Iranian Journal of Numerical Analysis and Optimization

Volume 8 , Number 1
Winter 2018
Serial Number: 13

Ferdowsi University of Mashhad, Iran

In the Name of God

Iranian Journal of Numerical Analysis and Optimization (IJNAO)

This journal is authorized under the registration No. 174/853 dated 1386/2/26, by the Ministry of Culture and Islamic Guidance.

Volume 8, Number 1, Winter 2018

ISSN: 2423-6977

Publisher: Faculty of Mathematical Sciences, Ferdowsi University of Mashhad

Published by: Ferdowsi University of Mashhad Press

Circulation: 100

Address: Iranian Journal of Numerical Analysis and Optimization

Faculty of Mathematical Sciences, Ferdowsi University of Mashhad

P.O. Box 1159, Mashhad 91775, Iran.

Tel. : +98-51-38806222 , **Fax:** +98-51-38807358

E-mail: mjms@um.ac.ir

Website: http://ijnao.um.ac.ir

This journal is indexed by:

- Zentralblatt
- ISC
- SID

به اطلاع کلیه محققان، پژوهشگران، اساتید ارجمند، دانشجویان تحصیلات تکمیلی و نویسندگان محترم می رساند که نشریه ایرانی آنالیز عددی و بهینه سازی - IJNAO - طبق مجوز شماره ۳/۱۸/۵۴۸۹۱۳/ مورخه ۱۳۹۲/۱۰/۲۵ مدیر کل محترم سیاست گذاری و برنامه ریزی امور پژوهشی وزارت علوم، تحقیقات و فناوری، علمی – پژوهشی میباشد.
بر اساس نامه شماره ۹۲/۱۵۳۶/پ مورخه ۱۳۹۲/۱۱/۲۱ سرپرست محترم معاونت پژوهشی وفناوری پایگاه استنادی علوم جهان اسلام، نشریه ایرانی آنالیز عددی و بهینه سازی در پایگاه ISC نیز نمایه می شود.

Iranian Journal of Numerical Analysis and Optimization

Volume 8, Number 1, Winter 2018

Ferdowsi University of Mashhad - Iran

Iranian Journal of Numerical Analysis and Optimization

This journal is published under the auspices of Ferdowsi University of Mashhad

* Full Professor
** Associate Professor

We would like to acknowledge the help of Narjes khatoon Zohorian in the preparation of this issue.

Letter from the Editor in Chief

I would like to welcome you to the Iranian Journal of Numerical Analysis and Optimization (IJNAO). This journal is published biannually and supported by the Faculty of Mathematical Sciences at the Ferdowsi University of Mashhad. Faculty of Mathematical Sciences with three centers of excellence and three research centers is well-known in mathematical communities in Iran.

The main aim of the journal is to facilitate discussions and collaborations between specialists in applied mathematics, especially in the fields of numerical analysis and optimization, in the region and worldwide.

Our vision is that scholars from different applied mathematical research disciplines, pool their insight, knowledge and efforts by communicating via this international journal.

In order to assure high quality of the journal, each article is reviewed by subject-qualified referees.

Our expectations for IJNAO are as high as any well-known applied mathematical journal in the world. We trust that by publishing quality research and creative work, the possibility of more collaborations between researchers would be provided. We invite all applied mathematicians especially in the fields of numerical analysis and optimization to join us by submitting their original work to the Iranian Journal of Numerical Analysis and Optimization.

Unfortunately, I need to announce that Professor H. R. Tareghian (1956-2018), the Editor in Charge of IJNAO, left us in January 2018. He was professor in Applied Mathematics in the Faculty of Mathematical Sciences, Ferdowsi University of Mashhad. Losing a friend is one of the most painful experiences must endure in life. It is hard for me and editors of IJNAO to accept that professor Tareghian , that we shared so many moments with him and trusted with so many things, is no longer with us. We really missed him.

Mohammad Hadi Farahi

Contents

Iranian Journal of Numerical Analysis and Optimization
Vol. 8, No. 1, (2018), pp 1–17
DOI:10.22067/ijnao.v8i1.60791

A computational method for solving weakly singular Fredholm integral equation in reproducing kernel spaces

D. Hamedzadeh and E. Babolian*

Abstract

In the present paper, we propose a method to solve a class of weakly singular Fredholm integral equations of the second kind in reproducing kernel spaces. The Taylor series of the unknown function is used to remove the singularity and bases of reproducing kernel spaces are used to solve this equation. Efficiency of the proposed method is investigated through various examples.

Keywords: Weakly singular kernel; Fredholm integral equations; Taylor series; Reproducing kernel space.

1 Introduction

The Fredholm integral equation with weakly singular kernel arises in different problems of mathematical physics such as potential problem, variational equilibrium, fracture mechanics, infrared radiation, and elastic contact problems [5,20,21]. Although in [11,16,17,22] authors obtained analytical solution for integral equations with weakly singular kernel in special cases, but generally this is not an easy task. So the numerical analysis standpoint plays major role for solving such equations.

In this paper, we consider Fredholm singular integral equation

---

*Corresponding author
Received 5 December 2016; revised 23 March 2017; accepted 14 June 2017
D. Hamedzadeh
Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran e-mail: d.hamedzadeh@srbiau.ac.ir and danial_hamedzadeh2010@yahoo.com

E. Babolian
Department of Mathematics, Science And Research Branch, Islamic Azad University, Tehran, Iran e-mail: babolian@khu.ac.ir

$$\mu(x)u(x) + \lambda(x)\int_{-1}^{1} \frac{k(x,y)u(y)}{(y-x)^{\alpha}}dy = f(x), \qquad |x| < 1 \text{ and } 0 < \alpha \leq 1, \quad (1)$$

in which $\mu(x) \neq 0$ and $\lambda(x) \neq 0$, for all $x \in [-1,1]$, and μ, λ, f, and k are known smooth functions and u is the solution of equation (1) to be determined. We assume that problem has a unique smooth solution u. Several authors considered the numerical solutions of Fredholm integral equations with weakly singular kernel. In [12] Jiang and Cui considered integral equation of first or third kind with weakly singular kernel of the form $k(x,y) = \frac{G(x,y)}{x^{\alpha}y^{\beta}}$; they solved the problem in reproducing kernel space $W^1[0,1]$. Chen and Zhou considered second kind Fredholm integral equation with Hilbert type singularity [8]. They used transform to remove singularity and solved problem in $W[0,2\pi]$ using reproducing kernel method. In [2] Babolian and Arzhang Hajikandi solved (1) with $k(x,y) = 1$. Du, Zhao, G., and Zhao, C. considered integro-differential equation with logarithmic kernel and Kalman kernel with boundary values [10]. They used smooth transform to remove singularity, solving the converted equation with reproducing kernel method in $W^3[0,1]$. Chen and Cheng in [7] used piecewise homotopy perturbation method (PHPM), for solving integro-differential equation with weakly singular kernel. In [3] authors solved (1), using Taylor series of the unknown function u to remove singularity, and then Taylor expansion of k together with Legendre polynomials as bases to implement Galerkin method. The Sinc-collocation method is studied by Maleknejad, Mollapourasl, and Ostadi, to solve nonlinear Fredholm integral equations with weakly singular kernel [15]. Beyrami, Lotfi, and Mahdiani solved Fredholm integral equation of the second kind with Cauchy kernel [6]; they removed singularity by smooth transform and used reproducing kernel Hilbert space (RKHS) method to solve problem in $W_o^3[0,1]$. Nili and Dastjerdi in [18] solved weakly singular Volterra–Hammerstein integral equation with operational Tau method. We use Taylor series expansion of $u(y)$, at point $y = x$, to remove singularity; then by use of reproducing kernel space $W^m[-1,1]$ we convert this equation to a system of linear equations. We demonstrate the method with convergence rate $O(h^m)$; so when we increase m and move from one reproducing kernel space to other, the rate of convergence will increase.

The rest of the paper is organized as follows. In section 2 we present definitions and some useful properties of reproducing kernel spaces. In section 3 we remove singularity, by use of reproducing kernel space to implement our method. Section 4 is devoted to error estimate and convergence analysis of our method. Section 5 contains some numerical examples illustrating the application of the proposed method. We end the paper with some conclusions.

2 Preliminaries and notations

In this section we briefly review reproducing kernel properties of the Hilbert space $W^m[a,b]$ and also fix notations used in this paper. The Hilbert function space, $W^m[a,b]$, is defined as the linear function space

$$W^m[a,b] = \{f | f, f^{(\prime)}, \ldots, f^{(m-1)} \text{are absolutely continuous}, f^{(m)} \in L^2[a,b]\},$$

which is equipped with the following inner product

$$\langle f, g \rangle_{W^m} = \sum_{i=0}^{m-1} f^{(i)}(a) g^{(i)}(a) + \int_a^b f^{(m)}(x) g^{(m)}(x) dx. \tag{2}$$

The inner product (2) induces the following Hilbert norm

$$\|f\|_{W^m} = \sqrt{\langle f, f \rangle_{W^m}}.$$

The following theorem presents an interesting property of the Hilbert space $W^m[a,b]$.

Theorem 1. [9] *The Hilbert function space $W^m[a,b]$ is a reproducing kernel space with the conjugate symmetric reproducing kernel $R_m(x,y)$, that is given by*

$$R_m(x,y) = \begin{cases} lR_m(x,y) = \sum_{i=1}^{2m} c_i(y) x^{i-1}, & x < y, \\ \\ rR_m(x,y) = \sum_{i=1}^{2m} d_i(y) x^{i-1}, & x \geq y, \end{cases}$$

in which coefficients $c_i(y)$ and $d_i(y)$ are the solutions of the following system of differential equations

$$\begin{cases} (-1)^m \frac{\partial^{2m} R(x,y)}{\partial x^{2m}} = \delta(x-y), \\ \frac{\partial^i R(a,y)}{\partial x^i} - (-1)^{m-i-1} \frac{\partial^{2m-i-1} R(a,y)}{\partial x^{2m-i-1}} = 0, \\ \frac{\partial^{2m-i-1} R(b,y)}{\partial x^{2m-i-1}} = 0, \qquad i = 0, 1, \ldots, m-1, \end{cases} \tag{3}$$

where $\delta(x-y)$ is the Dirac delta function.

Remark 1. [9] According to system (3), by use of the Dirac delta function properties, the equation

$$(-1)^m \frac{\partial^{2m} R(x,y)}{\partial x^{2m}} = \delta(x-y)$$

is convert to

$$\begin{cases} \left.\frac{\partial^i \ lR_m(x,y)}{\partial x^i}\right|_{x=y} = \left.\frac{\partial^i \ rR_m(x,y)}{\partial x^i}\right|_{x=y}, \qquad i=0,1,\ldots,2m-2, \\ \left.\frac{\partial^{2m-1} \ lR_m(x,y)}{\partial x^{2m-1}}\right|_{x=y^+} - \left.\frac{\partial^{2m-1} \ rR_m(x,y)}{\partial x^{2m-1}}\right|_{x=y^-} = (-1)^m. \end{cases} \tag{4}$$

By solving the system (4) with boundary condition of (3) with Mathematica, the coefficients $c_i(y)$ and $d_i(y)$ are computed.

In particular, each function $f \in W^m[a,b]$ satisfies the following reproducing property

$$f(y) = \langle f(.), R_m(.,y)\rangle_{W^m} \qquad \forall y \in [a,b].$$

Theorem 2. [9]*Let $R(x,y)$ be a reproducing kernel of $W^m[a,b]$; then*

$$\frac{\partial^{i+j}R(x,y)}{\partial x^i \partial y^j} \in L^2[a,b], \qquad i+j=2m-1$$

with respect to x and y and

$$\frac{\partial^{i+j}R(x,y)}{\partial x^i \partial y^j}, \qquad 0 \le i+j \le 2m-2,$$

are absolutely continuous functions in $[a,b]$, with respect to x and y; so we have

$$\frac{\partial^{i+j}R(x,y)}{\partial x^i \partial y^j} \in W^m[a,b], \qquad i+j=m-1.$$

Lemma 1. *Let f be a smooth function of order m on $[a,b]$; then, for $k = 1,2,\ldots,m$, we have $f \in W^k[a,b]$.*

Proof. Let f be a smooth function of order m on $[a,b]$; thus $f^{(k)}$, for $k = 0,1,\ldots,m$, are continuous functions on $[a,b]$. Therefore $f^{(k)} \in L^2[a,b]$, for $k=1,2,\ldots,m$, beside the real constants M_k exist which we have

$$|f^{(k)}(x)| \le \ M_k \qquad \forall x \in [a,b];$$

trivially $f^{(k)}$, for $k=0,1,\ldots,m-1$, are absolutely continuous functions, and proof is complete. □

3 Removing the singularity and implementation of our method

Throughout this section, we use the Taylor series to remove singularity of (1); then by use of reproducing kernel space property, we construct our method. With Taylor series of the unknown function $u(y)$, at $x \in [a, b]$, we have

$$\begin{aligned} u(y) = u(x) + (y-x)u^{'}(x) + \frac{(y-x)^2}{2!}u^{''}(x) + \cdots + \frac{(y-x)^k}{k!}u^{(k)}(x) \\ + \frac{(y-x)^{k+1}}{(k+1)!}u^{(k+1)}(\xi_{x,y}), \end{aligned} \tag{5}$$

in which $\xi_{x,y}$ is between x and y. Using (5), thus equation (1) converts to

$$\begin{aligned} \mu(x)u(x) + \lambda(x)\Bigg(\sum_{i=0}^{k}\frac{u^{(i)}(x)}{i!}\int_{-1}^{1}k(x,y)(y-x)^{i-\alpha}dy \\ + \int_{-1}^{1}\frac{u^{(k+1)}(\xi_{x,y})}{(k+1)!}k(x,y)(y-x)^{k+1-\alpha}dy\Bigg) = f(x), \end{aligned}$$

and singularity is removed; now we are in position to define operator $L : W^m[-1,1] \to W^m[-1,1]$,

$$\begin{aligned} L(u(.)) = \mu(.)u(.) + \lambda(.)\Bigg(\sum_{i=0}^{k}\frac{u^{(i)}(.)}{i!}\int_{-1}^{1}k(.,y)(y-.)^{i-\alpha}dy \\ + \int_{-1}^{1}\frac{u^{(k+1)}(\xi_{.,y})}{(k+1)!}k(.,y)(y-.)^{k+1-\alpha}dy\Bigg). \end{aligned}$$

By use of operator L, the equation (1) is convert to $Lu(x) = f(x)$. It is easy to prove that, for $0 \leq \alpha < 1$, operator L is a bounded linear operator on $[-1,1]$ and that, for $\alpha = 1$, L is a bounded linear operator on $[-1+\epsilon, 1-\epsilon]$ for $\epsilon > 0$. Let $\{x_i\}_{i=1}^{\infty}$ be a dense subset of interval $[a,b]$, for $i = 1, 2, \ldots$; take $y = x_i$, in $R_m(x,y)$, and define $\varphi_i(.) = R_m(., x_i)$. Moreover, assume that $\psi_i = L^*\varphi_i$, where L^* is the adjoint operator of L.

