floor heating need to be kept 1 – 2 ºC lower than the room temperatures given in the literature. Considering that a decrease of 1 ºC in the room temperature leads to a fuel economy of 7%, the superiority of the system on this matter can be emphasized.

In the current survey, a high powered density radiator using for the hydronic central heating applications has been developed for utilizing heat pipes. A heat pipe is hermetically sealed a light-water tube which exists inside the heat pipe shell as vapor and liquid at equilibrium [16]. In order to release hot weather from the distribution system into the building to save indoor energy and temperature, the heat emitters are used. Heat emitters which are commonly used are radiators, under-floor heating, fan-coil units (FCU) and air-handling units (AHU). This survey also showed that 95% of radiators were controlled by using TRVs (thermostatic radiator valves) and revealed that more than 65% of TRVs were performing very poorly [12].

2.3. Fan coil heating systems

First of all, fan coil system using is very useful and easy. Secondly, devices which could be hidden are comparatively aesthetic. Warming period is fastly reacting to the environment. Finally, system is relatively controllable.

2.4. Air-condition heating systems

Although the Turkish HVAC-R (heating, ventilating and air-conditioning & refrigeration) sector began to get organized in 1993, Turkey’s interest in the heating, ventilating, and air-conditioning sector dates back to the 1950’s. After that time, this industry has grown quickly both in terms of manufacturing and volume, expending its domestic and foreign markets. This growth has been expedited by a number of factors, including Turkey’s young population, the country’s steadily increasing GDP (gross domestic product), and the public’s growing demand for comfortable living standards [17].

The utilization of the system is very useful as fan coil heating systems. Also, system is fairly flexible due to the equipment could be camouflaged. The reaction of the system is very expeditious in terms of warm-up time. Besides, it can be simply controlled in terms of inspection. But, climate heating systems cannot be operating much more efficient in cold climate regions.

3. QFD and QFD methodology

Quality Function Deployment is a systematic approach to design based on a close awareness of customer desires, coupled with the integration of corporate functional groups. It consists in translating customer desires (for example, the ease of writing for a pen) into design characteristics (pen ink viscosity, pressure on ball-point) for each stage of the product development [18]. Figure 2 shows the quality house basic parts. Also, the main parts of a quality house matrix presented in Figure 3 is modeled.

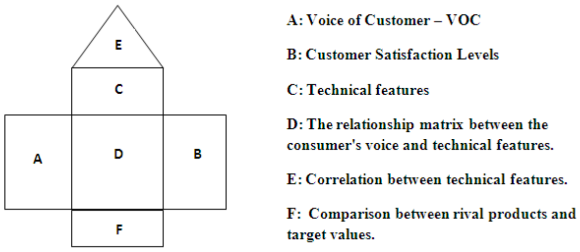


Figure 2. Quality house basic parts.

In this study, firstly the customer requirements were defined for quality house (QFD) application. Survey and double meetings were made when these requirements are defining. The quality house application was made with the data that was taken from surveys and AHP application in heating system selection was made in accordance with customer needs.

Beginning with the initial matrix, commonly referred as quality house (Figure 3), the QFD methodology focuses on the most important product or service attributes or qualities. These are composed of customer wants, and musts. Firstly, customer requests and technical requirements are determined. Then the relation between the customer request and technical requirements and the relation between the technical features are defined.

Evaluation of customers and technical evaluation according to competition are found and according to the firm goals technical importance values and normalized technical importance values are calculated as detailed in Section 4.

The methodology is summarized as shown in the Figure 4.

4. A QFD development for heating systems

The aim of this application is to see the applicability of QFD technique in the heating systems that based on customer expectations and customer satisfactions.

4.1. Forming the customer data part of the QFD matrix

4.1.1. Determining the customer demands

Expectations and demands from the heating systems and the selection criteria for the heating systems are asked to the customers and technical requirements are determined as shown in Table 1.

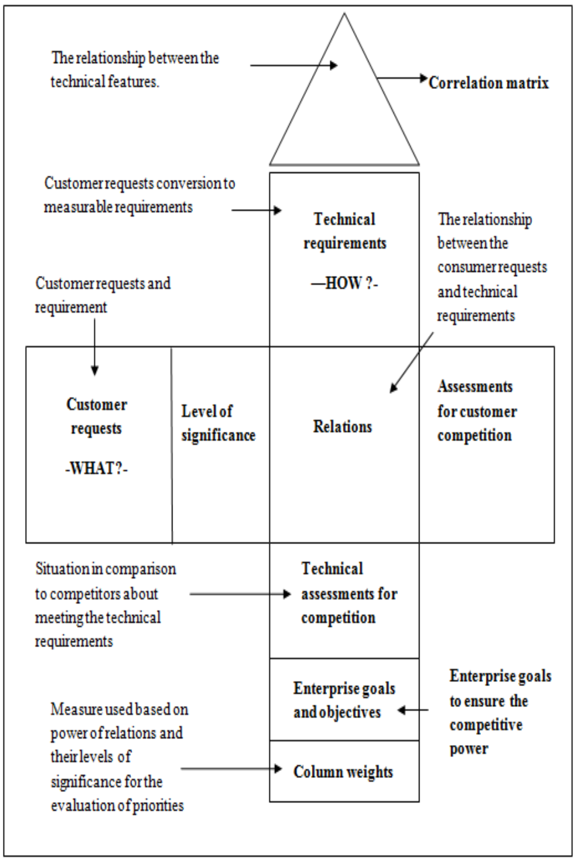


Figure 3. Quality house matrix basic parts.

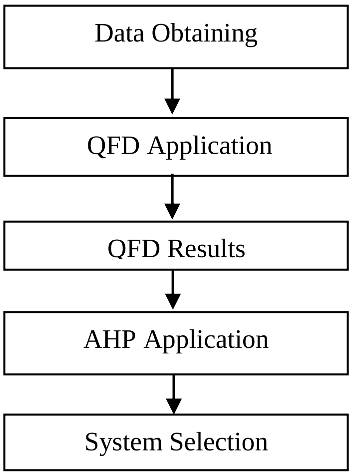


Figure 4. Methodology.

4.2. Arranging the relation between customer demands and the technical requirements

Customers' views are scaled with 1-9 scale that demonstrates 1-the least important, 9-the most important. Also firm experts are evaluated radiator system and floor heating systems with 1-5 scale that demonstrates 1-the worst, 5-the best as shown in Figure 5.