Theorem 3. *Let $\{x_i\}_{i=1}^{\infty}$ be a dense set in $[-1,1]$; then the functions $\psi_i \in W^m[-1,1]$, and $\{\psi_i\}_{i=1}^{\infty}$ is complete in $W^m[-1,1]$.*

Proof. By using the reproducing kernel properties of $W^m[-1,1]$, we have

$$\begin{aligned}\psi_i = L^*\varphi_i &= \langle L^*\varphi_i(y), R_m(y,.)\rangle_{W^m}\\ &= \langle \varphi_i(y), L_y R_m(y,.)\rangle_{W^m}\\ &= \langle \varphi_i(y), \overline{L_y R_m(.,y)}\rangle_{W^m}\\ &= \overline{\langle \overline{L_y R_m(.,y)}, \varphi_i(y)\rangle_{W^m}} = L_y R_m(.,y)\Big|_{y=x_i}.\end{aligned}$$

Therefore, we get

$$\psi_i = \mu(x_i)R_m(.,x_i) + \lambda(x_i)\sum_{i=0}^{2m-1}\frac{1}{i!}\,\frac{\partial^i R_m(.,s)}{\partial s^i}\Big|_{s=x_i}\int_{-1}^{1} k(x_i,s)(x_i-s)^{i-\alpha}ds;$$

hence $\psi_i \in W^m[-1,1]$. Let $u \in W^m[-1,1]$, and $i \in \mathbb{N}$. Then we have

$$\begin{aligned}\langle u,\psi_i\rangle_{W^m} &= \langle u, L^*\varphi_i\rangle_{W^m}\\ &= \langle Lu,\varphi_i\rangle_{W^m}\\ &= Lu(x_i) = 0.\end{aligned}$$

Since $\{x_i\}_{i=1}^{\infty}$ is dense in $[-1,1]$, we have $Lu = 0$. By the uniqueness of the solution of (1), we get $u = 0$. Hence, $\{\psi_i\}_{i=1}^{\infty}$ is complete in $W^m[-1,1]$. □

As a consequence of Theorem 3, we can expand each function as the following. Let $\{x_i\}_{i=1}^{\infty}$ be dense in $[-1,1]$, and let u be the solution of (1). Since $u \in W^m[-1,1]$, we get

$$u(x) = \sum_{i=0}^{\infty} a_i\psi_i(x).$$

In what follows, we use u_n as approximate solution of equation (1) which is defined by

$$u_{m,n}(x) = \sum_{i=0}^{n} a_i\psi_i(x), \qquad x \in [-1,1],$$

and by use of operator L, we get

$$\sum_{i=0}^{n} a_i\; L\psi_i(x) \simeq f(x), \tag{6}$$

and with inner product of equation (6), with φ_j, we have

$$\sum_{i=0}^{n} a_i\;\langle L\psi_i,\varphi_j\rangle_{W^m} \simeq \langle f,\varphi_j\rangle_{W^m}.$$

Thus by replacing $\simeq$ by $=$, the singular integral problem converts to the linear system of equations $Ba = F$, in which the coefficient matrix $B = [b_{i,j}] = [L\psi_i(x_j)]$, and the right hand side vector is $F = [F_j] = f(x_j)$.

4 Error estimate and convergence analysis

In this section first the convergence of the method is investigated, and then the error estimate of proposed method is studied.

Theorem 4. *Let $\{x_i\}_{i=1}^{\infty}$ be dense in $[-1,1]$, and let $\{\overline{\psi_i}\}_{i=1}^{\infty}$ be the orthonormal basis of $W^m[-1,1]$ produced by Gram–Schmidt process on $\{\psi_i\}_{i=1}^{\infty}$. Then $u_{m,n}$, the approximate solution of equation (1), and its derivatives are convergent uniformly to u; that is, the exact solution of equation (1), and its derivatives, respetively.*

Proof. With the Gram process we have

$$\psi_i(x) = \sum_{k=1}^{i} \beta_{ik}\overline{\psi_k(x)},$$

and hence we get

$$u(x) = \sum_{i=0}^{\infty}\sum_{k=1}^{i} a_i\beta_{ik}\overline{\psi_k(x)}.$$

Therefore the approximate solution of equation (1) is

$$u_{m,n}(x) = \sum_{i=0}^{n}\sum_{k=1}^{i} a_i\beta_{ik}\overline{\psi_k(x)},$$

and clearly we have $\lim_{n\to\infty}\|u-u_{m,n}\|_{W^m[-1,1]} = 0$. On the other hand, we have

$$u(x) = \langle u(.), R_m(.,x)\rangle = \sum_{i=0}^{m-1} u^{(i)}(1)R_m^{(i)}(1,x) + \int_{-1}^{1} u^{(m)}(z)R_m^{(m)}(z,x)dz.$$

Differentiating j times with respect to x, we get

$$\begin{aligned} u^{(j)}(x) &= \sum_{i=0}^{m-1} u^{(i)}(1)\frac{\partial^j R_m^{(i)}(1,x)}{\partial x^j} + \int_{-1}^{1} u^{(m)}(z)\frac{\partial^j R_m^{(m)}(z,x)}{\partial x^j}dz \\ &= \langle u(y), \frac{\partial^j R_m(y,x)}{\partial x^j}\rangle_{W^m}. \end{aligned}$$

Thus we have

$$\begin{aligned}| u^{(i)}(x) - u^{(i)}_{m,n}(x) | &= \langle u(.) - u_{m,n}(.), \frac{\partial^i R_m(.,x)}{\partial x^i}\rangle_{W^m} \\ &\leq \| u - u_{m,n} \|_{W^m[-1,1]} \| \frac{\partial^i R_m(.,x)}{\partial x^i} \|_{W^m[-1,1]} .\end{aligned}$$

Since $\frac{\partial^i R_m(.,x)}{\partial x^i} \in W^m[-1,1]$ and $\| \frac{\partial^i R_m(.,x)}{\partial x^i} \|_{W^m[-1,1]}$, is continuous with respect to x, on $[-1,1]$. For some $M > 0$, we have

$$\| \frac{\partial^i R_m(.,x)}{\partial x^i} \|_{W^m[-1,1]} \leq M,$$

which completes the proof. □

Let $\{x_i\}_{i=1}^n$ be a subset of $[-1,1]$, and let $-1 < x_1 < x_2 < \cdots < x_n < 1$. We define $h_j := x_{j+1} - x_j$, $j = 1, 2, \ldots, n-1$, and also put $h := \max_{1 \leq j \leq n-1} h_j$.

Theorem 5. [4]*Let $u \in W^m[-1,1]$ be a smooth solution, and let $u_{m,n}$ be the approximate solution of the equation* (1)*. Then*

$$\|u - u_{m,n}\|_\infty \leq ch^m,$$

where c is a constant.

5 Numerical examples

We examine accuracy and convergence of the proposed method through four examples, which indicate efficiency of the method. The errors were defined as

$$E_{m,n} = \|u - u_{m,n}\|_\infty \simeq \max_{1 \leq i \leq n} | u(x_i) - u_{m,n}(x_i) | \qquad m = 1, 2, \ldots.$$

Example 1. [3] Consider the following integral equation

$$u(x) + \int_{-1}^{1} \frac{(1 + xy + 2y^3)u(y)}{(y-x)^{\frac{3}{5}}} dy = f(x),$$

where f is chosen such that

$$u(x) = x^5, \qquad x \in [-1,1],$$

is the exact solution. The numerical results are given in Table 1. In Table 1, for fix n, by increasing m, the absolute errors converge to zero. In Figures 1, for $m = 6, 8, 10$, when $n = 8$ the absolute errors of u and its derivatives are

Table 1: Values of $E_{m,n}$, for Example 1.

m	$n = 8$	$n = 16$	$n = 32$
6	1.72675×10^{-1}	1.34164×10^{-1}	1.06141×10^{-1}
8	9.37476×10^{-4}	3.09832×10^{-4}	2.91592×10^{-5}
10	6.73973×10^{-6}	2.6006×10^{-6}	3.41363×10^{-6}

Table 2: Values of $E_{m,n}$, for Example 2.

m	$n = 8$	$n = 16$	$n = 32$
4	2.21811×10^{-2}	2.89879×10^{-2}	7.6194×10^{-2}
6	4.92002×10^{-4}	7.14302×10^{-4}	1.95649×10^{-3}
8	2.02376×10^{-5}	1.07565×10^{-5}	2.91592×10^{-5}
10	3.13506×10^{-6}	1.2141×10^{-7}	2.02656×10^{-7}

shown.

Example 2. Consider the following integral equation

$$u(x) + (x+2)\int_{-1}^{1} \frac{e^{x+y}u(y)}{(y-x)}dy = f(x),$$

where f is chosen such that

$$u(x) = xe^{x}, \qquad x \in [-1,1],$$

is the exact solution. The numerical results are given in Table 2. In Table 2, for fixed n, by increasing m, the absolute errors converge to zero. In Figures 2, for $m = 6, 8, 10$, when $n = 8$ the approximate solutions and its derivatives have high accuracies.

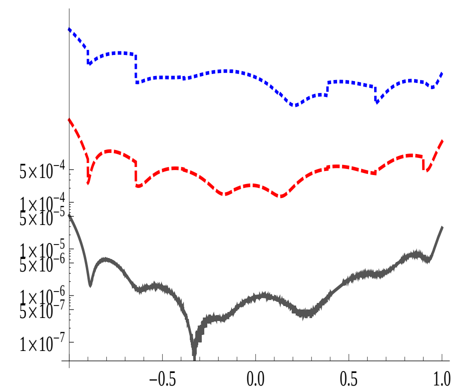
Example 3. $[1, 3, 13, 14]$Consider the following integral equation

$$u(x) - \int_{0}^{1} \frac{u(y)}{\mid x-y \mid^{\frac{1}{2}}}dy = f(x),$$

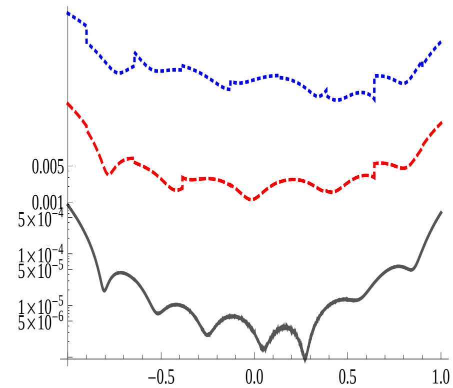
where f is chosen such that

$$u(x) = x, \qquad x \in [0,1]$$

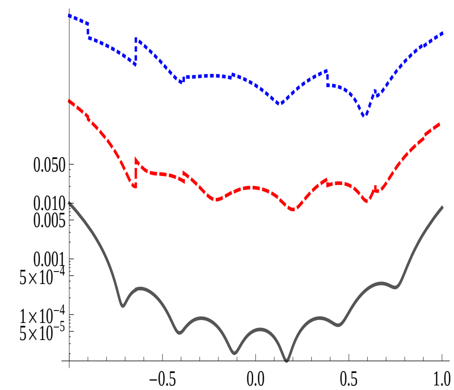
is the exact solution. For $m = 4, 6, 8$, when $n = 8$, the absolute errors of u and its derivatives are shown in Figures 3. In Table 3, the numerical results of proposed method are compared with $[1, 3, 13]$. where the Max Error is



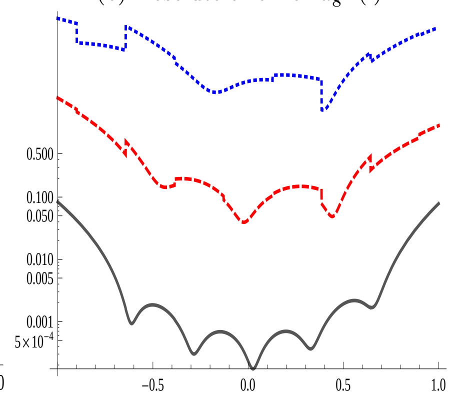
(a) Absolute error for $u_8(.)$.



(b) Absolute error for $u_8^{(1)}(.)$.

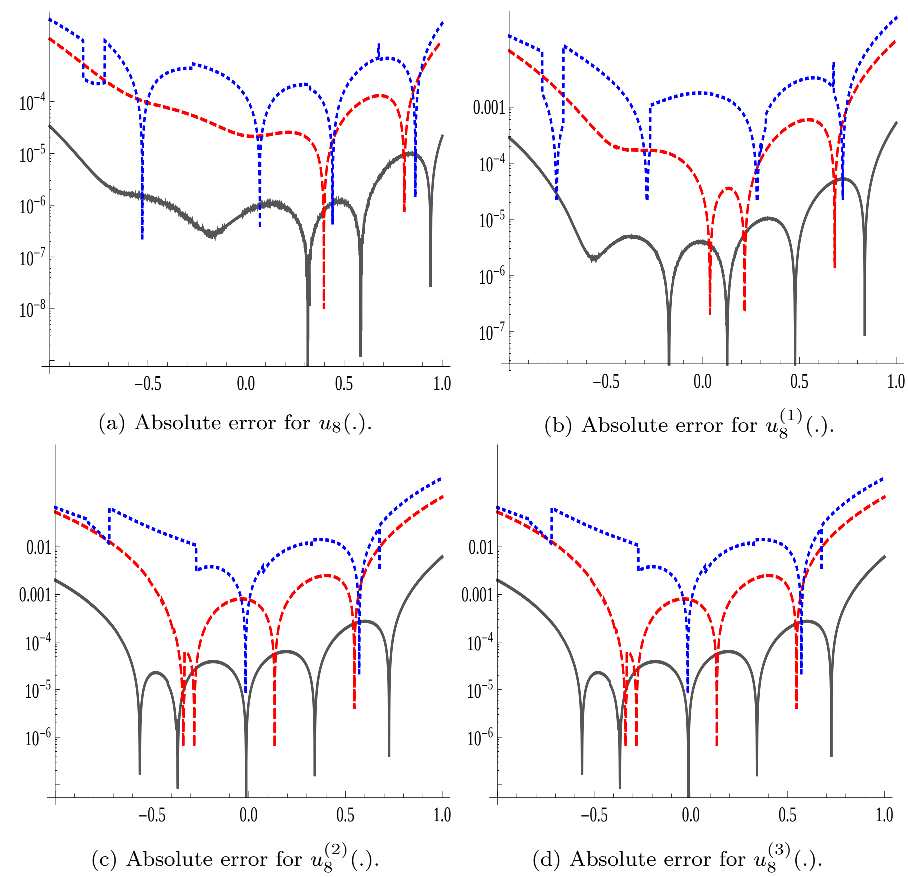


(c) Absolute error for $u_8^{(2)}(.)$.



(d) Absolute error for $u_8^{(3)}(.)$.

Figure 1: Example 1, $m = 6$(blue dotted), $m = 8$(red dash), $m = 10$(Gray line).



(a) Absolute error for $u_8(.)$.

(b) Absolute error for $u_8^{(1)}(.)$.

(c) Absolute error for $u_8^{(2)}(.)$.

(d) Absolute error for $u_8^{(3)}(.)$.

Figure 2: Example 2, $m = 6$(blue dotted), $m = 8$(red dash), $m = 10$(Gray line).

Table 3: The Max Error result of Example 3.

n	$n=16$	$n=32$
Babolian's method [3]	1.59×10^{-16}	
proposed method for $m=10$	5.29781×10^{-11}	4.75211×10^{-11}
Lakestani's method [13]	1.27×10^{-6}	2.71×10^{-8}
Product integration method [1]	1.5×10^{-5}	9.39×10^{-7}
Lagrangian interpolant [1]	2.12×10^{-5}	1.94×10^{-6}

Table 4: The Max Error result of $u(.)$, Example 4.

m	$n=8$	$n=16$	$n=32$
4	3.17985×10^{-4}	4.44036×10^{-4}	4.24008×10^{-4}
6	2.38288×10^{-6}	2.63628×10^{-6}	2.23585×10^{-6}
8	2.93877×10^{-9}	7.58042×10^{-9}	3.42293×10^{-8}

defined as Max Error $=\max_{0\leq x\leq 1} \mid u(x)-u_{m,n}(x)\mid$.