Table 1. Customer demands and the technical requirements for the heating systems.

Customer Demands	Technical Requirements
To be comfortable	Regular heat diffusion
To work with low operating costs	Heat insulation
To be aesthetic	Hidden devices and pipes
To be easy to use	Using thermostat
To response quickly	Ability to work in high-temperatures
To be responsive to the environment	Low CO_2 emission
Not to dust	Low temperature systems
To be compatible with the renewable energy sources	Ability to work with solar power
To save energy	High productive systems
To be easy to control	Using control equipment
To have smart appearance	Aesthetic devices
To be hide out	Hidden systems to the ceiling or to the ground
Ability to work with the outer air	Outer air temperature control
Ability to control each room detached	Thermostatic valves

4.3. Correlations and calculating the column weights

There can be positive or negative interactions between technical requirements that defined for covering the customer demands. Therefore "correlation matrix" is used for seeing these interactions.

In this matrix each cell represents the correlation between two different technical requirements and the positive relation can be shown with ✓ and the negative relation can be shown with X. The most important advantage of correlation matrix is being an indicator of negative relations. Each of the negative relation must be inspected while developing the product. Therefore the changes are determined that can be done for reducing the effect of any negative correlation. After determining the development direction the column weights are calculated (strong, middle, and low relations have 9, 3, 1 weights, respectively).

As an example, "to work with low operating costs" has 9 importance scales and has a strong relation with "heat insulation". Technical importance=Importance scale x Relation weight =9 x 9 = 81.

4.4. Analysis of the QFD matrix

As seen on Figure 6, the customer demands for heating systems are evaluated and the most important results of the house of quality for heating systems are; to be comfortable (8), to work with low operating costs (9), not to dust (8), to save energy (9), to have smart appearance (8), ability to control each room detached (8).

The analyzed firm in this study prefers mostly floor heating systems, therefore radiator systems are competitor for floor heating systems. But having regular heat diffusion, hidden devices and pipes and ability to work efficiently at low temperatures reinforce floor heating systems. Also ability to work efficiently at low temperatures causes to work with low operating costs, so one of the most important results for customer demands is satisfied.

From the technical importance point of view, the most important point is "using thermostat and control equipment". Therefore using these equipment causes to save energy, to be comfortable and ability to control the system according to the temperature of outer air.

As a development direction point of view, buildings that save more energy can be made with increasing the thickness of the insulation. Low operating costs can be obtained with increasing the number of thermostat and control equipment. For the purpose of reducing CO2 emission and being responsive to the environment, central boiler rooms must be enforced.

The analysis of the QFD matrix is concluded with the interpretation of the technical importance and normalized technical importance values. Technical requirements that have the maximum technical importance values are respectively;

- Using thermostat and control equipment
- Low temperature systems
- Hidden systems to the ceiling or to the ground
- Ability to work with solar power
- Hidden devices and pipes

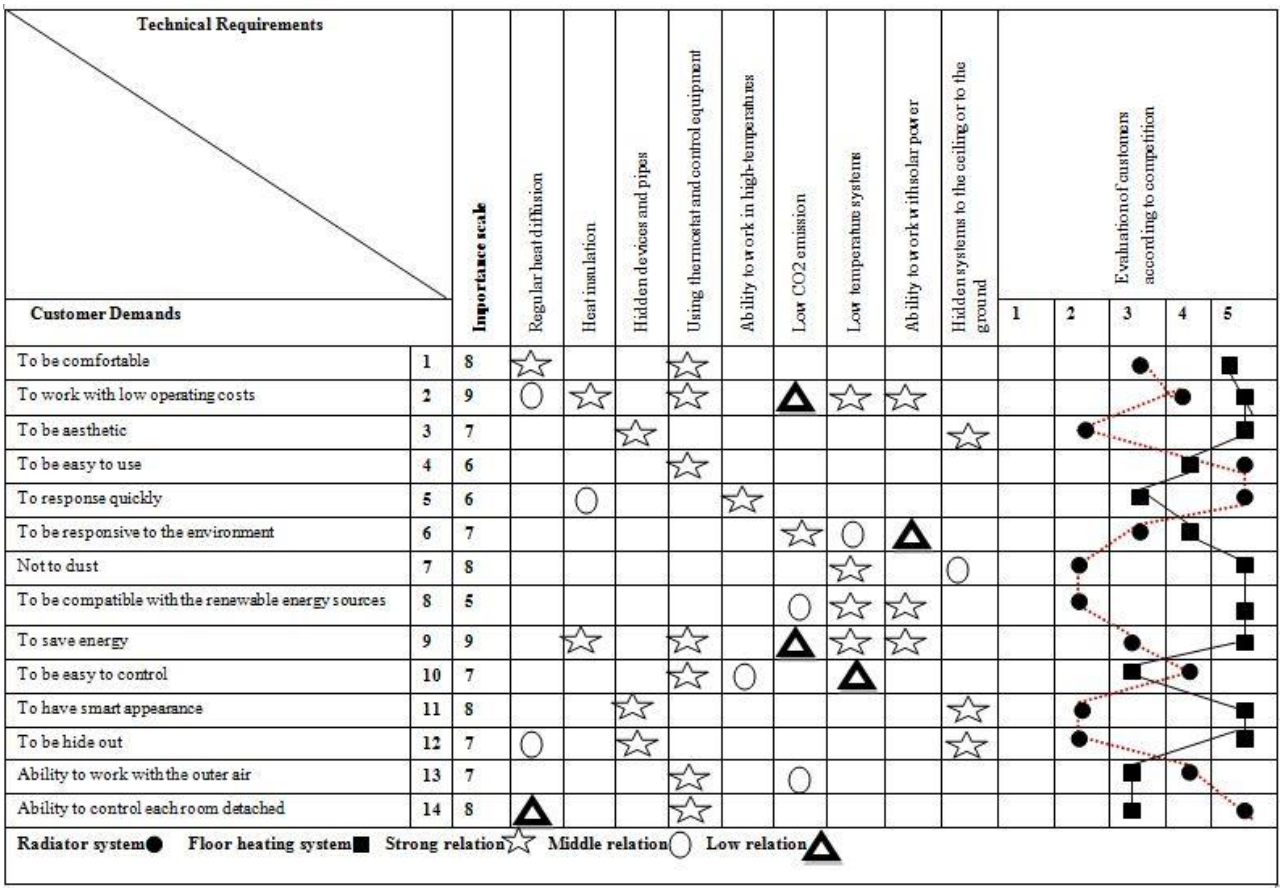


Figure 5. The relation between customer demands and the technical requirements.

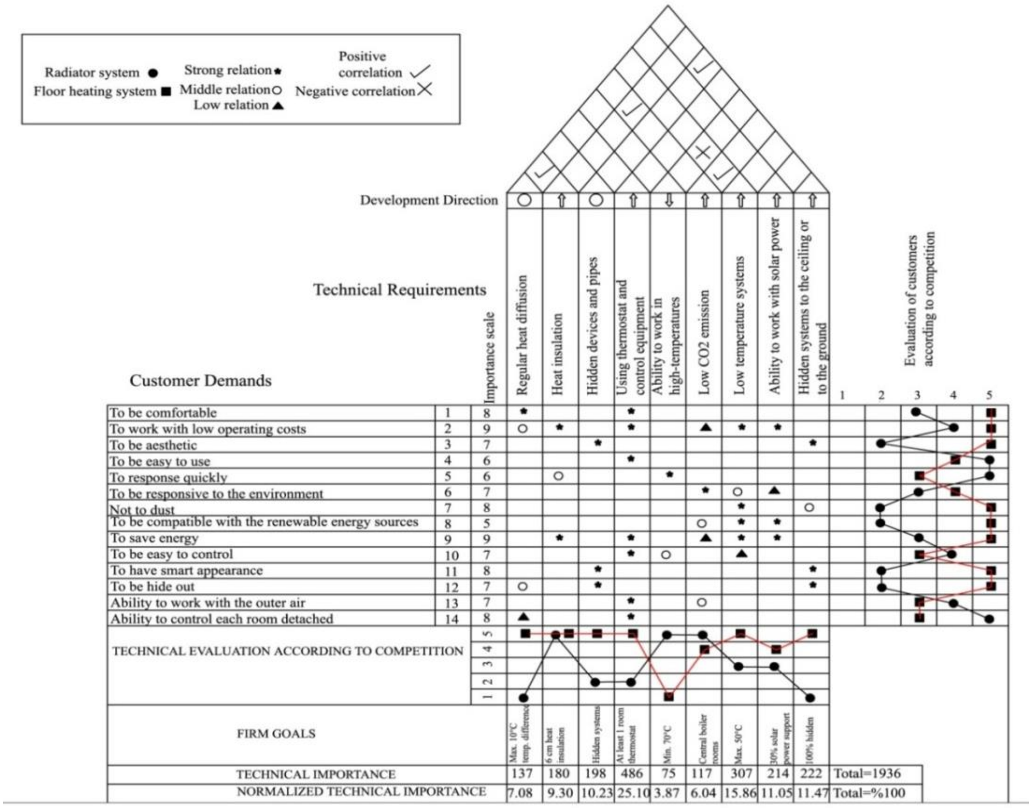


Figure 6. The house of quality for heating systems.

For the purpose of satisfying these technical requirements and customer demands, some heating systems alternatives will be evaluated and prioritized in the Section 6 using Analytic Hierarchy Process (AHP). Customer demands with high importance values will be the criteria for the heating system evaluation process. Therefore we will combine QFD with AHP.

5. Analytic Hierarchy Process (AHP)

The Analytic Hierarchy Process (AHP) is a multi-criteria decision-making method that has been widely used since 1970s. It separates a problem into smaller pieces and examines the effect of these parts on each other. As a result of this process, the weight of parts and the importance order of parts are obtained. For this purpose, a benchmark scale was established that quantitatively assesses the effects of parts on each other. Parts of the problem are compared pair wise and effects of each part on the target are quantitatively obtained. The AHP method can be used in both social and physical areas to make measurement [19].

Steps of AHP is given below;

1. Identification of the problem and determination of the desired information,
2. Formation of the hierarchy of decision-making from top to bottom determination of the goal and criteria,
3. Obtaining pair wise comparison matrix,
4. Finding weights of criteria.

There is a need for a scale to make comparisons. This scale shows how important an element is compared to the other element. The scale used in AHP can be seen in Table 2 [20].

Table 2. AHP Scale.

Importance Values	Value Definitions	Explanation
1	Both factors are equally important	Both activities have an equal importance.
3	Factor 1 is slightly more important than Factor 2	Experience and judgment shows that Factor 1 is slightly more important than the other.
5	Factor 1 is more important than Factor 2	Experience and judgment shows that Factor 1 is more important than the other.
7	Factor 1 is strongly more important than Factor 2	Experience and judgment shows that Factor 1 is strongly more important than Factor 2.
9	Factor 1 has absolute superiority over Factor 2	Experience and judgment shows that Factor 1 is absolutely more important than the other.

The mathematical realization of AHP will be explained in the following steps [21].

1. First, the problem and elements (criteria) to be decided are defined. Using these elements, a comparison matrix is constructed. The comparison

matrix for "n" elements contains "nxn" elements and the values on the diagonal (where i = j) are 1.

$$K = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ k_{21} & k_{22} & \dots & k_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ k_{n1} & k_{n2} & \dots & k_{nn} \end{bmatrix} \quad (1)$$

In the comparison matrix there is such relation, between the elements above the diagonal and the elements below the diagonal;

$$k_{ji} = \frac{1}{k_{ij}} \quad (2)$$

For example, if the third criterion more important than the second criterion, the value of element k_{23} is 5 and k_{32} element has a value of 1/5.

2. This matrix shows us the importance of each criterion, but does not allow us to see the weight of each criterion in total. We need to get the column vectors for this. Each element is divided by the sum of the values in its column, and if the value is substituted, n column vectors of n elements are obtained.

$$s_{ij} = \frac{k_{ij}}{\sum_{i=1}^{n} k_{ij}} \quad (3)$$

The above formula is used when the values of the column vector are calculated.

$$S_i = \begin{bmatrix} s_{11} \\ s_{21} \\ . \\ . \\ . \\ s_{n1} \end{bmatrix} \quad (4)$$

3. To create column matrix, n column vectors are are formed in a matrix. This matrix is as follows

$$S = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ s_{n1} & s_{n2} & \dots & s_{nn} \end{bmatrix} \quad (5)$$

4. Finally, using the S matrix, we need to obtain the weight vector to obtain the percentage of the elements. This is obtained by taking the arithmetic mean of the elements in the rows of the column matrix.

$$A_i = \frac{\sum_{j=1}^{n} s_{ij}}{n} \quad (6)$$

The sum of the elements of the weight vector is 1. The weight vector is as follows;

$$A = \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \end{bmatrix} \quad (7)$$

Consistency analysis is used to measure the consistency after weight results are found. This analysis shows whether there is an error in the work done or the result is consistent within itself. The following steps are taken to calculate the consistency rate [21, 22].