Example 4. [13, 19]Consider the following integral equation

$$u(x)-\frac{1}{10}\int_0^1 \frac{u(y)}{\mid x-y\mid^{\frac{1}{3}}}dy=f(x),$$

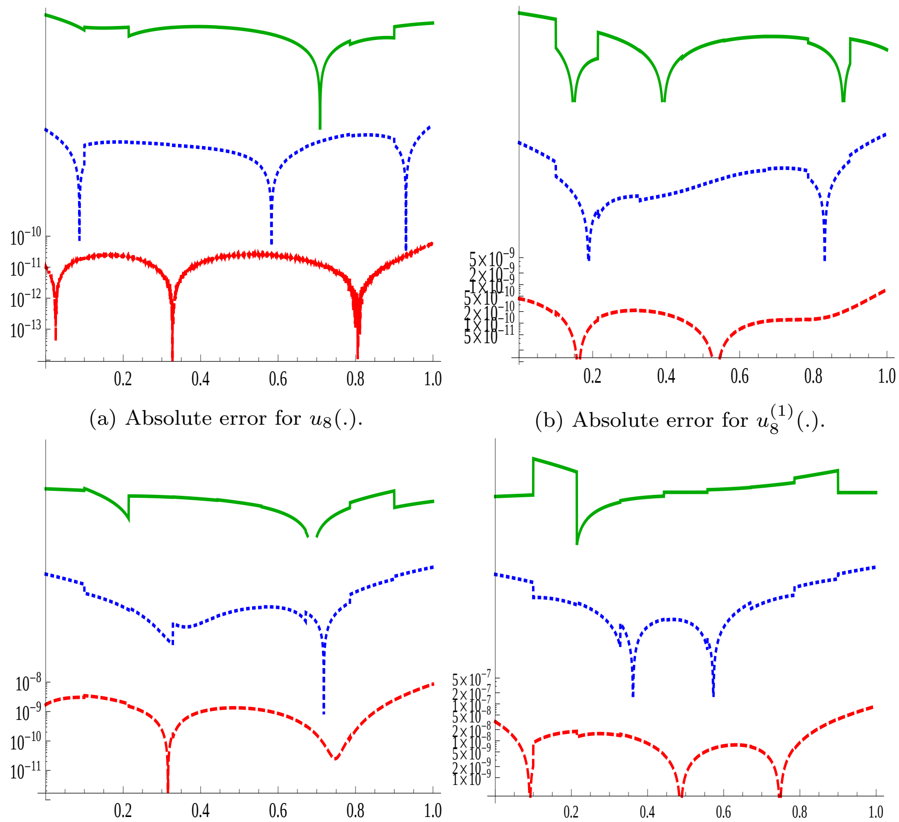
where f is chosen such that

$$u(x)=x^2(1-x^2), \qquad x\in[0,1],$$

is the exact solution. For $m=4,6,8$, when $n=8$, the absolute errors of u and its derivatives are shown in Figures 4. In Tables 4, 5, and 6 the numerical results of proposed method are given. where the Max Error is defined as Max Error $=\max_{0\leq x\leq 1} \mid u^{(i)}(x)-u^{(i)}_{m,n}(x)\mid$.

Table 5: The Max Error result of $u^{(1)}(.)$, Example 4.

m	$n=8$	$n=16$	$n=32$
4	2.10369×10^{-1}	7.59437×10^{-3}	7.61556×10^{-2}
6	8.89262×10^{-6}	2.43063×10^{-6}	1.37019×10^{-6}
8	3.66355×10^{-7}	2.21809×10^{-7}	1.20238×10^{-7}



(a) Absolute error for $u_8(.)$.

(b) Absolute error for $u_8^{(1)}(.)$.

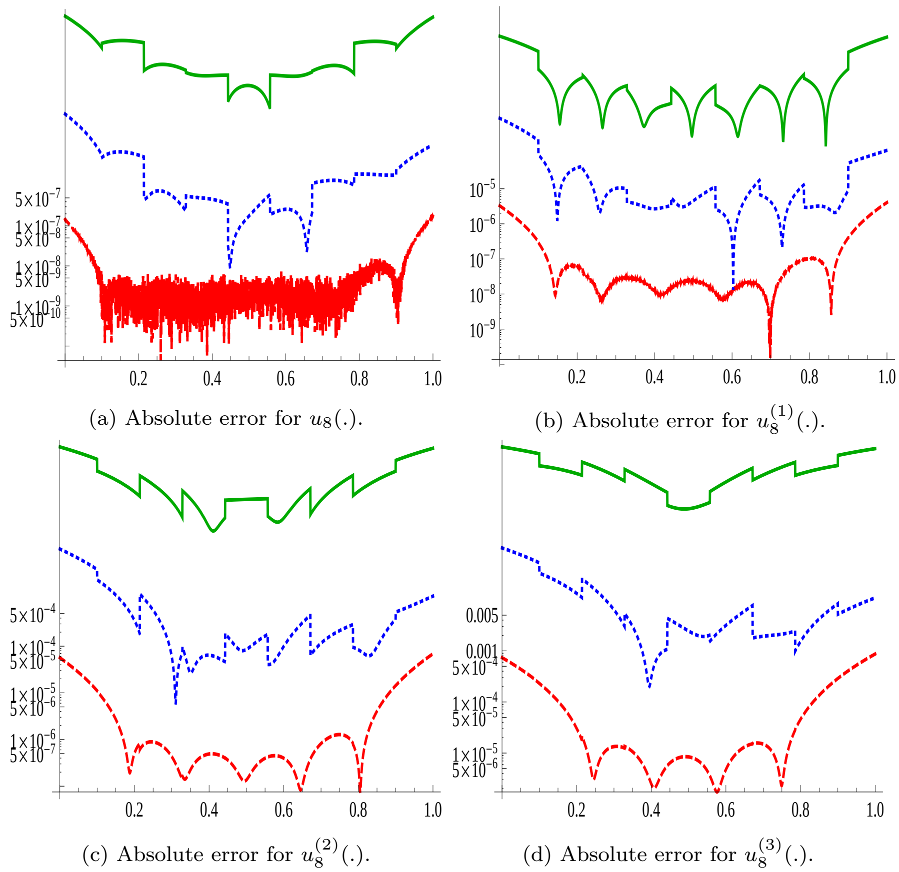
(c) Absolute error for $u_8^{(2)}(.)$.

(d) Absolute error for $u_8^{(3)}(.)$.

Figure 3: Example 3, $m = 4$(green line), $m = 6$(blue dotted), $m = 8$(red dash).

Table 6: The Max Error result of $u^{(2)}(.)$, Example 4.

m	$n = 8$	$n = 16$	$n = 32$
4	4.6866×10^{-1}	3.45622×10^{-1}	3.75916×10^{-1}
6	1.18157×10^{-3}	1.25931×10^{-4}	1.20977×10^{-3}
8	1.3664×10^{-5}	2.79569×10^{-7}	1.67734×10^{-5}



(a) Absolute error for $u_8(.)$.

(b) Absolute error for $u_8^{(1)}(.)$.

(c) Absolute error for $u_8^{(2)}(.)$.

(d) Absolute error for $u_8^{(3)}(.)$.

Figure 4: Example 4, $m = 4$(green line), $m = 6$(blue dotted), $m = 8$(red dash).

6 Conclusion

Different problems have been solved by researchers using reproducing kernel Hilbert spaces (RKHS). They assume the unknown solution belongs to special fix W_2^m and use Gram–Schmidt process to implement RKHS method. In this paper, we attempted to solve Fredholm integral equations of the second kind with weakly singular kernel. We used the Taylor series to remove the singularity. According to the problem, we investigated the problem in reproducing kernel spaces $W^m[-1,1]$. In the proposed method by use of reproducing kernel space property, the problem was converted to a system of linear equations. In the proposed method we do not use Gram process, we use RKHS property and use W_2^m for different m. Though we suppose the unknown solution is smooth, but this is a very strong assumption, and it is enough for the solution to be in W_2^m for some m. In this method for fix m, when the number of bases n increases the coefficient matrix of the linear system will be ill-conditioned; so we increase m and move from one reproducing kernel to other instead, to find approximate solution with high accuracy. Several test examples were used to show the efficiency and applicability of the method.

Acknowledgements

Authors are grateful to there anonymous referees and editor for their constructive comments.

References

1. Allouch, C., Sablonnèire, P., Sbibih, D., and Tahrichi, M. *Product integration methods based on discrete spline quasi-interpolants and application to weakly singular integral equations*, J. Comput. Appl. Math., 233 (2010), no. 11, 2855–2866.

2. Babolian, E. and Arzhang Hajikandi, A. *The approximate solution of a class of Fredholm integral equations with a weakly singular kernel*, J. Comput. Appl. Math., 235 (2011), no. 5, 1148–1159.

3. Babolian, E., Hamedzadeh, D., Jafari, H., Arzhang Hajikandi, A., and Baleanu, D. *On integral equations with weakly singular kernel by using Taylor series and Legendre polynomials*, Open Physics, 13 (2015), no. 1, 1–7.

4. Babolian, E., Javadi, S., and Moradi, E. *Error analysis of reproducing kernel Hilbert space method for solving functional integral equations*, J. Comput. Appl. Math., 300 (2016), 300–311.

5. Badr, A. A. *Integro-differential equation with Cauchy kernel*, J. Comput. Appl. Math., 134 (2001), no. 1-2, 191–199.

6. Beyrami, H., Lotfi, T., and Mahdiani, K. *A new efficient method with error analysis for solving the second kind Fredholm integral equation with Cauchy kernel*, J. Comput. Appl. Math., 300 (2016), 385–399.

7. Chen, Z. and Cheng, X. *An efficient algorithm for solving Fredholm integro-differential equations with weakly singular kernels*, J. Comput. Appl. Math., 257 (2014), 57–64.

8. Chen, Z. and Zhou, Y. *An efficient algorithm for solving Hilbert type singular integral equations of the second kind*, Comput. Math. Appl., 58 (2009), no. 4, 632–640.

9. Cui, M. and Lin, Y. *Nonlinear numerical analysis in the reproducing kernel space*, Nova Science Publishers, Inc., New York, 2009.

10. Du, H., Zhao, G., and Zhao, C. *Reproducing kernel method for solving Fredholm integro-differential equations with weakly singularity*, J. Comput. Appl. Math., 255 (2014), 122–132.

11. Estrada, R. and Kanwal, Ram P. *Singular Integral Equations*, Birkhäuser Boston, Inc., Boston, MA, 2000.

12. Jiang, W. and Cui, M. *The exact solution and stability analysis for integral equation of third or first kind with singular kernel*, Appl. Math. Comput., 202 (2008), no. 2, 666–674.

13. Lakestani, M., Nemati Saray, B., and Dehghan, M. *Numerical solution for the weakly singular Fredholm integro-differential equations using Legendre multiwavelets*, J. Comput. Appl. Math., 235 (2011), no. 11, 3291–3303.

14. Long, G., Nelakanti, G., and Zhang, X., *Iterated fast multiscale Galerkin methods for Fredholm integral equations of second kind with weakly singular kernel*, Appl. Numer. Math., 62 (2012), no. 3, 201–211.

15. Maleknejad, K., Mollapourasl, R., and Ostadi, A. *Convergence analysis of Sinc-collocation methods for nonlinear Fredholm integral equations with weakly singular kernel*, J. Comput. Appl. Math., 278 (2015), 1–11.

16. Mandal, B. N. and Chakrabarti, A. *Applied singular integral equations*, Published by Science Publishers, Enfield, NH; Marketed and distributed by CRC Press, Boca Raton, FL, 2011.

17. Muskhelishvili, N. I. *Singular Integral Equations. Boundary problems of function theory and their application to mathematical physics*, Translation by J. R. M. Radok. P. Noordhoff N. V., Groningen, Holland, 1953.

18. Nili Ahmadabadi, M. and Laeli Dastjerdi, H. *Tau approximation method for the weakly singular Volterra-Hammerstein integral equations*, Appl. Math. Comput., 285 (2016), 241–247.

19. Okayama, T., Matsuo, T., and Sugihara, M. *Sinc-Collocation methods for weakly singular Fredholm integral equations of the second kind*, J. Comput. Appl. Math., 234 (2010), no. 4, 1211–1227.

20. Schneider, C. *Product integration for weakly singular integral equations*, Math. Comp., 36 (1981), no. 153, 207–213.

21. Schneider, C. *Regularity of the solution to a class of weakly singular Fredholm integral equations of the second kind*, Integral Equations Operator Theory, 2 (1979), no. 1, 62–68.

22. Tricomi, F. G. *Integral Equations*, Dover Publications, Inc., New York, 1985.

Iranian Journal of Numerical Analysis and Optimization
Vol. 8, No. 1, (2018), pp 19–39
DOI:10.22067/ijnao.v8i1.49742

A new dwindling nonmonotone filter method without gradient information for solving large-scale systems of equations

F. Arzani and M.R. Peyghami*

Abstract

In this paper, we present a new derivative-free spectral residual method for solving large-scale systems of equations. Our algorithm is equipped with a dwindling multidimensional nonmonotone filter in which whose envelope is dwindling as the step-length of line search is decreasing. The proposed algorithm is also combined with a relaxed nonmonotone line search technique which allows the algorithm to enjoy the nonmonotone property from scratch. Under some standard assumptions, the global convergence property of the proposed algorithm is established. Numerical results on some test problems show the efficiency and effectiveness of the new algorithm in practice.

Keywords: Dwindling filter technique; Systems of equations; Nonmonotone line search; Global convergence.

1 Introduction

In this paper, we deal with the following nonlinear systems of equations:

*Corresponding author
Received 9 September 2015; revised 12 April 2016; accepted 12 July 2017
F. Arzani
Faculty of Mathematics, K.N. Toosi University of Technology, P.O.Box 16315-1618, Tehran, Iran. Scientific Computations in OPtimization and Systems Engineering (SCOPE), K.N. Toosi University of Technology, Tehran, Iran. e-mail: farzani@dena.kntu.ac.ir

M.R. Peyghami
Faculty of Mathematics, K.N. Toosi University of Technology, P.O.Box 16315-1618, Tehran, Iran. Scientific Computations in OPtimization and Systems Engineering (SCOPE), K.N. Toosi University of Technology, Tehran, Iran. e-mail: peyghami@kntu.ac.ir

$$F(x) = 0, \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable map. We assume that (1) is symmetric; that is, its Jacobian is a symmetric matrix. Associated with (1), the norm function is defined by

$$f(x) = \frac{1}{2}\|F(x)\|^2.$$

Several methods in the literature have been proposed for solving (1), such as Newton and quasi-Newton methods, Gauss–Newton method, Levenberg–Marquardt method [20], trust-region methods, spectral gradient, and residual methods [10, 24] and conjugate gradient methods [9]. The main factor in these methods is how the procedure deals with large-scale settings. The spectral gradient and conjugate gradient methods are a class of methods that can suitably cope with large-scale systems of equations. The spectral gradient method was first introduced in [7] for unconstrained optimization and has been successfully extended for solving large-scale nonlinear systems of equations; see, for example, [25, 29], and the references therein.

The so-called nonmonotone line search techniques for unconstrained optimization was introduced by Grippo, Lampariello, and Lucidi in [17]. In their approach, at the point x_k, for a given positive integer M, the stepsize $\lambda \in (0, 1]$ is chosen such that

$$f(x_k + \lambda d_k) \le \max_{0 \le j \le \min\{k,M\}} f(x_{k-j}) + \gamma\lambda\nabla f(x_k)^T d_k,$$

where $\gamma \in (0, 1)$ is a given scalar. An extension and adaption of this technique in the framework of a globalization technique for the Barzilai-Browein (BB) gradient method has been proposed in [18]. It is well known that the Grippo's nonmonotone technique suffers from some drawbacks; some of them have been listed in [1]. In order to overcome these difficulties, Ahookhosh, Amini, and Peyghami in [2] introduced a nonmonotone term as follows:

$$R_k = \epsilon_k f_{\ell(k)} + (1 - \epsilon_k) f_k, \tag{2}$$

where

$$f_{\ell(k)} = \max_{0 \le j \le \min\{k,M\}} f_{k-j}, \tag{3}$$

$f_k = f(x_k)$, and $\epsilon_k \in [\epsilon_{\min}, \epsilon_{\max}] \subset [0, 1]$. The motivation behind this nonmonotone term is that the best convergence results are obtained by stronger nonmonotone strategy whenever the iterates are far from the minimizer and by weaker nonmonotone strategy whenever iterates are close enough to that, see, for example, [30]. Although, (2) enjoys the strong and weak nonmonotone properties; it does not allow the iterates to act as a nonmonotone iterate in the first iterations. Ataee Tarzanagh, Saeidian, Peyghami, and Mesgarani in [6] proposed a nonmonotone term by replacing R_k with $(1 + \psi_k)R_k$ in the trust region ratio, where ψ_k is determined by

$$\psi_k = \begin{cases} \eta_k & \text{if } R_k > 0, \\ 0 & \text{if } R_k \leq 0, \end{cases}$$

and $\{\eta_k\}$ is a positive sequence satisfying the following condition:

$$\sum_{k=1}^{\infty} \eta_k \leq \eta < \infty. \tag{4}$$

This substitution causes an increase in the function value in the first iterations.