1. In order to calculate the consistency ratio, firstly the comparison matrix and the weight matrix are multiplied to obtain the T column vector.

$$T = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ k_{21} & k_{22} & \dots & k_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ k_{n1} & k_{n2} & \dots & k_{nn} \end{bmatrix} x \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ . \\ . \\ . \\ t_n \end{bmatrix} \quad (8)$$

2. After obtaining the T vector, basic value elements are obtained by dividing each element of the T vector by the weight vector A of the T vector.

$$E_i = \frac{t_i}{a_i} \quad i = 1,2,\dots,n \quad (9)$$

3. The arithmetic mean of these elements gives the basic value of pair wise comparison of this problem λ.

$$\lambda = \frac{\sum_{i=1}^{n} E_i}{n} \quad (10)$$

4. After obtaining λ, consistent indicator Cl should be obtained.

$$CI = \frac{\lambda - n}{n - 1} \quad (11)$$

5. The following formula is used to calculate the consistency ratio at the last step.

$$CR = \frac{CI}{RI} \quad (12)$$

The result is consistent if the consistency ratio (CR) is less than 0.1. If it exceeds 0.1, either there is a mistake in applying the AHP, or the operation is inconsistent.

In this study, AHP application was made using the Super Decisions software and the consistency ratio for all comparisons were found less than 0.1.

6. Heating system selection using Analytic Hierarchy Process (AHP)

For selecting the best heating system alternative for indoor-use, according to the results of QFD, customer demands with high importance values are the criteria for AHP. These are operating costs, to be easy to use, appearance, comfort, and saving energy. Also, the alternatives for the selection process are floor heating systems, radiator, air-condition, and fan-coil.

6.1. Comparing the alternatives

After the purpose, criteria and alternatives have determined, binary comparisons have done with 3 different experts from the sector and the academia. After all of binary comparisons have completed, the averages of their views are entered to Super Decisions software as shown in Table 3-8. After all of data have entered the program, lastly the result can be found as shown in Table 9.

Table 3. Comparing the alternatives according to the "saving energy" criteria.

	Radiator	Fan-Coil	Floor Heating	Air-Condition
Radiator	1	1/3	1/4	3
Fan-Coil	3	1	1/2	2
Floor Heating	6	4	1	5
Air-Condition	1/3	1/4	1/5	1

Table 4. Comparing the alternatives according to the "appearance" criteria.

	Radiator	Fan-Coil	Floor Heating	Air-Condition
Radiator	1	1/4	1/8	1/3
Fan-Coil	4	1	1/6	2
Floor Heating	8	6	1	6
Air-Condition	3	1/2	1/6	1

Table 5. Comparing the alternatives according to the "to be easy to use" criteria.

	Radiator	Fan-Coil	Floor Heating	Air-Condition
Radiator	1	1/4	1/5	1/4
Fan-Coil	4	1	1/3	2
Floor Heating	5	3	1	3
Air-Condition	4	1/2	1/3	1

Table 6. Comparing the alternatives according to the "operating costs" criteria.

	Radiator	Fan-Coil	Floor Heating	Air-Condition
Radiator	1	1/3	1/5	2
Fan-Coil	3	1	1/3	2
Floor Heating	5	3	1	4
Air-Condition	1/2	1/2	1/4	1

Table 7. Comparing the alternatives according to the "comfort" criteria.

	Radiator	Fan-Coil	Floor Heating	Air-Condition
Radiator	1	1/4	1/6	1/3
Fan-Coil	4	1	1/3	2
Floor Heating	6	3	1	5
Air-Condition	3	1/2	1/5	1

Table 8. Comparing the alternatives.

	Saving Energy	Operating Costs	Appeara nce	Comfort	To Be Easy To Use
Saving Energy	1	2	3	3	4
Operating Costs	1/2	1	4	3	5
Appearance	1/3	1/4	1	1/4	1/3
Comfort	1/3	1/3	4	1	3
To Be Easy To Use	1/4	1/5	3	1/3	1

Table 9. AHP results.

Alternatives	Total	Normal	Ideal	Ranking
Radiator	0.0537	0.1075	0.2002	4
fan-coil	0.1214	0.2428	0.4521	2
floor heating	0.2685	0.5370	10.000	1
air-condition	0.0564	0.1128	0.2101	3

As a result, according to the criteria and the evaluation, the most appropriate heating system is floor heating system (53.7%), then fan-coil (24.28%), air-condition (11.28%), and radiator (10.75%), respectively.

7. Conclusion

Heating systems directly affect customers' comfort and life quality; therefore construction companies must pay attention to quality and market research. For this reason several techniques were developed for several purposes; using QFD methodology, customer demands are emphasized and the technical requirements are listed and the comparison can be made. Using the AHP methodology, the decision maker can make decisions according to the criteria and the alternatives.

In this study, firstly QFD analysis has done for the heating systems and with this analysis, customer demands, technical requirements, correlation between them, and the technical importance have determined. For the purpose of satisfying these technical requirements and customer demands, some heating system alternatives have evaluated and prioritized using AHP.

The general limitation of the proposed model is the costly and exhausting information requested from experts (approx. 105 pairwise comparisons per one expert). Other limitations of the model are the preferences of the expert including uncertainty and conflicts and there is often needed more than one expert to make decisions.

According to the results, the most appropriate heating system alternative is floor heating system (53.7%), then fan-coil (24.28%), air-condition (11.28%) and radiator (10.75%), respectively. Also we have to say that, this is the first paper in the literature that combines QFD with AHP methodology in the heating system sector. As a further research, we think to improve this study with fuzzy numbers and also we consider combining QFD with other selection methodologies, such as Analytic Network Process (ANP), TOPSIS and ELECTRE. Besides, we will compare the results that found in this paper.