Since the Jacobian of (1) might not available or requires a prohibitive amount of storage, derivative-free methods have been widely developed for solving large-scale nonlinear systems of equations. For the square systems, La Cruz, Martínez, and Raydan in [24] proposed a derivative-free spectral residual method, named as DF-SANE, based on a combination of the Li and Fukushima's line search [26] and the Grippo's nonmonotone technique. In their method, the step-length α_k is determined by one and only one of the following two conditions:

$$\begin{aligned} f(x_k - \alpha_k d_k) &\leq f_\ell(k) + \eta_k - \gamma \alpha_k \|F(x_k)\|^2, \\ f(x_k + \alpha_k d_k) &\leq f_\ell(k) + \eta_k - \gamma \alpha_k \|F(x_k)\|^2, \end{aligned}$$

where $d_k := \sigma_k F(x_k)$ and σ_k is the BB spectral coefficient, which is computed by [7]

$$\sigma_k = \frac{s_k^T s_k}{y_k^T s_k},$$

where $s_k = x_{k+1} - x_k$ and $y_k = F(x_{k+1}) - F(x_k)$.

Several variants of DFSANE algorithm for the square systems have been proposed and applied on some specific systems of equations; see, for example, [21–23]. Cheng and Li in [11] introduced another derivative-free spectral residual method for the square systems, named as NDFSANE, in which the nonmonotone term $f_{\ell(k)}$ is replaced by that of proposed in [30]. Later, the performance of DFSANE algorithm for the square systems is improved in SDFSANE algorithm [10] by using an approximation of the steepest descent direction.

Recently, La Cruz in [22] presented a residual spectral algorithm for solving monotone equations on the Hilbert space which is an extension of the SANE and DFSANE algorithms for systems of equations. This algorithm uses in a systematic way the residual $d = F(x_k)$ as a search direction, combined with a suitable step-length and a nonmonotone line search globalization strategy. The main objective of this algorithm is to guarantee the convergence without any necessity to the differentiability of F.

The concept of filter for constrained optimization problems was first introduced by Fletcher and Leyffer in [15] in order to avoid the difficulty of adjusting penalty parameter in penalty methods. Accepting/rejecting of a trial step in the filter methods is inspired from the same idea in multiobjective optimization. Indeed, a trial point is accepted by the filter if it is not dominated by any other points in the filter. In this situation, the filter is updated by removing all the points that are dominated by the new point; see, for example, [15, 16, 27], and the references therein. The fixed envelope of a multidimensional filter keeps an acceptable trial point far from iterates in the filter. However, the trial point with a tiny step-length may get close to the current iterate and be rejected by the filter due to its fixed envelope. Due to this fact, Chen and Sun in [8] proposed a dwindling multidimensional filter line search method for unconstrained optimization. Their idea is to incorporate step-length in the framework of multidimensional filter, and let the envelope of the filter become thin as the step-length approaches zero. Recently, Arzani and Peyghami in [4] proposed a dwindling filter technique for solving positive definite generalized eigenvalue problem.

Our aim in this paper is to introduce a new dwindling nonmonotone filter DFSANE method for solving large-scale nonlinear systems of equations. Our approach is a new version of DFSANE algorithm which is installed by a filter technique. The concept of filter leads to save some points that are eliminated by the DFSANE algorithm. The accumulated points in the filter helps the algorithm to reach the optimal point as quickly as possible. Moreover, our approach uses some relaxations in its structure and is equipped with the nonmonotone term as proposed in [6]. Besides, by using dwindling technique, we have more flexibility for the acceptance of the trial step. Under some standard and suitable assumptions, the global convergence property of the new proposed algorithm is constructed. Numerical results on some test problems validate that the proposed algorithm is practically efficient and robust and it has some priorities to that provided in [31].

The rest of the paper is organized as follows. Section 2 is devoted to describe the new algorithm in details. In Section 3, we establish the global convergence property of the new proposed algorithm under some suitable assumptions. Some preliminary numerical results of applying the new algorithm on some test problems are given in Section 4. Finally, we end up the paper by some concluding remarks in Section 5.

2 The new algorithm

In this section, we propose a new dwindling nonmonotone filter DFSANE algorithm for solving large-scale nonlinear systems of equations. To do so, we first recall the filter that is used in this paper and then describe the new algorithm.

Recently, Fatemi and Mahdavi-Amiri in [13] introduced a new filter that controls the size of the filter. In their proposed filter, a tentative point x_k is acceptable with respect to x_l if there exists $j \in \{1, \dots, n\}$ such that

$$|g_j(x_k)|^{\mu_2} + \theta_2 \|g(x_k)\|^{\mu_1} \leqslant |g_j(x_l)|^{\mu_2} + \theta_1 \|g(x_l)\|^{\mu_1}, \tag{5}$$

where $g(x) = \nabla f(x)$, μ_1, and μ_2 are positive constants and θ_1 and θ_2 satisfy the following relation:

$$0 \leqslant \theta_1 < \theta_2 < \frac{1}{\sqrt{n}}. \tag{6}$$

Therefore, a point x_k is accepted to the filter $\mathcal{F}$ if it is accepted with respect to every x_l with $g(x_l) \in \mathcal{F}$. In the case of acceptance of x_k, $g(x_k)$ is added to the filter and any $g(x_l) \in \mathcal{F}$ with the following property is removed:

$$|g_j(x_k)|^{\mu_2} + \theta_2 \|g(x_k)\|^{\mu_1} \leqslant |g_j(x_l)|^{\mu_2} + \theta_1 \|g(x_l)\|^{\mu_1} \qquad \forall j \in \{1, \dots, n\}.$$

In our dwindling filter, the parameters θ_1 and θ_2, satisfying (6), are considered as a function of the step-length. Since the step-length α in line search techniques is normally less than or equal to one, in our setting, we use the following parameters in the filter (5):

$$\hat{\theta}_1 = \phi(\alpha)\theta_1 \quad \text{and} \quad \hat{\theta}_2 = \phi(\alpha)\theta_2,$$

where θ_1 and θ_2 satisfy (6) and $\phi : [0, 1] \to \mathbb{R}^+$ is a monotonically increasing continuous function with the following properties:

$$\begin{aligned} &\lim_{\alpha \to 0^+} \frac{\phi(\alpha)}{\alpha} = 0, \\ &\phi(\alpha) = 0 \quad \text{if and only if} \quad \alpha = 0 \end{aligned} \tag{7}$$

such that (6) holds for $\hat{\theta}_1$ and $\hat{\theta}_2$.

Now, let us briefly describe one iteration of our new algorithm. Given x_k, if the stopping criteria do not hold, the trial step $d_k = -\sigma_k F_k$ is computed where σ_k is the so called spectral coefficient. Assuming $\alpha_+ = \alpha_- = 1$, the trial point $x_{k+1}^+ = x_k + \alpha_+ d_k$ and $x_{k+1}^- = x_k - \alpha_- d_k$ are introduced and with the priority of x_{k+1}^+ are verified for acceptance to the filter. In case of rejecting both trial points by the filter, these points with the same priority are checked for acceptance by a nonmonotone line search conditions. If the trial points are rejected by the filter and line search conditions, then a backtracking scheme is done on α_+ and α_- and the above mentioned procedure is repeated until one of the trial points is accepted by the filter or nonmonotone line search conditions.

The structure of the new proposed algorithm is outlined in Algorithm 1.

Algorithm 1. *: A new dwindling derivative free filter algorithm*

Step 0. *Given* $x_0 \in \mathbb{R}^n$*, integer* $M \geq 0$*,* $0 < \gamma < 1$*,* $\epsilon > 0$*,* $0 < \sigma_{\min} < \sigma_{\max} < \infty$*,* $0 \leqslant \theta_1 < \theta_2 < \frac{1}{\sqrt{n}}$*,* $0 \leq \tau_{\min} < \tau_{\max} < 1$*, and a monotonically increasing continuous function* ϕ *satisfying* (7)*; set* $\mathcal{F}_0 = \emptyset$ *and* $k := 0$.

Step 1. *If* $\|F(x_k)\| \leq \epsilon$*, then stop.*

Step 2. *Choose the spectral coefficient* σ_k *such that* $|\sigma_k| \in [\sigma_{\min}, \sigma_{\max}]$*, and compute* $f_{\ell(k)}$ *by using* (3)*. Set* $d_k = -\sigma_k F(x_k)$*,* $\alpha_+ = 1$ *and* $\alpha_- = 1$.

Step 3. *Compute* $x^+_{k+1} = x_k + \alpha_+ d_k$*,* $x^-_{k+1} = x_k - \alpha_- d_k$*, and* R_k *by* (2)*.*
If x^+_{k+1} *is accepted by the filter, then set* $x_{k+1} = x^+_{k+1}$ *and add* x_{k+1} *to the filter* $\mathcal{F}_k$*. Update the filter to form* $\mathcal{F}_{k+1}$*. Set* $k := k+1$*,* $\alpha_k = \alpha_+$*, and goto Step 1.*
Else if x^-_{k+1} *is accepted by the filter, then set* $x_{k+1} = x^-_{k+1}$ *and add* x_{k+1} *to the filter* $\mathcal{F}_k$*. Update the filter to form* $\mathcal{F}_{k+1}$*. Set* $k := k+1$*,* $\alpha_k = \alpha_-$*, and goto Step 1.*
Else if

$$f(x^+_{k+1}) \leq (1+\psi_k)R_k - \gamma {\alpha_+}^2 f(x_k), \tag{8}$$

then set $x_{k+1} = x^+_{k+1}$*,* $k := k+1$*,* $\alpha_k = \alpha_+$*, and goto Step 1.*
Else if

$$f(x^-_{k+1}) \leq (1+\psi_k)R_k - \gamma {\alpha_-}^2 f(x_k), \tag{9}$$

then set $x_{k+1} = x^-_{k+1}$*,* $k := k+1$*,* $\alpha_k = \alpha_-$*, and goto Step 1.*

Step 4. *Choose* $\alpha_+ \in [\tau_{\min}\alpha_+, \tau_{\max}\alpha_+]$ *and* $\alpha_- \in [\tau_{\min}\alpha_-, \tau_{\max}\alpha_-]$*, and goto Step 3.*

Remark 1. Since $\psi_k > 0$, $R_k > 0$, and $R_k \geq f(x_k)$, due to Lemma 2.2 in [2], it is easily seen that after finite number of reductions of α_+ or α_- in Step 4 of Algorithm 1, one of the conditions (8) and (9) is necessarily satisfied. This shows that the proposed algorithm is well-defined.

3 Convergence analysis

In this section, our goal is to analyze the global convergence property of Algorithm 1. For this purpose, let

$$P1 = \{k | x_k \in \mathcal{F}_k\},$$

and let

$$P2 = \{k | x_k \text{ satisfies (8) or (9)}\}.$$

It is easily seen that $P1 \cap P2 = \emptyset$. Before going further, we first consider the following assumptions on the problem (1):

A1. F is a continuously differentiable function map over $\mathbb{R}^n$.

A2. For a given η, as in (4), the level set $L^0 = \{x \in \mathbb{R}^n | f(x) \leq e^{\eta} f(x_0)\}$ is closed and bounded.

A3. The sequence $\{x_k\}$, generated by Algorithm 1, is contained in a bounded and closed set $\Omega \subset \mathbb{R}^n$. Let $\{x_k\}$ be the sequence generated by Algorithm 1. Then, Algorithm 1 either stops at a certain iteration k_0 with $\|F(x_{k_0})\| = 0$, or generates an infinite sequence. In the latter case, it is shown that there exists a subsequence of $\{x_k\}$ which converges to a stationary point of F. Therefore, from now on, we assume that $|P1 \cup P2| = \infty$. Two possibilities are recognized for the cardinality of $P1$ and $P2$, $|P1| = \infty$ or $|P2| = \infty$.

3.1 Analysis of the case $|P1| = \infty$

Consider the filter $\mathcal{F}_k$ and its acceptance criterion, which is given by (5). Due to Lemma 4 in [13], the size of $\mathcal{F}_k$ is finite. Although, $\mathcal{F}_k$ contains finite number of elements; it may happen that an infinite number of iteration points is accepted by the filter. The following lemma states the behaviour of the algorithm in this case.

Lemma 1. *Let Assumptions A1–A3 hold, and let $|P1| = \infty$. Then,*

$$\lim_{k \to \infty,\, k \in P1} \|F_{k+1}\| = 0. \tag{10}$$

Proof. The proof is similar to the proof of Lemma 5.4 in [14] by setting $\theta(x) = F(x)$ and using the fact that based on the definition of $\phi(\alpha)$ and the initial values of $\alpha_+ = \alpha_- = 1$ in Step 2, the condition $0 \leqslant \theta_1 < \theta_2 < \frac{1}{\sqrt{n}}$ implies that $0 \leqslant \hat{\theta}_1 < \hat{\theta}_2 < \frac{1}{\sqrt{n}}$. □

From Lemma 1, for the case $|P1| = \infty$, it is concluded that there exists a limit point of subsequence $\{x_k\}_{k\in P1}$ which is a stationary point of f; that is,

$$\lim_{k\to\infty,\, k\in P1} \inf \|F_k\| = 0. \tag{11}$$

3.2 Analysis of the case $|P2| = \infty$

First of all, without loss of generality, we may assume that $|P2| = \infty$ and that $|P1|$ is finite, as for the case $|P1| = \infty$ (11) holds. Under this assumption, we show that there exists a limit point of subsequence $\{x_k\}_{k\in P2}$ satisfying the first-order necessary condition; that is,

$$\lim_{k\to\infty,\, k\in P2} \inf \|F_k\| = 0.$$

Since $|P1|$ is finite and $|P2| = \infty$, there exists a positive integer M_0 such that for every $k \geq M_0$, we have $k \in P2$. Without loss of generality and in order to establish simple analysis, in what follows, we assume that $M_0 = 0$, and therefore

$$P2 = \{0, 1, 2, \ldots\}. \tag{12}$$

To construct convergence property of Algorithm 1 in this case, we first state some technical lemmas.

Lemma 2. *Suppose that $\{x_k\}_{k\in P2}$ is generated by Algorithm 1. Then, we have*

$$f_{k+1} \leq |f_0| \prod_{i=0}^{k} (1 + \psi_i) - \omega_k,$$

where $\omega_k = \gamma {\alpha_k}^2 f(x_k)$.

Proof. The proof is similar to the proof of Lemma 3 in [5], and therefore we omit it here. □

The following lemma shows that all the iteration points that are generated by Algorithm 1 are contained in the level set L^0. Using Lemma 2, its proof is the same as the proof of Lemma 4 in [5] and Lemma 3.4 in [28].

Lemma 3. *For the sequence $\{x_k\}_{k\in P2}$, generated by Algorithm 1, we have $\{x_k\} \subseteq L^0$.*

Remark 2. Let $\{x_k\}_{k\in P2}$ be generated by Algorithm 1, and let $|P2| = \infty$. Then, due to the (12), for a given positive integer M, one can rewrite $P2$ as below:

$$P2 = \{LM + r \mid L \in \mathbb{N} \cup \{0\}, 0 \le r \le M-1\}.$$

Lemma 4. *Assume that $\{x_k\}_{k\in P2}$ is an infinite sequence, generated by Algorithm 1 and that M is the constant as given by (3). Then, for every $k \in P2$, there exist a non-negative integer L and $0 \le r \le M-1$ such that $k = LM + r$ and*

$$f_{k+1} = f_{LM+r+1} \le |f_0| \prod_{i=0}^{LM+r} (1+\psi_i) - \sum_{i=0}^{L} \omega_{s(i)}, \qquad L = 0, 1, 2, \ldots,$$

where $\omega_{s(i)} = \min_{iM \le j \le (i+1)M-1} \omega_j$ and ω_j is defined as in Lemma 2.

Proof. The proof is similar to the proof of Lemma 6 in [5]. □

Lemma 5. *Suppose that Assumptions* **A1** *and* **A2** *hold and that $\{x_k\}_{k\in P2}$ is the infinite sequence generated by Algorithm 1. Then, one has*

$$\lim_{i\to\infty} \omega_{s(i)} = 0,$$

where $\omega_{s(i)}$ is defined as in Lemma 4.

Proof. Using Lemmas 3 and 4, for $1 \le r \le M$ and $LM + r \in I$, we have

$$\sum_{i=0}^{L} \omega_{s(i)} \le |f_0| \prod_{i=0}^{LM+r-1} (1+\psi_i) - f_{LM+r} \le e^{\eta}|f_0| - f_{LM+r}.$$

By taking limit from both side of this inequality, as $L \to \infty$, and using Assumption **A1**, we conclude that $\sum_{i=0}^{\infty} \omega_{s(i)} < \infty$, which implies that

$$\lim_{i\to\infty} \omega_{s(i)} = 0. \tag{13}$$

This completes the proof of the lemma. □

From now on, we define $K = \{s(i) | i = 1, 2, \ldots\}$. The following theorem establishes the global convergence of Algorithm 1.

Theorem 1. *Let Assumptions* **A1**–**A3** *hold, and let $\{x_k\}$ be the sequence generated by Algorithm 1. Then, for every limit point x^* of $\{x_k\}_{k\in K}$, one has*

$$\langle J(x^*)^T F(x^*), F(x^*) \rangle = 0,$$

where $\langle .,. \rangle$ stands for the inner product of two vectors in $\mathbb{R}^n$.

Proof. From Lemma 5, we have

$$\lim_{k\in K,\, k\to\infty} \omega_{s(k)} = \lim_{k\in K,\, k\to\infty} \min_{kM \le j \le (k+1)M-1} \gamma \alpha_j^2 \|F(x_j)\|^2 = 0,$$

which implies that

$$\lim_{k\in K,\, k\to\infty} \alpha_k^2 \|F(x_k)\|^2 = 0, \tag{14}$$

using the fact that $0 < \gamma < 1$. Since x^* is a limit point of $\{x_k\}_K$, there exists an infinite index set $K_1 \subset K$ such that $\lim_{k\in K_1,\, k\to\infty} x_k = x^*$. Besides, (14) implies that

$$\lim_{k\in K_1,\, k\to\infty} \alpha_k^2 \|F(x_k)\|^2 = 0. \tag{15}$$

To proceed, we consider the following two possible cases:
Case 1: $\lim_{k\in K_1} \alpha_k \neq 0$.
In this case, there exist an infinite sequence of indices $K_2 \subset K_1$ and a positive constant c such that $\alpha_k \geq c > 0$ for all $k \in K_2$. Therefore, (14) implies that

$$\lim_{k\in K_2} \|F(x_k)\|^2 = 0. \tag{16}$$

Now, (16) along with the continuity of F imply $F(x^*) = 0$.
Case2: $\lim_{k\in K_1} \alpha_k = 0$.
In this case, there exists $k_0 \in K_1$ such that $\alpha_k < 1$, for all $k_0 \leq k \in K_1$. Therefore, for infinitely many iterations, the line search is not immediately successful and α_+ and α_- should be updated at least once based on Step 4 of Algorithm 1. Suppose that, in step $k \in K_1$, α_+ and α_- are adapted m_k times in the line search procedure in Steps 3–4. Let α_k^+ and α_k^- be the values of α^+ and α^-, respectively, in the last unsuccessful line search process; that is, inequalities (8) and (9) are violated. Now, form Step 4 of Algorithm 1, for all $k_0 \leq k \in K_1$, we have

$$\alpha_k \geq \tau_{\min}^{m_k}.$$

This inequality together with $\lim_{k\in K_1} \alpha_k = 0$ and the fact that $\tau_{\min} < 1$ imply that

$$\lim_{k\in K_1} m_k = \infty.$$

On the other hand, Step 4 of Algorithm 1 implies that $\alpha_k^+ \leq \tau_{\max}^{m_k - 1}$ and $\alpha_k^- \leq \tau_{\max}^{m_k-1}$. Therefore, using the fact that $\tau_{\max} < 1$, we deduce that

$$\lim_{k\in K1} \alpha_k^+ = \lim_{k\in K1} \alpha_k^- = 0.$$

Now, from the definition of α_k^+ and α_k^-, for all $k \geq k_0$, $k \in K_1$, we have

$$f(x_k - \alpha_k^+ \sigma_k F(x_k)) > (1+\psi_k)R_k - \gamma(\alpha_k^+)^2 f(x_k), \tag{17}$$
$$f(x_k + \alpha_k^- \sigma_k F(x_k)) > (1+\psi_k)R_k - \gamma(\alpha_k^-)^2 f(x_k). \tag{18}$$

Since $\psi_k > 0$ and $R_k > 0$, then $(1+\psi_k)R_k > R_k \geq 0$. Besides, from Lemma 2.2 in [2], one has $R_k > f(x_k)$. Thus, (17) implies that

$$f(x_k - \alpha_k^+ \sigma_k F(x_k)) - f(x_k) > -\gamma(\alpha_k^+)^2 f(x_k).$$

From Lemma 3 and Assumptions **A1**–**A2**, $\{f(x_k)\}$ is a bounded above sequence. Thus

$$f(x_k - \alpha_k^+ \sigma_k F(x_k)) - f(x_k) > -\gamma(\alpha_k^+)^2 e^\eta f(x_0),$$

which implies that

$$\left\|F(x_k - \alpha_k^+ \sigma_k F(x_k))\right\|^2 - \|F(x_k)\|^2 > -\gamma(\alpha_k^+)^2 e^\eta f(x_0),$$

and therefore,

$$\frac{\left\|F(x_k - \alpha_k^+ \sigma_k F(x_k))\right\|^2 - \|F(x_k)\|^2}{\alpha_k^+} > -\gamma \alpha_k^+ e^\eta f(x_0).$$

Now, by Mean Value theorem, there exists $\xi_k \in [0,1]$ such that

$$\sigma_k \langle g(x_k - \xi_k \alpha_k^+ F(x_k)), -F(x_k)\rangle \geq -e^\eta f(x_0)\gamma\alpha_k^+. \tag{19}$$

Due to Algorithm 1, two possible cases may happen for σ_k, it might be positive or negative for infinitely many indices. If $\sigma_k > 0$ for infinitely many indices k, then (19) implies that

$$\langle g(x_k - \xi_k \alpha_k^+ F(x_k)), F(x_k)\rangle \leq e^\eta f(x_0)\frac{\gamma\alpha_k^+}{\sigma_k} \leq e^\eta f(x_0)\frac{\gamma\alpha_k^+}{\sigma_{\min}}. \tag{20}$$

Using an analogous approach, from (18), we conclude that

$$\langle g(x_k + \xi_k{}' \alpha_k^- F(x_k)), F(x_k)\rangle \geq -e^\eta f(x_0)\frac{\gamma\alpha_k^-}{\sigma_k} \geq -e^\eta f(x_0)\frac{\gamma\alpha_k^-}{\sigma_{\min}}, \tag{21}$$

for some $\xi_k{}' \in [0,1]$.
Now, as $\alpha_k^+ \to 0$, $\alpha_k^- \to 0$, and $\|\sigma_k F(x_k)\|$ is bounded, taking limit from both sides in (20) and (21) leads to the following equality:

$$\langle J(x^*)^T F(x^*), F(x^*)\rangle = 0. \tag{22}$$

For the case in which $\sigma_k < 0$, for infinitely many indices, proceeding in a similar approach, we can deduce (22) as well. $\square$

The following corollary is an immediate consequence of Theorem 1.

Corollary 1. *Assume that x^* is a limit point of the sequence $\{x_k\}_{k\in K}$ generated by Algorithm 1 and that $\langle J(x^*)v, v\rangle \neq 0$ for all $0 \neq v \in \mathbb{R}^n$. Then, $F(x^*) = 0$.*

Lemma 6. *For $k \in \mathbb{N}$, let*

$$L_k := f(x_{\ell(kM)}) = f_{\ell(kM)},$$

where $\ell(kM) \in \{(k-1)M+1, \ldots, kM\}$ *and* $f_{\ell(j)}$ *is defined as in* (3). *Then, we have*

$$L_{k+1} := f(x_{\ell((k+1)M)}) \leq \left[1 + \sum_{i=kM}^{\infty} \eta_i + \prod_{i=kM}^{\infty} \eta_i\right] L_k.$$

Proof. Using Step 3 of Algorithm 1, we have

$$\begin{aligned}
f(x_{kM+1}) &\leq (1+\psi_{kM})R_{kM} \leq (1+\psi_{kM})L_k \\
f(x_{kM+2}) &= (1+\psi_{kM+1})[\varepsilon_k \max\{L_k, f(x_{kM+1})\} + (1-\varepsilon_k)f(x_{kM+1})] \\
&\leq (1+\psi_{kM+1})(1+\psi_{kM})L_k \\
&= [1+\psi_{kM}+\psi_{kM+1}+\psi_{kM}\psi_{kM+1}]L_k,
\end{aligned}$$

and so on. Therefore, by an inductive argument,

$$f(x_{kM+j}) \leq \left[1 + \sum_{i=kM}^{kM+j} \eta_i + \prod_{i=kM}^{kM+j} \eta_i\right] L_k.$$

Since $\ell(k+1) \in \{kM+1, \ldots, (k+1)M\}$, then, we have

$$L_{k+1} = f(x_{l((k+1)M)}) \leq \left[1 + \sum_{i=kM}^{\infty} \eta_i + \prod_{i=kM}^{\infty} \eta_i\right] L_k.$$

This completes the proof of the lemma. □

Let us define K_+ as below:

$$K_+ = \{\ell(1), \ell(2), \ldots\}.$$

One can easily observe that

$$\ell(j+1) \leq \ell(j) + 2M - 1, \qquad j = 1, 2, \ldots. \tag{23}$$

The following theorem states that in case of existing a limit point for the sequence generated by Algorithm 1, all the limit points are the solutions of the nonlinear system (1).

Theorem 2. *Let* x^* *be a limit point of the sequence* $\{x_k\}$, *generated by Algorithm 1, such that* $F(x^*) = 0$. *Then, we have*

$$\lim_{k\to\infty} F(x_k) = 0.$$

Proof. Let K_1 be an infinite subset of $\mathbb{N}$ such that

$$\lim_{k\in K_1,\, k\to\infty} x_k = x^*.$$

This relation together with Assumption **A1** imply that

$$\lim_{k\in K_1,\, k\to\infty} F(x_k) = 0. \tag{24}$$

Now, since $x_{k+1} = x_k \pm \alpha_k\sigma_k F(x_k)$ and $|\alpha_k\sigma_k| < \sigma_{\max}$, for all $k \in \mathbb{N}$, using (24), one can deduce that

$$\lim_{k\in K_1,\, k\to\infty} \|x_{k+1} - x_k\| = 0,$$

and therefore,

$$\lim_{k\in K_1,\, k\to\infty} x_{k+1} = x^*.$$

Proceeding by induction, one can easily show that for the fixed $j \in \{0, 1, \ldots, 2M-1\}$, we have

$$\lim_{k\in K_1,\, k\to\infty} x_{k+j} = x^*. \tag{25}$$

Now, using (23), for a given $k \in K_1$, there exists $\mu(k) \in \{0, 1, \ldots, 2M-1\}$ such that $k + \mu(k) \in K_+$. Since K_+ is infinite, there exists $\bar{i} \in \{0, 1, \ldots, 2M-1\}$ such that, $\mu(k) = \bar{i}$ for infinitely many $k \in K_1$. Assume that

$$K_2 = \{k + \mu(k) | k \in K_1 \text{ and } \mu(k) = \bar{i}\}.$$

It is easily seen that $K_2 \subset K_+$, and by (25)

$$\lim_{k\in K_2,\, k\to\infty} x_{k+1} = x^*.$$

Thus

$$\lim_{k\in K_2,\, k\to\infty} F(x_k) = 0.$$

Now, as $K_2 \subset K_+$, there exists an infinite subsequence $\{x_{\ell(j)}\}_{j\in J}$ such that

$$\lim_{j\in J} x_{\ell(j)} = x^* \quad \text{and} \quad \lim_{j\in J} f(x_{\ell(j)}) = \lim_{j\in J} L_j = 0. \tag{26}$$

Let us write $J = \{j_1, j_2, j_3, \ldots\}$, where $j_1 < j_2 < j_3 < \cdots$. Then, (26) implies that

$$\lim_{i\to\infty} L_{j_i} = 0. \tag{27}$$

Now, from Lemma 6, we have

$$L_p \le \left[1 + \sum_{t=Mj_i}^{\infty} \eta_t + \prod_{t=Mj_i}^{\infty} \eta_t\right] L_{j_i} \qquad \forall\, p \in \mathbb{N}, p > j_i.$$

Thus,

$$\sup_{p\geq j_i} L_p \leq \left[1+\sum_{t=Mj_i}^{\infty}\eta_t+\prod_{t=Mj_i}^{\infty}\eta_t\right]L_{j_i}. \tag{28}$$

On the other hand, from (4), we have

$$\lim_{i\to\infty}\left[\sum_{t=Mj_i}^{\infty}\eta_t+\prod_{t=Mj_i}^{\infty}\eta_t\right]=0.$$

Therefore, using (27) and (28), we conclude that

$$\lim_{i\to\infty}\sup_{p\geq j_i} L_p = 0,$$

which in turn implies that

$$\lim_{j\to\infty} L_j = 0.$$

Now, using the definition of L_j, we have

$$\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty}\frac{1}{2}\left\|F(x_k)\right\|^2 = 0.$$

This completes the proof of the theorem. □

4 Numerical results

In this section, our aim is to investigate the practical performance of Algorithm 1, denoted by DF-DFSANE, along with the following algorithms on some test problems:

- DFSANE: Algorithm 1 in [24].
- NF-DFSANE: Algorithm 1 in this paper in which the concept of filter has been removed.

All the considered algorithms are implemented in MATLAB 7.10.0 (R2010a) environment on a PC with CPU Intel 2.33 GHz Quad, 4GB RAM memory, and double precision format. Test problems are mostly taken from [3, 19], and the problem dimensions vary from 2 to 10000.

All the considered algorithms are being stopped whenever $\| F(x_k) \| \leq 10^{-6}$. Moreover, we declare that an algorithm is failed whenever the number of iterations and the number of function evaluations exceed 10000 and 50000, respectively. We have also utilized the advantages of the performance profile of Dolan and Moré [12] in order to properly compare the considered algorithms. The following parameters are considered in the relevant algorithms:

$$\epsilon = 10^{-6}, \mu_1 = 0.25, \mu_2 = 0.75, \sigma_{\min} = 10^{-6}, \sigma_{\max} = 10^{+6},$$
$$M = 20, \gamma = 10^{-4}, \tau_{\min} = 0.1, \tau_{\max} = 0.5, \eta_k = \frac{1}{(1+k)^2}, \; k \geq 0.$$

Moreover, we utilized $\phi(\alpha) = \alpha^{\frac{3}{2}}$ in the dwindling technique. It is worth mentioning that the parameters are chosen from a set of values in which reasonably better results are achieved while performing an algorithm. Numerical results are given in Table 1. In this table, $Problem$, n, n_i, and n_f represent the problem name, the problem size, the number of iterations, and the number of function evaluations, respectively. Based on the results of Table 1, we have plotted the performance profile of the considered algorithms in Figures 1 and 2 in terms of n_i and n_f, respectively.