References

[1] Lin, M.C., Wang, C.C., Chen, M.S., & Chang, A. (2008). Using AHP and TOPSIS approaches in customer-driven product design process. *Computers in Industry,* 59(1), 17-31.

[2] Saaty, T.L. (1980). *The analytic hierarchy process*. McGraw-Hill, New York.

[3] Saaty, T.L. (1990). How to make a decision: the analytic hierarchy process. *European Journal of Operational Research*, 48, 9-26.

[4] De Felice, F., & Petrillo, A. (2010). A multiple choice decision analysis: an integrated QFD – AHP model for the assesment of customer needs. *International Journal of Engineering, Science and Technology*, 2, 25-38.

[5] Mohsen, M.S., & Akash, B.A. (1997). Evaluation of solar water heating system in Jordan using Analytic Hierarchy Process. *Energy Conversion and Management*, 38, 1815-1822.

[6] Nieminen, J., & Huovilla, P. (2000). QFD in design process decision making. In: *International Symposium on Intergrated Life-Cycle Design of Materials and Structures*. 22-24 May, Helsinki, 51-56.

[7] Moisiadis, F. (2002). The Fundamentals of prioritising requirements. In: *Systems Engineering, Test & Evaluation Conference*. October, Sydney, Australia.

[8] Alanne, K. (2007). Applications of decision analysis in the assesment of energy technologies for buildings. System Analysis Laboratory Research Reports, Helsinki University of Technology.

[9] Alanne, K., Saho, A., Saari, A., & Gustafsson, S. (2007). Multi Criteria Evaluation of Residential Energy Supply Systems. *Energy and Buildings,* 39(12), 1218-1226.

[10] Zhou, P., Ang, B.W., & Poh, K.L. (2006). Decision analysis in energy and environmental modeling: An update. *Energy,* 31, 2604-2622.

[11] Nebel, B., Bayne, K., & Crookshanks, C. (2008). Prioritisation of building systems. *Beacon Pathway Limited Report*.

[12] Liao, Z., Swainson, M., & Dexter, A.L. (2004). On the control of heating systems in the UK. *Building and Environment*, 40(3), 343-351.

[13] http://www.turkishheating.com/main/36/heating_industry [Accessed 1 November 2016]

[14] Weitzmann, P., Kragh, J., Roots, P., & Svendsen, S. (2005). Modelling floor heating systems using a validated two-dimensional ground-coupled numerical model. *Building and Environment,* 40, 153-163.

[15] Hasan, A., Kurnitski, J., & Jokiranta, K. (2009). A combined low temperature water heating system consisting of radiators and floor heating. *Energy and Buildings*, 41, 470-479.

[16] Kerrigan, K., Jouhara, H., O'Donnell, G.E., & Robinson, A.J. (2011). Heat pipe-based radiator for low grade geothermal energy conversion in domestic space heating. *Simulation Modelling Practice and Theory*, 19, 1154-1163.

[17] http://www.turkishairconditioning.com/main/38/isib [Accessed 1 November 2016]

[18] Rosenthal, S.R. (1992). Effective product design and development, How to cut lead time and increase customer satisfaction. McGraw-Hill Professional Publishing, Business One Irwin.

[19] Yaralıoğlu, K. (2001). Performans Değerlendirmede Analitik Hiyerarşi Prosesi. *Journal of Dokuz Eylül University Faculty of Economics and Administrative Sciences*, 16(1), 129-142.

[20] Saaty, R.W. (1987). The Analytical Hierarchy Process-What it is and how it is used. *Math Modelling*, 9 (3-5), 161-175.

[21] Kaan, Y. *Analytic Hierarchy Process, Lecture Note [online]*. Available from: http://www.deu.edu.tr/userweb/k.yaralioglu/dosyalar/Analitik_Hiyerarsi_Proses.doc

[22] Chang, J.I., & Liang, C. (2009). Performance evaluations of process safety management systems of paint manufacturing facilities. *Journal of Loss Prevention in the Process Industries*, 22, 398-402.

***Yavuz Ozdemir** was born in İstanbul (Bakirkoy) in 1985. He graduated from Vefa High School, and received his BSc, MSc and PhD from Yildiz Technical University, in the Department of Industrial Engineering. He also works as an engineer at Yildiz Technical University. His research areas are logistics, optimization, evaluation of investments, multi criteria decision making and airline transportation.*

***Pelin Alcan** was born in İstanbul (Üsküdar) in 1981. She graduated from Haydarpaşa High School and received her BSc, MSc and PhD from Yildiz Technical University in the Department of Industrial Engineering. She also worked as a Research Assistant at Yildiz Technical*

University in the year of between 2007 and 2014. She is now working in Istanbul Gelisim University as Asst. Prof. Her research areas are scheduling, optimization and operations research, meta-heuristics and fuzzy logic.

Huseyin Basligil *was born at 1949. He received both BSc and MSc from Istanbul University in the Department of Physics, and a PhD from the same university in the Department of Chemical Engineering. He has been a Professor at Yildiz Technical University in the Department of Industrial Engineering since 1996. He has several academic and administrative duties. His research areas are production planning and control, systems engineering, simulation, project management. He has presented several national and international papers and also written a number of books.*

Dursun Cakrak *was born in Giresun in 1984. He graduated from Hamdi Bozbag Anatolian High School, and received his BSc from Yildiz Technical University, in the Department of Mechanical Engineering and MSc from Istanbul Technical University, in the Department of Industrial Engineering. He is now a PhD candidate in the Yildiz Technical University, Department of Industrial Engineering. He works as a team manager at Yapı Proje Uygulama. His research areas are HVAC systems, project management and quality management.*

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

INSTRUCTIONS FOR AUTHORS

Aims and Scope

This journal shares the research carried out through different disciplines in regards to optimization, control and their applications.

The basic fields of this journal are linear, nonlinear, stochastic, parametric, discrete and dynamic programming; heuristic algorithms in optimization, control theory, game theory and their applications. Problems such as managerial decisions, time minimization, profit maximizations and other related topics are also shared in this journal.

Besides the research articles expository papers, which are hard to express or model, conference proceedings, book reviews and announcements are also welcome.

Journal Topics

- Applied Mathematics,
- Financial Mathematics,
- Control Theory,
- Game Theory,
- Fractional Calculus,
- Fractional Control,
- Modeling of Bio-systems for Optimization and Control,
- Linear Programming,
- Nonlinear Programming,
- Stochastic Programming,
- Parametric Programming,
- Conic Programming,
- Discrete Programming,
- Dynamic Programming,
- Optimization with Artificial Intelligence,
- Operational Research in Life and Human Sciences,
- Heuristic Algorithms in Optimization,
- Applications Related to Optimization on Engineering.

Submission of Manuscripts

New Submissions

Solicited and contributed manuscripts should be submitted to IJOCTA via the journal's online submission system. You need to make registration prior to submitting a new manuscript (please click here to register and do not forget to define yourself as an "Author" in doing so). You may then click on the "New Submission" link on your User Home.

IMPORTANT: If you already have an account, please click here to login. It is likely that you will have created an account if you have reviewed or authored for the journal in the past.

On the submission page, enter data and answer questions as prompted. Click on the “Next” button on each screen to save your work and advance to the next screen. The names and contact details of at least four internationally recognized experts who can review your manuscript should be entered in the "Comments for the Editor" box.

You will be prompted to upload your files: Click on the “Browse” button and locate the file on your computer. Select the description of the file in the drop down next to the Browse button. When you have selected all files you wish to upload, click the “Upload” button. Review your submission before sending to the Editors. Click the “Submit” button when you are done reviewing. Authors are responsible for verifying all files have uploaded correctly.

You may stop a submission at any phase and save it to submit later. Acknowledgment of receipt of the manuscript by IJOCTA Online Submission System will be sent to the corresponding author, including an assigned manuscript number that should be included in all subsequent correspondence. You can also log-on to submission web page of IJOCTA any time to check the status of your manuscript. You will receive an e-mail once a decision has been made on your manuscript.

Each manuscript must be accompanied by a statement that it has not been published elsewhere and that it has not been submitted simultaneously for publication elsewhere.

Manuscripts can be prepared using LaTeX (.tex) or MSWord (.docx). However, manuscripts with heavy mathematical content will only be accepted as LaTeX files.

Preferred first submission format (for reviewing purpose only) is Portable Document File (.pdf). Please find below the templates for first submission.

Click here to download Word template for first submission (.docx)

Click here to download LaTeX template for first submission (.tex)

Revised Manuscripts

Revised manuscripts should be submitted via IJOCTA online system to ensure that they are linked to the original submission. It is also necessary to attach a separate file in which a point-by-point explanation is given to the specific points/questions raised by the referees and the corresponding changes made in the revised version.

To upload your revised manuscript, please go to your author page and click on the related manuscript title. Navigate to the "Review" link on the top left and scroll down the page. Click on the "Choose File" button under the "Editor Decision" title, choose the revised article (in pdf format) that you want to submit, and click on the "Upload" button to upload the author version. Repeat the same steps to upload the "Responses to Reviewers/Editor" file and make sure that you click the "Upload" button again.

To avoid any delay in making the article available freely online, the authors also need to upload the source files (Word or LaTex) when submitting revised manuscripts. Files can be compressed if necessary. The two-column final submission templates are as follows:

Click here to download Word template for final submission (.docx)

Click here to download LaTeX template for final submission (.tex)

Authors are responsible for obtaining permission to reproduce copyrighted material from other sources and are required to sign an agreement for the transfer of copyright to IJOCTA.

Article Processing Charges

There are no charges for submission and/or publication.

English Editing

Papers must be in English. Both British and American spelling is acceptable, provided usage is consistent within the manuscript. Manuscripts that are written in English that is ambiguous or incomprehensible, in the opinion of the Editor, will be returned to the authors with a request to resubmit once the language issues have been improved. This policy does not imply that all papers must be written in "perfect" English, whatever that may mean. Rather, the criteria require that the intended meaning of the authors must be clearly understandable, i.e., not obscured by language problems, by referees who have agreed to review the paper.

Presentation of Papers

Manuscript style

Use a standard font of the **11-point type: Times New Roman** is preferred. It is necessary to single line space your manuscript. Normally manuscripts are expected not to exceed 25 single-spaced pages including text, tables, figures and bibliography. All illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end.

During the submission process you must enter: (1) the full title, (2) names and affiliations of all authors and (3) the full address, including email, telephone and fax of the author who is to check the proofs. Supply a brief **biography** of each author at the end of the manuscript after references.

- Include the name(s) of any **sponsor(s)** of the research contained in the paper, along with **grant number(s).**
- Enter an **abstract** of no more than 250 words for all articles.

Keywords

Authors should prepare no more than 5 keywords for their manuscript.

Maximum five **AMS Classification number** (http://www.ams.org/mathscinet/msc/msc2010.html) of the study should be specified after keywords.

Writing Abstract

An abstract is a concise summary of the whole paper, not just the conclusions. The abstract should be no more than 250 words and convey the following:

1. An introduction to the work. This should be accessible by scientists in any field and express the necessity of the experiments executed.
2. Some scientific detail regarding the background to the problem.
3. A summary of the main result.
4. The implications of the result.
5. A broader perspective of the results, once again understandable across scientific disciplines.

It is crucial that the abstract conveys the importance of the work and be understandable without reference to the rest of the manuscript to a multidisciplinary audience. Abstracts should not contain any citation to other published works.

Reference Style

Reference citations in the text should be identified by numbers in square brackets "[ ]". All references must be complete and accurate. Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Online citations should include date of access. References should be listed in the following style:

Journal article

Author, A.A., & Author, B. (Year). Title of article. Title of Journal, Vol(Issue), pages.

Castles, F.G., Curtin, J.C., & Vowles, J. (2006). Public policy in Australia and New Zealand: The new global context. Australian Journal of Political Science, 41(2), 131–143.

Book

Author, A. (Year). Title of book. Publisher, Place of Publication.

Mercer, P.A., & Smith, G. (1993). Private Viewdata in the UK. 2nd ed. Longman, London.

Chapter

Author, A. (Year). Title of chapter. In: A. Editor and B. Editor, eds. Title of book. Publisher, Place of publication, pages.

Bantz, C.R. (1995). Social dimensions of software development. In: J.A. Anderson, ed. Annual review of software management and development. CA: Sage, Newbury Park, 502–510.

Internet document

Author, A. (Year). Title of document [online]. Source. Available from: URL [Accessed (date)].

Holland, M. (2004). Guide to citing Internet sources [online]. Poole, Bournemouth University. Available from: http://www.bournemouth.ac.uk/library/using/guide_to_citing_internet_sourc.html [Accessed 4 November 2004].