At a glance to Figure 1, one can easily find out that DF-DFSANE algorithm solves roughly 93% of the test problems successfully while this percentage for NF-DFSANE and DFSANE algorithms is below 90%. Moreover, DF-DFSANE algorithm solves about 62% of problems at the lowest value of n_i while this percentage for the NF-DFSANE and DFSANE algorithms are 40% and 10%, respectively.

Figure 2 reveals that when the DF-DFSANE, NF-DFSANE, and DFSANE algorithms are applied on the test problems; the DF-DFSANE algorithm solves roughly 49% of the test problems in the lowest value of n_f while this number for the NF-DFSANE and DFSANE algorithms are 38% and 10%, respectively. Moreover, DF-DFSANE algorithm solves about 90% of the test problems without any failure. It has to be noticed that in the cases that DF-DFSANE is not the best algorithm; its performance index is roughly close to the performance index of the best algorithm.

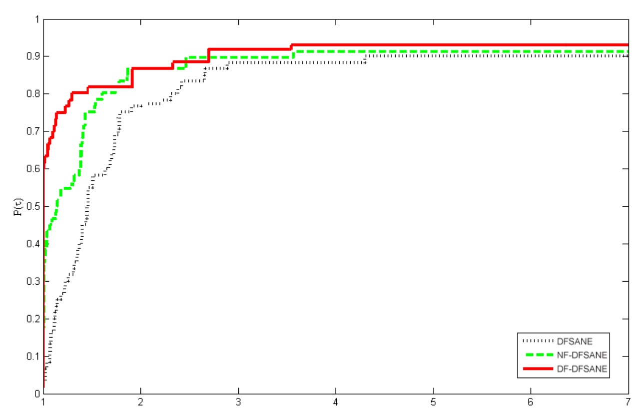


Figure 1: Performance profile of the considered algorithms in terms of n_i.

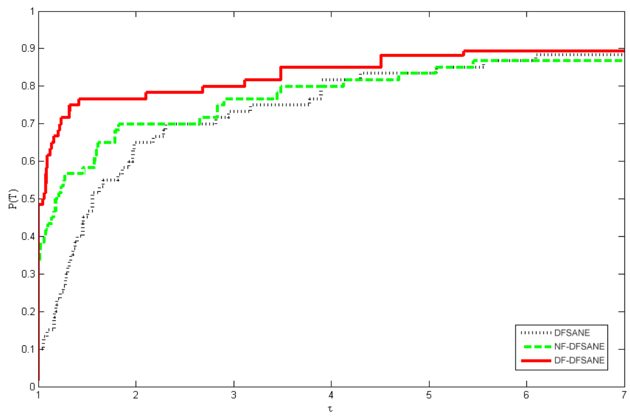


Figure 2: Performance profile of the considered algorithms in terms of n_f.

5 Conclusion

In this paper, a new dwindling nonmonotone filter DFSANE method for solving large-scale nonlinear systems of equations is proposed. In our approach, a filter is set up on the so-called DFSANE algorithm which has been developed in [24]. Indeed, the concept of filter helps us to save some appropriate points that are eliminated by the DFSANE algorithm. Moreover, the accumulated points in the filter cause the algorithm to reach the optimal point as quickly as possible. The proposed approach also uses some relaxations in its structure and is equipped with the nonmonotone term as proposed in [6]. Furthermore, we have more flexibility for the acceptance of the trial step by using a dwindling technique. Under some standard assumptions, the global convergence property is established. Numerical results on some test problems confirm that our proposed algorithm is practically efficient, robust, and has some priorities to that provided in [31].

Acknowledgements

The authors would like to thank the Research Council of K.N. Toosi University of Technology and the SCOPE research center for supporting this work.

Table 1: Test problems and numerical results.

Problem	n	DFSANE n_i/n_f	NF-DFSANE n_i/n_f	DF-DFSANE n_i/n_f
Extended Beale [3]	1000	38/52	33/46	29/39
Extended Beale [3]	5000	53/79	47/57	38/52
Extended Beale [3]	10000	62/90	48/72	38/64
Extended penalty [3]	1000	17/39	13/35	10/24
Extended penalty [3]	5000	21/49	17/45	21/47
Extended penalty [3]	10000	23/49	21/51	22/50
Extended Three Exponential [3]	1000	18/32	16/24	15/19
Extended Three Exponential [3]	5000	29/43	18/32	14/30
Extended Three Exponential [3]	10000	30/62	20/34	20/32
Generalized Tridiagonal-2 [3]	1000	58/68	58/72	58/72
Generalized Tridiagonal-2 [3]	5000	Failed	193/296	126/258
Generalized Tridiagonal-2 [3]	10000	Failed	789/3299	440/1083
Extended PSC1 Function [3]	1000	18/26	14/21	14/20
Extended PSC1 Function [3]	5000	22/42	18/26	17/23
Extended PSC1 Function [3]	10000	36/42	22/34	17/23
Extended Block Diagonal BD1 [3]	1000	17/23	17/23	15/23
Extended Block Diagonal BD1 [3]	5000	19/25	15/23	20/24
Extended Block Diagonal BD1 [3]	10000	99/131	114/213	105/285
DQDRTIC (CUTE) [3]	1000	64/116	52/74	50/70
DQDRTIC (CUTE) [3]	5000	89/209	60/90	35/61
DQDRTIC (CUTE) [3]	10000	126/301	44/66	33/47
LIARWHD function [3]	1000	1/1	1/1	1/1
LIARWHD function [3]	5000	18/90	18/90	16/74
LIARWHD function [3]	10000	87/233	111/401	105/341
Extended DENSCHNF [3]	1000	1/1	1/1	1/1
Extended DENSCHNF [3]	5000	25/41	21/49	21/49
Extended DENSCHNF [3]	10000	52/76	25/41	22/40
Generalized Quartic [3]	1000	17/23	10/12	10/10
Generalized Quartic [3]	5000	20/28	10/12	2/4
Generalized Quartic [3]	10000	50/98	21/29	12/16
Diagonal 8 [3]	1000	8/9	6/7	5/6
Diagonal 8 [3]	5000	20/37	16/17	7/9
Diagonal 8 [3]	10000	38/95	5/38	7/9
Full Hessian FH3 [3]	1000	4/24	4/24	4/22
Full Hessian FH3 [3]	5000	5/31	4/30	4/28
Full Hessian FH3 [3]	10000	Failed	Failed	Failed
SINCOS [3]	1000	18/26	14/21	11/16
SINCOS [3]	5000	27/33	18/26	17/23
SINCOS [3]	10000	36/42	30/36	17/23
HIMMELH (CUTE) [3]	1000	1/1	1/1	1/1
HIMMELH (CUTE) [3]	5000	14/18	14/18	8/10
HIMMELH (CUTE) [3]	10000	Failed	Failed	Failed

Table 1: Test problems and numerical results (Continued).

Problem	n	DFSANE n_i/n_f	NF-DFSANE n_i/n_f	DF-DFSANE n_i/n_f
Power [3]	1000	75/91	74/90	54/62
Power [3]	5000	152/176	123/142	120/134
Power [3]	10000	224/345	210/311	129/135
FLETCHCR function (CUTE) [3]	1000	750/1404	712/1201	425/689
FLETCHCR function (CUTE) [3]	5000	Failed	1012/4865	994/4752
FLETCHCR function (CUTE) [3]	10000	Failed	Failed	Failed
Problem 2 [19]	2	46/49	25/31	2/6
problem 4 [19]	5	1/2	1/2	1/2
Problem 6 [19]	2	3/7	4/15	3/6
Problem 7 [19]	5	268/271	62/74	33/42
Problem 8 [19]	10	15/18	19/25	14/16
Problem 26 [19]	5	149/157	21/21	21/21
Problem 27 [19]	2	1/1	1/1	1/1
Problem 39 [19]	5	1383/2100	24/28	20/34
Problem 40 [19]	5	Failed	Failed	Failed
Problem 42 [19]	2	1/1	1/1	1/1
Problem 46 [19]	5	Failed	Failed	Failed
Problem 47 [19]	10	324/841	260/642	222/423
Problem 48 [19]	5	423/1108	112/504	53/71
Problem 53 [19]	5	1/1	1/1	1/1
Problem 56 [19]	10	22/24	3/5	3/3
Problem 61 [19]	10	394/5784	357/5650	357/5650
Problem 63 [19]	5	22/26	47/66	21/38
Problem 77 [19]	10	149/219	51/126	17/24
Problem 78 [19]	2	Failed	Failed	Failed
Problem 79 [19]	5	970/4130	952/4025	844/3441
Problem 81 [19]	5	Failed	Failed	Failed
Problem 81 [19]	10	Failed	843/4344	432/3456
problem 111 [19]	5	344/1815	310/1012	72/770
problem 111 [19]	10	Failed	572/4567	424/3245

References

1. Ahookhosh, M. and Amini, K. *A nonmonotone trust region method with adaptive radius for unconstrained optimization*, Comput. Math. Appl. 60 (2010), no. 3, 411–422.

2. Ahookhosh, M., Amini, K., and Reza Peyghami, M. *A nonmonotone trust-region line search method for large-scale unconstrained optimization*, Appl. Math. Model. 36 (2012), no. 1, 478–487.

3. Andrei, N. *An unconstrained optimization test functions collection*, Adv. Model. Optim. 10 (2008), no. 1, 147–161.

4. Arzani, F. and Reza Peyghami, M. *An approach based on dwindling filter method for positive definite generalized eigenvalue problem*, Comp. Appl. Math. (2016). doi:10.1007/s40314-016-0391-z

5. Arzani, F. and Reza Peyghami, M. *A new nonmonotone filter Barzilai-Borwein method for solving unconstrained optimization problems*, Int. J. Comput. Math. 93 (2016), no. 3, 596–608.

6. Ataee Tarzanagh, D., Saeidian, Z., Reza Peyghami, M., and Mesgarani, H. *A new trust region method for solving least-square transformation of system of equalities and inequalities*, Optim. Lett. 9 (2015), no. 2, 283–310.

7. Barzilai, J. and Borwein, J. M. *Two-point step size gradient methods*, IMA J. Numer. Anal. 8 (1988), 141–148.

8. Chen, Y. and Sun, W. *A dwindling filter line search method for unconstrained optimization*, Math. Comp. 84 (2015), no. 291, 187–208.

9. Cheng, W. Y. *A two-term PRP based-descent method*, Numer. Funct. Anal. Optim. 28 (2007), no. 11-12, 1217–1230.

10. Cheng, W. and Chen, Z. *Nonmonotone spectral method for large-scale symmetric nonlinear equations*, Numer. Algorithms 62 (2013), 149–162.

11. Cheng, W. and Li, D. H. *A derivative-free nonmonotone line search and its application to the spectral residual method*, IMA J. Numer. Anal. 29 (2009), no. 3, 814–825.

12. Dolan, E. and Moré, J. J. *Benchmarking optimization software with performance profiles*, Math. Program. 91 (2002), no. 2, Ser. A, 201–213.

13. Fatemi, M. and Mahdavi-Amiri, N. *A filter trust-region algorithm for unconstrained optimization with strong global convergence properties*, Comput. Optim. Appl. 52 (2012), no. 1, 239–266.

14. Fatemi, M. and Mahdavi-Amiri, N. *A non-monotone trust region algorithm for unconstrained optimization with dynamic reference iteration updates using filter*, Optimization 61 (2012), no. 6, 733–763.

15. Fletcher, R. and Leyffer, S. *Nonlinear programming without a penalty function*, Math. Program. 91 (2002), no. 2, Ser. A, 239–269.

16. Fletcher, R., Leyffer, S., and Toint, Ph.L. *A brief history of filter methods*, SIAG/OPT Views-and-News, 18 (2006), 2–12.

17. Grippo, L., Lampariello, F., and Lucidi, S. *A class of nonmonotone stabilization methods in unconstrained optimization*, Numer. Math. 59 (1991), no. 8, 779–805.

18. Grippo, L. and Sciandrone, M. *Nonmonotone globalization techniques for the Barzilai-Borwein gradient method*, Comput. Optim. Appl. 23 (2002), no. 2, 143–169.

19. Hock, W. and Schittkowski, K. *Test examples for nonlinear programming codes*, Lecture Notes in Economics and Mathematical Systems, 187. Springer-Verlag, Berlin-New York, 1981.

20. Kanzow, C., Yamashita, N., and Fukushima, M. *Levenberg–Marquardt methods for constrained nonlinear equations with strong local convergence properties*, J. Comput. Appl. Math. 172 (2004), no. 2, 375–397.

21. La Cruz, W. *A projected derivative-free algorithm for nonlinear equations with convex constraints*, Optim. Methods Softw. 29 (2014), no. 1, 24–41.

22. La Cruz, W. *Residual spectral algorithm for solving monotone equations on a Hilbert space*, Appl. Math. Comput. 219 (2013), no. 12, 6633–6644.

23. La Cruz, W. *A residual method for solving nonlinear operator equations and their application to nonlinear integral equations using symbolic computation*, Appl. Math. Comput. 217 (2010), no. 1, 11–24.

24. La Cruz, W., Martínez, J.M., and Raydan, M. *Spectral residual method without gradient information for solving large-scale nonlinear systems of equations*, Math. Comp. 75 (2006), no. 255, 1429–1448.

25. La Cruz, W. and Raydan, M. *Nonmonotone spectral methods for large-scale nonlinear systems*, Optim. Methods Softw. 18 (2003), no. 5, 583–599.

26. Li, D. H. and Fukushima, M. *A derivative-free line search and global convergence of Broyden-like method for nonlinear equations*, Optim. Methods Softw. 13 (2000), no. 3, 181–201.

27. Li, C. and Sun, W. *On filter-successive linearization methods for nonlinear semidefinite programming*, Sci. China Ser. A 52 (2009), no. 11, 2341–2361.

28. Peyghami, M.R. and Ataee Tarzanagh, D. *A relaxed nonmonotone adaptive trust region method for solving unconstrained optimization problems*, Comput. Optim. Appl. 61 (2015), no. 2, 321–341.

29. Sun, W. and Yuan, Y. *Optimization Theory and Methods*, Nonlinear programming. Springer Optimization and Its Applications, 1. Springer, New York, 2006.

30. Zhang, H. and Hager, W. W. *A nonmonotone line search technique and its application to unconstrained optimization*, SIAM J. Optim. 14 (2004), no. 4, 1043–1056.

31. Zhang, Y., Sun, W., and Qi, L. *A nonmonotone filter Barzilai-Borwein method for optimization*, Asia-Pac. J. Oper. Res. 27 (2010), no. 1, 55–69.

Iranian Journal of Numerical Analysis and Optimization
Vol. 8, No. 1, (2018), pp 41–61
DOI:10.22067/ijnao.v8i1.62683

WENO schemes for multidimensional nonlinear degenerate parabolic PDEs

R. Abedian*

Abstract

In this paper, a scheme is presented for approximating solutions of nonlinear degenerate parabolic equations which may contain discontinuous solutions. In the one-dimensional case, following the idea of the local discontinuous Galerkin method, first the degenerate parabolic equation is considered as a nonlinear system of first order equations, and then this system is solved using a fifth-order finite difference weighted essentially nonoscillatory (WENO) method for conservation laws. This is the first time that the minmod-limiter combined with weighted essentially nonoscillatory procedure has been applied to the degenerate parabolic equations. Also, it is necessary to mention that the new scheme has fifth-order accuracy in smooth regions and second-order accuracy near singularities. The accuracy, robustness, and high-resolution properties of the new scheme are demonstrated in a variety of multidimensional problems.

Keywords: WENO schemes; Finite difference scheme; Multidimensional nonlinear degenerate parabolic equation; Porous medium equation.