Newspaper article

Author, A. (or Title of Newspaper) (Year). Title of article. Title of Newspaper, day Month, page, column.

Independent (1992). Picking up the bills. Independent, 4 June, p. 28a.

Thesis

Author, A. (Year). Title of thesis. Type of thesis (degree). Name of University.

Agutter, A.J. (1995). The linguistic significance of current British slang. PhD Thesis. Edinburgh University.

Illustrations

Illustrations submitted (line drawings, halftones, photos, photomicrographs, etc.) should be clean originals or digital files. Digital files are recommended for highest quality reproduction and should follow these guidelines:

- 300 dpi or higher
- Sized to fit on journal page
- TIFF or JPEG format only
- Embedded in text files and submitted as separate files (if required)

Tables and Figures

Tables and figures (illustrations) should be embedded in the text at the appropriate points, rather than at the end. A short descriptive title should appear above each table with a clear legend and any footnotes suitably identified below.

Proofs

Page proofs are sent to the designated author using IJOCTA EProof system. They must be carefully checked and returned within 48 hours of receipt.

Offprints/Reprints

Each corresponding author of an article will receive a PDF file of the article via email. This file is for personal use only and may not be copied and disseminated in any form without prior written permission from IJOCTA.

Submission Preparation Checklist

As part of the submission process, authors are required to check off their submission's compliance with all of the following items, and submissions may be returned to authors that do not adhere to these guidelines.

1. The submission has not been previously published, nor is it before another journal for consideration (or an explanation has been provided in Comments for the Editor).
2. The submission file is in Portable Document Format (.pdf).
3. Where available, URLs for the references have been provided.
4. The text is single line spaced; uses a 11-point font; employs italics, rather than underlining (except with URL addresses); and all illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end.
5. The text adheres to the stylistic and bibliographic requirements outlined in the Author Guidelines, which is found in "About the Journal".
6. Maximum five AMS Classification number (http://www.ams.org/mathscinet/msc/msc2010.html) of the study have been provided after keywords.
7. After the acceptance of manuscript (before copy editing), Word (.docx) or LaTeX (.tex) version of the paper will be presented.
8. The names and email addresses of at least four (4) possible reviewers have been indicated in "Comments for the Editor" box in Paper Submission Step 1. Please note that at least two of the recommendations should be from different countries. Avoid suggesting reviewers who are at arms-length from you or your co-authors. This includes graduate advisors, people in your current department, or any others with a conflict of interest.

Peer Review Process

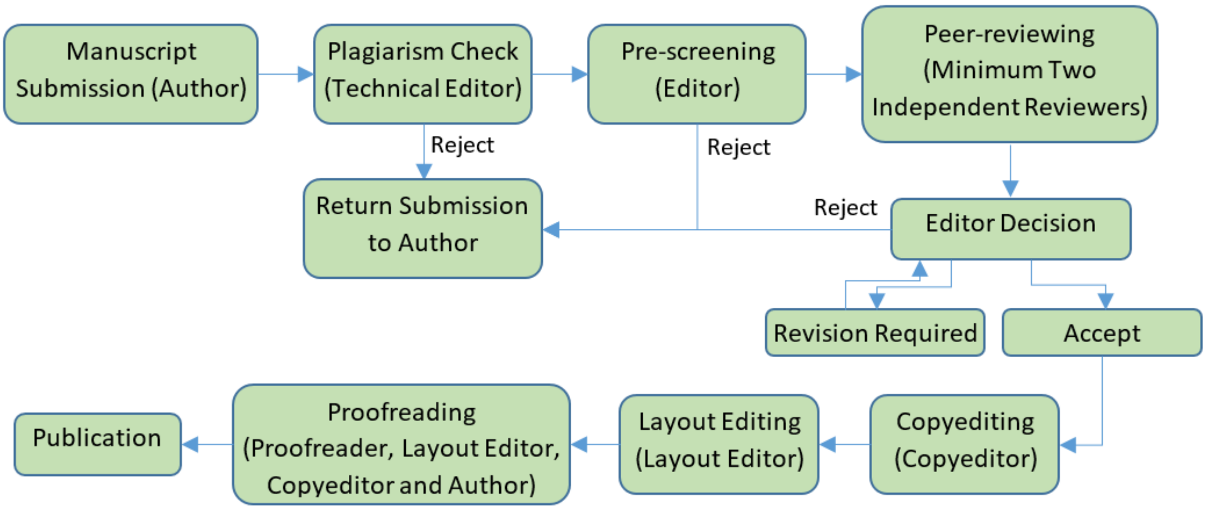
All contributions, prepared according to the author guidelines and submitted via IJOCTA online submission system are evaluated according to the criteria of originality and quality of their scientific content. The corresponding author will receive a confirmation e-mail with a reference number assigned to the paper, which he/she is asked to quote in all subsequent correspondence.

All manuscripts are first checked by the Technical Editor using plagiarism detection software (iThenticate) to verify originality and ensure the quality of the written work. If the result is not satisfactory (i.e. exceeding the limit of 30% of overlapping), the submission is rejected and the author is notified.

After the plagiarism check, the manuscripts are evaluated by the Editor-in-Chief and can be rejected without reviewing if considered not of sufficient interest or novelty, too preliminary or out of the scope of the journal. If the manuscript is considered suitable for further evaluation, it is first sent to the Area Editor. Based on his/her opinion the paper is then sent to at least two independent reviewers. Each reviewer is allowed up to four weeks to return his/her feedback but this duration may be extended based on his/her availability. IJOCTA has instituted a blind peer review process where the reviewers' identities are not known to authors. When the reviews are received, the Area Editor gives a decision and lets the author know it together with the reviewer comments and any supplementary files.

Should the reviews be positive, the authors are expected to submit the revised version usually within two months the editor decision is sent (this period can be extended when the authors contact to the editor and let him/her know that they need extra time for resubmission). If a revised paper is not resubmitted within the deadline, it is considered as a new submission after all the changes requested by reviewers have been made. Authors are required to submit a new cover letter, a response to reviewers letter and the revised manuscript (which ideally shows the revisions made in a different color or highlighted). If a change in authorship (addition or removal of author) has occurred during the revision, authors are requested to clarify the reason for change, and all authors (including the removed/added ones) need to submit a written consent for the change. The revised version is evaluated by the Area editor and/or reviewers and the Editor-in-Chief brings a decision about final acceptance based on their suggestions. If necessary, further revision can be asked for to fulfil all the requirements of the reviewers.

When a manuscript is accepted for publication, an acceptance letter is sent to the corresponding author and the authors are asked to submit the source file of the manuscript conforming to the IJOCTA two-column final submission template. After that stage, changes of authors of the manuscript are not possible. The manuscript is sent to the Copyeditor and a linguistic, metrological and technical revision is made, at which stage the authors are asked to make the final corrections in no more than a week. The layout editor prepares the galleys and the authors receive the galley proof for final check before printing. The authors are expected to correct only typographical errors on the proofs and return the proofs within 48 hours. After the final check by the layout editor and the proofreader, the manuscript is assigned a DOI number, made publicly available and listed in the forthcoming journal issue. After printing the issue, the corresponding metadata and files published in this issue are sent to the databases for indexing.



Publication Ethics and Malpractice Statement

IJOCTA is committed to ensuring ethics in publication and quality of articles. Conforming to standards of expected ethical behavior is therefore necessary for all parties (the author, the editor(s), the peer reviewer) involved in the act of publishing.

International Standards for Editors

The editors of the IJOCTA are responsible for deciding which of the articles submitted to the journal should be published considering their intellectual content without regard to race, gender, sexual orientation, religious belief, ethnic origin, citizenship, or political philosophy of the authors. The editors may be guided by the policies of the journal's editorial board and constrained by such legal requirements as shall then be in force regarding libel, copyright infringement and plagiarism. The editors may confer with other editors or reviewers in making this decision. As guardians and stewards of the research record, editors should encourage authors to strive for, and adhere themselves to, the highest standards of publication ethics. Furthermore, editors are in a unique position to indirectly foster responsible conduct of research through their policies and processes.

To achieve the maximum effect within the research community, ideally all editors should adhere to universal standards and good practices.

- Editors are accountable and should take responsibility for everything they publish.
- Editors should make fair and unbiased decisions independent from commercial consideration and ensure a fair and appropriate peer review process.
- Editors should adopt editorial policies that encourage maximum transparency and complete, honest reporting.
- Editors should guard the integrity of the published record by issuing corrections and retractions when needed and pursuing suspected or alleged research and publication misconduct.
- Editors should pursue reviewer and editorial misconduct.
- Editors should critically assess the ethical conduct of studies in humans and animals.
- Peer reviewers and authors should be told what is expected of them.
- Editors should have appropriate policies in place for handling editorial conflicts of interest.

Reference:

Kleinert S & Wager E (2011). Responsible research publication: international standards for editors. A position statement developed at the 2nd World Conference on Research Integrity, Singapore, July 22-24, 2010. Chapter 51 in: Mayer T & Steneck N (eds) Promoting Research Integrity in a Global Environment. Imperial College Press / World Scientific Publishing, Singapore (pp 317-28). (ISBN 978-981-4340-97-7) [Link].

International Standards for Authors

Publication is the final stage of research and therefore a responsibility for all researchers. Scholarly publications are expected to provide a detailed and permanent record of research. Because publications form the basis for both new research and the application of findings, they can affect not only the research community but also, indirectly, society at large. Researchers therefore have a responsibility to ensure that their publications are honest, clear, accurate, complete and balanced, and should avoid misleading, selective or ambiguous reporting. Journal editors also have responsibilities for ensuring the integrity of the research literature and these are set out in companion guidelines.

- The research being reported should have been conducted in an ethical and responsible manner and should comply with all relevant legislation.
- Researchers should present their results clearly, honestly, and without fabrication, falsification or inappropriate data manipulation.
- Researchers should strive to describe their methods clearly and unambiguously so that their findings can be confirmed by others.
- Researchers should adhere to publication requirements that submitted work is original, is not plagiarised, and has not been published elsewhere.
- Authors should take collective responsibility for submitted and published work.
- The authorship of research publications should accurately reflect individuals' contributions to the work and its reporting.

- Funding sources and relevant conflicts of interest should be disclosed.
- When an author discovers a significant error or inaccuracy in his/her own published work, it is the author's obligation to promptly notify the journal's Editor-in-Chief and cooperate with them to either retract the paper or to publish an appropriate erratum.

Reference:

Wager E & Kleinert S (2011) Responsible research publication: international standards for authors. A position statement developed at the 2nd World Conference on Research Integrity, Singapore, July 22-24, 2010. Chapter 50 in: Mayer T & Steneck N (eds) Promoting Research Integrity in a Global Environment. Imperial College Press / World Scientific Publishing, Singapore (pp 309-16). (ISBN 978-981-4340-97-7) [Link].

Basic principles to which peer reviewers should adhere

Peer review in all its forms plays an important role in ensuring the integrity of the scholarly record. The process depends to a large extent on trust and requires that everyone involved behaves responsibly and ethically. Peer reviewers play a central and critical part in the peer-review process as the peer review assists the Editors in making editorial decisions and, through the editorial communication with the author, may also assist the author in improving the manuscript.

Peer reviewers should:

- respect the confidentiality of peer review and not reveal any details of a manuscript or its review, during or after the peer-review process, beyond those that are released by the journal;
- not use information obtained during the peer-review process for their own or any other person's or organization's advantage, or to disadvantage or discredit others;
- only agree to review manuscripts for which they have the subject expertise required to carry out a proper assessment and which they can assess within a reasonable time-frame;
- declare all potential conflicting interests, seeking advice from the journal if they are unsure whether something constitutes a relevant conflict;
- not allow their reviews to be influenced by the origins of a manuscript, by the nationality, religion, political beliefs, gender or other characteristics of the authors, or by commercial considerations;
- be objective and constructive in their reviews, refraining from being hostile or inflammatory and from making libellous or derogatory personal comments;
- acknowledge that peer review is largely a reciprocal endeavour and undertake to carry out their fair share of reviewing, in a timely manner;
- provide personal and professional information that is accurate and a true representation of their expertise when creating or updating journal accounts.

Reference:

Homes I (2013). COPE Ethical Guidelines for Peer Reviewers, March 2013, v1 [Link].

Copyright Notice

An International Journal of Optimization and Control: Theories & Applications

Volume: 8 Number: 1
January 2018

CONTENTS

ISSN 2146-0957

www.ijocta.com
info@ijocta.com