1 Introduction

This paper is concerned with multidimensional of nonlinear degenerate parabolic equations of the form

$$u_t = \Delta \mathbf{b}(u) + S(\mathbf{x}, t, u), \qquad \mathbf{x} = (x_1, \ldots, x_d) \in \Omega \text{ and } \ t \geq 0, \tag{1}$$

satisfying the initial condition

*Corresponding author
Received 17 February 2017; revised 7 may 2017; accepted 12 July 2017
R. Abedian
School of Engineering Science, College of Engineering, University of Tehran, Iran. e-mail: rabedian@ut.ac.ir

$$u(\mathbf{x},0) = u_0(\mathbf{x}), \qquad \mathbf{x} \in \Omega, \tag{2}$$

and suitable boundary conditions. The condition $D_j(u) = b'_j(u) \geq 0$, $j = 1,\ldots,d$ is needed to make the equation formally parabolic. Whenever $D_j(u) = 0$, for some $u \in \mathbb{R}$, is said that the equation degenerates at that $u-$level, since it ceases to be strictly parabolic. For $b_j(u) = u^m$, $m > 1$, and $S(\mathbf{x},t,u) = 0$, the equation (1) is called porous medium equation(PME) [36]. As is said in [36], if the initial condition has compact support, then (1) and (2) have at most one weak solution that has compact support in space (for more detail, see Theorem 5.3 of [36]). The degenerate parabolic equations appear often in applications, for example, spread of viscous fluids [11], mathematical biology [7], groundwater [6], heat transfer [16], flow of an isentropic gas through a porous medium [36], and other fields.

Methods for approximating solutions to (1) have attracted a lot of attention in recent years, for example, local discontinuous Galerkin finite element method [39], relaxation scheme [9], and kinetic scheme [4] have been proposed to find its solution. Also, Liu, Shu, and Zhang [26] proposed two different formulations of WENO schemes for solving nonlinear degenerate parabolic equations. The first formulation approximates directly the second derivative term by using a conservative flux difference, while the second formulation is similar to the formulation of the local discontinuous Galerkin (LDG) schemes [10].

As we know, WENO schemes were introduced in the literature to approximate hyperbolic conservation laws. Also, WENO schemes are based upon the essentially nonoscillatory (ENO) schemes [17]. The first version of WENO schemes was introduced in finite volume formulation for the one-dimensional conservation laws in 1994 by Liu, Osher, and Chan [25]. In 1996, Jiang and Shu observed that the ENO stencil selection is sensitive to the round-off perturbation near zeros of the solution and its derivatives [21]. Then, they improved the ENO reconstruction and proposed a general framework for designing arbitrary order accurate finite difference WENO schemes, which are more efficient for multidimensional calculations.

The degenerate parabolic equation (1) may have discontinuous solution, possible existence of sharp fronts, and finite speed of propagation of wave fronts [36]. The degenerate parabolic equations usually have features similar to those of a hyperbolic conservation laws. Therefore for solving (1), it is reasonable to generalize numerical methods for solving hyperbolic conservation laws, such as the WENO schemes [26].

In this paper, a WENO finite difference scheme is proposed by following the idea of the local discontinuous Galerkin method [10]. First rewrite (1) in one-dimensional case as a nonlinear system of first order equations

$$\begin{aligned} &u_t = v_x + S(x,t,u) \\ &v = b(u)_x, \end{aligned}$$

and then to solve the two first order equations using a fifth order finite difference WENO method for conservation laws. This approach has the advantage of simplicity, and it is easier to generalize it to higher than second order PDEs.

The structure of this paper is as follows: In section 2, we describe by detailing the construction and implementation of the scheme, for nonlinear degenerate parabolic equations. In section 3, the results of numerical experiments conducted with the proposed scheme are given. Some remarks are made in section 4.

2 Nonoscillatory reconstruction to the second derivative

In this section, a fifth-order WENO finite difference scheme for multidimensional nonlinear degenerate parabolic equation is described. Using the dimension by dimension approach [32], the presented scheme in this paper for multidimensional problems is implemented.

2.1 General framework

Consider the one-dimensional nonlinear degenerate parabolic equation

$$\begin{aligned} &u_t = b(u)_{xx} + S(x,t,u), \qquad && (x,t) \in \Omega \times (0,\infty), \\ &u(x,0) = u_0(x), \end{aligned} \tag{3}$$

satisfying suitable boundary conditions.

This section describe the formulation of a WENO finite difference scheme by following the idea of the local discontinuous Galerkin method [10]. First, the equation (3) is considered as a nonlinear system of first order equations

$$u_t = v_x + S(x,t,u) \tag{4}$$

$$v = b(u)_x, \tag{5}$$

and then to solve the two first order equations, respectively, using the fifth order finite difference WENO method for conservation laws [1]. This approach has the advantage of simplicity, and it is easy to generalize it to higher than second order PDEs. The effective stencil, which is a composition of two successive WENO procedures, is also wider in comparison with the approach that is said in [2]. To make the final effective stencil smaller, a right-biased stencil for (4) followed by a left-biased stencil for (5) is used.

2.2 The fifth order finite difference WENO method

Assuming that the cell averages $\bar{u}_j$ are known at time t, van Leer [35] proposed a piecewise polynomial reconstruction $u(x,t) = \sum_j R_j(x)\chi_j(x)$ to solve 1-dimension hyperbolic conservation laws, where $\chi_j(x)$ is the characteristic function of the cell $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$. Let $U(x)$ be the primitive function of $u(x,t)$; that is, $U(x) = \int_{-\infty}^{x} u(\xi,t)d\xi$. To begin, an optimal polynomial is selected, which is denoted by $p_{\text{opt},j}(x)$, to approximate $u(x,t)$ on cell I_j of degree four on the five-cells central stencil $S \equiv \{I_{j-2}, \ldots, I_{j+2}\}$. In [30, 32], this optimum polynomial is obtained. First, $U(x)$ is interpolated at the cell boundaries $\{x_{j-\frac{5}{2}}, \ldots, x_{j+\frac{5}{2}}\}$ by using Newton's interpolation formula. Afterwards, differentiating the new polynomial, we get

$$p_{\text{opt},j}(x) = \sum_{i=1}^{5} U[x_{j-\frac{5}{2}}, \ldots, x_{j+i-\frac{5}{2}}] \sum_{m=0}^{i-1} \prod_{l=0,l\neq m}^{i-1} (x - x_{j+l-\frac{5}{2}}),$$

where $U[\cdots]$ is a divided difference of the function $U(x)$. High-order ENO reconstructions decrease damping of the solution, but generate significant oscillations when solving hyperbolic systems unless costly characteristic decompositions are used. For solving this problem, three supplementary polynomials are defined, $\{p_0^j(x), p_1^j(x), \text{and } p_2^j(x)\}$, approximating $u(x)$ with a lower accuracy on I_j. The three-cells stencils of ENO3 are obtained by either choosing r cells to the left and s cells to the right of I_j. Then, the three-cells stencils of ENO3 can be explicitly determined.

$$S^r \equiv \{I_{j-r}, \ldots, I_{j+s}\}, \qquad (r,s) = (0,2), (1,1), \text{ and } (2,0).$$

Having obtained $p_{opt}(x)$ as above, $p^{(r),j}(x)$ is obtained from cell boundaries of the stencil S^r

$$p^{(r),j}(x) = \sum_{i=1}^{3} U[x_{j-r-\frac{1}{2}}, \ldots, x_{j-r+i-\frac{1}{2}}] \sum_{m=0}^{i-1} \prod_{l=0,l\neq m}^{i-1} (x - x_{j-r+l-\frac{1}{2}}). \quad (6)$$

In order to obtain a more efficient reconstruction, the ENO3 reconstruction is modified. We interpolate equation (6) over an additional point lying within the same stencil and the same number of selection steps as ENO3. The new reconstruction starts by seeking a new high-order interpolating polynomial, $p_r^j(x)$, such that equation (6) passes over the additional point $x_j \in I_j$,

$$p_r^j(x) = p^{(r),j}(x) + U[x_{j-r-\frac{1}{2}}, \ldots, x_{j-r+\frac{5}{2}}, x_j] \sum_{m=0}^{3} \prod_{l=0,l\neq m}^{3} (x - x_{j-r+l-\frac{1}{2}}), \; r = 0,1,2. \quad (7)$$

The divided difference $U[x_{j-r+\frac{5}{2}}, x_j]$, for $r = 0, 1$, and 2 of (7), is given by

$$\begin{aligned} U[x_{j-r+\frac{5}{2}}, x_j] &= \tfrac{1}{x_j - x_{j-r+\frac{5}{2}}} \left[\int_{-\infty}^{x_j} u(\xi) d\xi - \int_{-\infty}^{x_{j-r+\frac{5}{2}}} u(\xi) d\xi \right] \\ &= \tfrac{1}{x_j - x_{j-r+\frac{5}{2}}} \int_{x_{j-r+\frac{5}{2}}}^{x_j} \left(\textstyle\sum_j L_j(x) \chi_j(x) \right) dx. \end{aligned} \tag{8}$$

As it can be seen from the equation (8), a polynomial that retains information within the cell I_j is needed. The NT scheme [29] is based on the first-order Lax–Friedrichs scheme [13] and it involves the reconstruction of piecewise-linear MUSCL-type interpolates from piecewise constant data and uses nonlinear limiters to prevent oscillations

$$L_j(x) = \bar{u}_j + (x - x_j) \frac{1}{h} u'_j, \quad x \in I_j. \tag{9}$$

Integrating over (9), the divided differences in (8) are given by

$$\begin{aligned} U[x_{j-r+\frac{5}{2}}, x_j] = \frac{1}{5-2r} \Bigg(&\bar{u}_j - (r^2 - r - 2) \bar{u}_{j+1} + \\ &(r^2 - 3r + 2) \bar{u}_{j+2} + \frac{1}{4} u'_j \Bigg), \quad r = 0, 1, 2. \end{aligned}$$

For the numerical derivative u'_j, the UNO limiter [18] is chosen which is given by

$$\begin{aligned} u'_j = \mathbf{MM} \Bigg(&\Delta \bar{u}_{j-\frac{1}{2}} + \frac{1}{2} \mathbf{MM}(\Delta^2 \bar{u}_{j-1}, \Delta^2 \bar{u}_j), \\ &\Delta \bar{u}_{j+\frac{1}{2}} - \frac{1}{2} \mathbf{MM}(\Delta^2 \bar{u}_j, \Delta^2 \bar{u}_{j+1}) \Bigg), \end{aligned}$$

where $\Delta^2 \bar{u}_j = \Delta \bar{u}_{j+\frac{1}{2}} - \Delta \bar{u}_{j-\frac{1}{2}}$, and $\Delta \bar{u}_{j+\frac{1}{2}} = \bar{u}_{j+1} - \bar{u}_j$. Here, the MinMod limiter ($\mathbf{MM}$) is defined by

$$\mathbf{MM}(x_1, x_2, \ldots) = \begin{cases} \min_p \{x_p\} & \text{if } x_p > 0, \\ \max_p \{x_p\} & \text{if } x_p < 0, \\ 0 & \text{otherwise.} \end{cases}$$

2.3 Essentially nonoscillatory reconstruction

Spurious oscillations can appear in the numerical solution, if the big stencil which defines $p_{\mathrm{opt},j}(x)$, contains a discontinuity or large gradients. To avoid such a problem a WENO procedure is designed that smoothly adapts the stencil in the neighbourhood of the singularity.

The principle of CWENO (central WENO) procedure which is defined in [8, 24] is extended, and an ENO interpolant is constructed as a convex combination of polynomials that are based on different stencils; that is,

$$R_j(x) \equiv \sum_i w_i^j p_i^j(x), \qquad w_i^j \geq 0, \ \sum_i w_i^j = 1, \ i \in \{0, 1, 2, c\}. \tag{10}$$

Here, $R_j(x)$ is the nonoscillatory reconstruction on I_j. $p_c^j(x)$ is a polynomial which is defined on stencil S, also the polynomials $p_0^j(x), p_1^j(x)$, and $p_2^j(x)$ are computed in (7). In order to simplify the notations, the upper index j will be omitted; remembering that the weights and the four polynomials change from cell to cell. Now, the polynomial $p_c(x)$ is computed such that the convex combination (10), will be fifth-order accurate in smooth regions. It must, therefore, satisfy the following equation:

$$p_{\text{opt}}(x) = C_0 p_0(x) + C_1 p_1(x) + C_2 p_2(x) + C_c p_c(x), \quad \sum_i C_i = 1 \ i \in \{0, 1, 2, c\}, \tag{11}$$

where constants C_i represent linear or ideal weights for (10). A straightforward calculation shows that any symmetric choice of constants C_i in (11) provides the desired accuracy. This property must be contrasted with classical upwind WENO schemes. Therefore, we make the choice $C_0 = C_2 = \frac{1}{8}, C_1 = \frac{1}{4}$, and $C_c = \frac{1}{2}$. Then, the polynomial $p_c(x)$ is obtained explicitly as follows:

$$p_c(x) = [p_{\text{opt}}(x) - C_0 p_0(x) - C_1 p_1(x) - C_2 p_2(x)]/C_c \qquad \forall x \in I_j. \tag{12}$$

In order to complete the reconstruction of $R_j(x)$, the ideal weights $C_i, \ i \in \{0, 1, 2, c\}$, are changed to nonlinear weights w_i, with objective of maintaining the fifth-order accuracy for smooth solutions and nonoscillatory performance near discontinuities. In smooth regions for achieving the optimal interpolation (11), the weights w_i must smoothly converge to the ideal weights C_i as h approaches zero. Also, in regions where a discontinuity does exist, the weights should effectively remove the contribution of stencils that contain the discontinuity.

2.4 The nonoscillatory weights

In order to fully determine the scheme, it is required to specify the nonoscillatory weights,. Then, following [21, 32], the nonlinear weights are written as

$$w_i = \frac{\alpha_i}{\sum_k \alpha_k}, \quad \alpha_i = \frac{C_i}{(\epsilon + IS_i)^2}, \quad i,k \in \{0,1,2,c\}. \tag{13}$$

The constants $C_i,\ i \in \{0,1,2,c\}$, in Eq. (13) are chosen to be the same as in the equation (11); that is, $C_0 = C_2 = \frac{1}{8}, C_1 = \frac{1}{4}$, and $C_c = \frac{1}{2}$.

The *smoothness indicators*, IS_i, are responsible for detecting large gradients or discontinuities and automatically switch to the stencil that generates the least oscillatory reconstruction in such cases. The smoothness indicator is defined as [21]

$$IS_i = \sum_k h^{2k-1} \times \int_{I_j} \left(\frac{d^k}{dx^k} p_i(x)\right)^2 dx, \qquad i \in \{0,1,2,c\}.$$

The constant ϵ, taken to prevent the denominator from vanishing, can range from 10^{-5} to 10^{-7}, [21]. As is said in [12] and [19], different values of ϵ change the order of convergence of the scheme and yield different sharpness near discontinuities. In this paper, $\epsilon = 10^{-6}$ is used. The smoothness indicators of smooth and discontinuous stencils yield $IS_r = \mathcal{O}(h^2)$ and $IS_r = \mathcal{O}(1)$, respectively. A direct computation based on equations (7) and (12) yields:

$$\begin{aligned} IS_r =& \frac{1}{8100}(9\bar{u}_j - 12\bar{u}_{j+(1-r)} + 3\bar{u}_{j+2(1-r)} + (1-r)96u'_j)^2 \\ &+ \frac{13}{2700}(57\bar{u}_j - 66\bar{u}_{j+(1-r)} + 9\bar{u}_{j+2(1-r)} + (1-r)48u'_j)^2 \\ &+ \frac{781}{162000}(72\bar{u}_j - 96\bar{u}_{j+(1-r)} + 24\bar{u}_{j+2(1-r)} + (1-r)48u'_j)^2, \quad r = 0,2, \end{aligned}$$

$$\begin{aligned} IS_1 =& \frac{9}{2916}(\bar{u}_{j-1} - \bar{u}_{j+1} - 16u'_j)^2 + \frac{13}{12}(\bar{u}_{j-1} - 2\bar{u}_j + \bar{u}_{j+1})^2 \\ &+ \frac{449856}{58320}(\bar{u}_{j-1} - \bar{u}_{j+1} + 2u'_j)^2, \end{aligned}$$

and

$$\begin{aligned} IS_c =& (\frac{23957}{31500}\bar{u}_{j+2} - \frac{690709}{126000}\bar{u}_{j+1} + \frac{34913}{8400}\bar{u}_j - \frac{4051}{126000}\bar{u}_{j-1} - \frac{20591}{126000}\bar{u}_{j-2})\bar{u}_{j+2} \\ &+ (\frac{1478263}{126000}\bar{u}_{j+1} - \frac{7423}{350}\bar{u}_j + \frac{68419}{21000}\bar{u}_{j-1} - \frac{4051}{126000}\bar{u}_{j-2})\bar{u}_{j+1} \\ &+ (\frac{143239}{8400}\bar{u}_j - \frac{7423}{350}\bar{u}_{j-1} + \frac{34913}{8400}\bar{u}_{j-2})\bar{u}_j \\ &+ (\frac{1478263}{126000}\bar{u}_{j-1} - \frac{690709}{126000}\bar{u}_{j-2})\bar{u}_{j-1} + (\frac{23957}{31500}\bar{u}_{j-2})\bar{u}_{j-2} \\ &+ (\frac{2464}{1125}u'_j - \frac{649}{375}\bar{u}_{j-2} + \frac{2486}{375}\bar{u}_{j-1} - \frac{2486}{375}\bar{u}_{j+1} + \frac{649}{375}\bar{u}_{j+2})u'_j. \end{aligned}$$

2.5 The new scheme for nonlinear degenerate parabolic PDEs

From equations (4) and (5), it is required to approximate $b(u)_x$ and v_x. The final form of the left- and right-biased points are given by

$$u^-_{j+\frac{1}{2}} = \sum_{i\in\{0,1,2,c\}} w_i \times u^{-(i)}(x_{j+\frac{1}{2}}),$$

$$v^+_{j-\frac{1}{2}} = \sum_{i\in\{0,1,2,c\}} w_i \times v^{+(i)}(x_{j-\frac{1}{2}}),$$

where

$$u^{-(0)}_{j+\frac{1}{2}} := p_0(x_{j+\frac{1}{2}}) = (22\bar{u}_j + 9\bar{u}_{j+1} - \bar{u}_{j+2} + 8u'_j)/30,$$

$$u^{-(1)}_{j+\frac{1}{2}} := p_1(x_{j+\frac{1}{2}}) = (\bar{u}_{j-1} + 15\bar{u}_j + 2\bar{u}_{j+1} + 8u'_j)/18,$$

$$u^{-(2)}_{j+\frac{1}{2}} := p_2(x_{j+\frac{1}{2}}) = (-2\bar{u}_{j-2} + 13\bar{u}_{j-1} + 19\bar{u}_j + 24u'_j)/30,$$

$$\begin{aligned} u^{-(c)}_{j+\frac{1}{2}} &:= p_c(x_{j+\frac{1}{2}}) \\ &= (30\bar{u}_{j-2} - 205\bar{u}_{j-1} + 291\bar{u}_j + 277\bar{u}_{j+1} - 33\bar{u}_{j+2} - 176u'_j)/360, \end{aligned}$$

and

$$v^{+(0)}_{j-\frac{1}{2}} := p_0(x_{j-\frac{1}{2}}) = (19\bar{v}_j + 13\bar{v}_{j+1} - 2\bar{v}_{j+2} - 24v'_j)/30,$$

$$v^{+(1)}_{j-\frac{1}{2}} := p_1(x_{j-\frac{1}{2}}) = (2\bar{v}_{j-1} + 15\bar{v}_j + \bar{v}_{j+1} - 8v'_j)/18,$$

$$v^{+(2)}_{j-\frac{1}{2}} := p_2(x_{j-\frac{1}{2}}) = (-\bar{v}_{j-2} + 9\bar{v}_{j-1} + 22\bar{v}_j - 8v'_j)/30,$$

$$\begin{aligned} v^{+(c)}_{j-\frac{1}{2}} &:= p_c(x_{j-\frac{1}{2}}) \\ &= (-33\bar{v}_{j-2} + 277\bar{v}_{j-1} + 291\bar{v}_j - 205\bar{v}_{j+1} + 30\bar{v}_{j+2} + 176v'_j)/360. \end{aligned}$$

Therefore,

$$\begin{cases} \bar{v}_j = \frac{b(u^-_{j+\frac{1}{2}}) - b(u^-_{j-\frac{1}{2}})}{h}, \\ \frac{d\bar{u}_j(t)}{dt} = \frac{v^+_{j+\frac{1}{2}} - v^+_{j-\frac{1}{2}}}{h} + S(x_j, t, u_j) = F(\bar{u}_j(t)), \end{cases} \tag{14}$$

and the semidiscrete scheme (14) is discretized in time by a third order TVD RungeKutta method [33], which is given by

$$
\begin{aligned}
\bar{u}_j^{(1)} &= \bar{u}_j^n + \Delta t F(\bar{u}_j^n),\\
\bar{u}_j^{(2)} &= \tfrac{3}{4}\bar{u}_j^n + \tfrac{1}{4}\bar{u}_j^{(1)} + \tfrac{1}{4}\Delta t F(\bar{u}_j^{(1)}), \qquad (15)\\
\bar{u}_j^{n+1} &= \tfrac{1}{3}\bar{u}_j^n + \tfrac{2}{3}\bar{u}_j^{(2)} + \tfrac{2}{3}\Delta t F(\bar{u}_j^{(2)}).
\end{aligned}
$$

3 Numerical results

This section gives some numerical examples to confirm the accuracy, robustness, and high-resolution properties of the proposed method. Example 1 is three-dimensional heat equation with a known exact solution and is used for checking the accuracy and order of convergence [15]. Example 2 is the two-dimensional porous medium equation [26]. In Example 3, a nonlinear problem is presented and suggested as test for multidimensional nonlinear convection-diffusion problems [3, 22]. Next, in Example 4, a scalar two-dimensinal reaction-diffusion equation is considered [7]. Example 5 taken from [20, 23], is presented to cover the case strongly degenerate parabolic equation. Finally, Example 6 is considered to solve a Turing model of biological pattern formation [7]. This example will demonstrate that the new technique also handles systems of degenerate equations. The constant CFL number $c = 0.3$ is considered. Then, the time step is determined by

$$\max_u |b'(u)| \frac{\Delta t}{h^2} = 0.3.$$

Also, in the following, CWENO5 is used to denote the new proposed method in the current paper.

Example 1. For $\mathbf{x} = (x, y, z)$, $S(\mathbf{x}, t, u) \equiv 0$, and $b_j(u) = u$, $j = 1, 2, 3$, consider the three-dimensional linear initial-value problem with periodic boundary condition [15],

$$\begin{cases} u_t = u_{xx} + u_{yy} + u_{zz}, \ \mathbf{x} \in \Omega = (-\pi, \pi)^3, \ t > 0, \\ u(x, y, z, 0) = \sin(x)\sin(y)\sin(z). \end{cases}$$

The closed analytical form solution for this problem is

$$u(x, y, z, t) = \exp(-t)\sin(x)\sin(y)\sin(z).$$

The error is computed in the discrete L_∞ and L_1 norms and, respectively, defined by

$$||u||_\infty = \max_{i,j,k} |u_{i,j,k}|,$$

$$||u||_1 = \sum_{i,j,k} |u_{i,j,k}| \Delta x^3.$$

Table 1: Errors for Example 1 at T=2

N	L_∞ error	L_∞ order	L_1 error	L_1 order	CPU
10	0.192(-2)	-	0.847(-3)	-	08.97
20	0.601(-4)	5.000	0.266(-4)	4.999	15.01
40	0.189(-5)	4.999	0.822(-6)	5.021	32.84
80	0.579(-7)	5.030	0.257(-7)	5.000	68.10
160	0.163(-8)	5.151	0.804(-9)	5.000	142.90

Table 2: Errors and orders of convergence for Example 1 with different linear weights at T=2. Top:$C_0 = C_1 = C_2 = C_c = \frac{1}{4}$, Middle:$C_0 = C_2 = \frac{1}{16}, C_1 = \frac{1}{2}, C_c = \frac{3}{8}$, Bottom:$C_0 = \frac{1}{4}, C_1 = C_2 = \frac{1}{8}, C_c = \frac{1}{2}$.

N	L_∞ error	L_∞ order	L_1 error	L_1 order	CPU
10	0.115(-2)	-	0.788(-3)	-	09.12
20	0.400(-4)	4.845	0.295(-4)	4.739	15.15
40	0.110(-5)	5.184	0.910(-6)	5.018	33.01
80	0.347(-7)	4.986	0.290(-7)	4.971	70.11
160	0.107(-8)	5.019	0.863(-9)	5.070	139.98
10	0.149(-2)	-	0.102(-2)	-	10.01
20	0.453(-4)	5.039	0.341(-4)	4.902	14.93
40	0.190(-5)	4.575	0.971(-6)	5.134	30.99
80	0.338(-7)	5.812	0.281(-7)	5.110	73.12
160	0.103(-8)	5.036	0.877(-9)	5.001	138.88
10	0.171(-2)	-	0.145(-2)	-	11.55
20	0.148(-3)	3.530	0.139(-3)	3.382	16.01
40	0.170(-4)	3.122	0.182(-4)	2.933	32.18
80	0.125(-5)	3.765	0.201(-5)	3.178	69.87
160	0.873(-7)	3.839	0.156(-6)	3.687	140.11

In Table 1, the respective errors and convergence rates by CWENO5 are given. The results verify the fifth-order accuracy both in the L_∞ and in the L_1 norms. The L_1 and L_∞ errors and orders of convergence by CWENO5 with different linear weights are reported in Table 2. As is expected, any symmetric choice of constants C_i enables the fifth-order accuracy.

Example 2. Consider the porous medium equation (PME) as [26]

$$u_t = \Delta(u^m), \qquad m > 1, \tag{16}$$

where $u = u(\mathbf{x}, t)$, $\mathbf{x} \in \Omega \subset \mathbb{R}^d$. This equation describes various diffusion processes, such as the flow of an isentropic gas through a porous medium where u is the density of the gas required to be non-negative and u^{m-1} is the pressure of the gas. The PME equation degenerates at points where $u = 0$,

Table 3: Errors and orders of convergence for Example 2 at T=2. Top to Bottom respectively: $m=2,\ m=4,\ m=6$, and $m=8$.

N	L_∞ error	L_∞ order	L_1 error	L_1 order	CPU
50	0.012(-1)	-	0.429(-1)	-	10.62
100	0.010(-1)	0.263	0.172(-1)	1.318	19.88
200	0.006(-1)	0.737	0.036(-1)	2.256	35.17
400	0.002(-1)	1.585	0.006(-1)	2.585	78.41
50	0.117(-1)	-	0.248	-	13.51
100	0.083(-1)	0.495	0.639(-1)	1.956	23.23
200	0.069(-1)	0.266	0.256(-1)	1.319	37.11
400	0.054(-1)	0.353	0.092(-1)	1.476	86.12
50	0.216(-1)	-	0.353	-	16.55
100	0.232(-1)	-0.103	0.135	1.386	27.51
200	0.179(-1)	0.374	0.725(-1)	0.896	41.68
400	0.134(-1)	0.417	0.248(-1)	1.547	95.90
50	0.337(-1)	-	0.583	-	23.55
100	0.248(-1)	0.442	0.201	1.536	39.77
200	0.213(-1)	0.219	0.752(-1)	1.418	53.13
400	0.218(-1)	-0.033	0.368(-1)	1.031	115.55

resulting in the phenomenon of finite speed of propagation and sharp fronts. The classical solutions to the PME may not exist in general, even if the initial condition is smooth. The closed analytical form which is a weak solution for PME, is obtained by Zel'dovich and Kompaneets [38] and Barenblatt [5] in years 1950 and 1952, respectively. The solution of [5] has the following explicit form

$$u(\mathbf{x},t) = t^{-\alpha}\left[\left(1-k|\mathbf{x}|^2 t^{-\frac{2\alpha}{d}}\right)_+\right]^{\frac{1}{m-1}}, \tag{17}$$

where $\alpha = \frac{d}{(m-1)d+2}$, $k = \frac{\alpha(m-1)}{2md}$, and $u_+ = \max(u,0)$. Now, we put $d=2$; therefore, Barenblatt solution has no derivative at the points of the circle $x^2+y^2 = \sqrt{\frac{4m}{\alpha(m-1)}}t^{\alpha}$, where $\alpha = \frac{1}{m}$. The CWENO5 scheme is used to solve the PME (16) for $m = 2,4,6$ and $m = 8$, where $\Omega = [-10,10]^2$. The initial condition is taken as the Barenblatt solution (17) at $t=1$, and the boundary condition is chosen to be $u=0$ on $\partial\Omega$ for $t>1$.

Table 3 indicates the magnitude of errors obtained with CWENO5 scheme for different mesh sizes. Figure 1 shows the results of CWENO5 scheme at the Final time $T=3$ with the computational domain Ω divided into 120 uniform cells. The CWENO5 scheme prevents the appearance of spurious solutions close to the circles. The contour plots of absolute difference between approximated solution and Barenblatt solution with $m=2,4$ are shown in Figure 1 top right and bottom right, respectively. The CWENO5 scheme is sufficiently

Table 4: Errors and orders of convergence for Example 2 ($m = 2$) with different linear weights at T=2. Top:$C_0 = C_1 = C_2 = C_c = \frac{1}{4}$, Middle:$C_0 = C_2 = \frac{1}{16}, C_1 = \frac{1}{2}, C_c = \frac{3}{8}$, Bottom:$C_0 = \frac{1}{4}, C_1 = C_2 = \frac{1}{8}, C_c = \frac{1}{2}$.

N	L_∞ error	L_∞ order	L_1 error	L_1 order	CPU
50	0.012(-1)	-	0.422(-1)	-	10.89
100	0.009(-1)	0.415	0.165(-1)	1.354	20.85
200	0.005(-1)	0.848	0.037(-1)	2.156	34.37
400	0.002(-1)	1.322	0.007(-1)	2.402	80.51
50	0.014(-1)	-	0.418(-1)	-	11.11
100	0.008(-1)	0.807	0.167(-1)	1.324	22.20
200	0.006(-1)	0.415	0.039(-1)	2.098	33.01
400	0.003(-1)	1.000	0.009(-1)	2.116	81.52
50	0.222(-1)	-	0.338	-	09.95
100	0.218(-1)	0.026	0.165	1.035	22.51
200	0.157(-1)	0.473	0.722(-1)	1.192	35.68
400	0.131(-1)	0.261	0.243(-1)	1.571	79.95

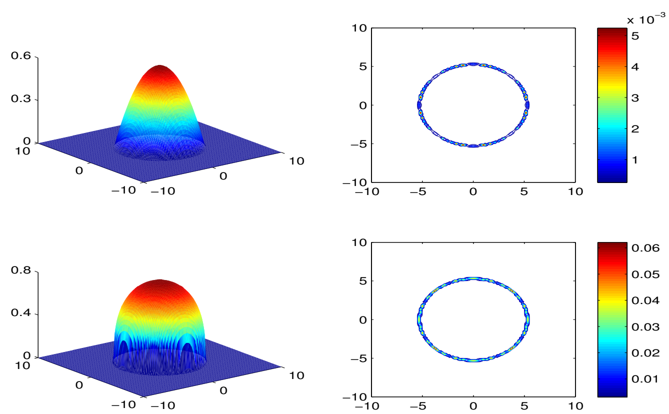


Figure 1: Example 2 (two-dimensional porous medium equation). Top left: PME with $m = 2$ and $N = 120$, Top right: 30 contour lines of $|u_{\text{Barenblatt}} - u|$ with $m = 2$, Bottom left: PME with $m = 4$ and $N = 120$, Bottom right: 30 contour lines of $|u_{\text{Barenblatt}} - u|$ with $m = 4$.

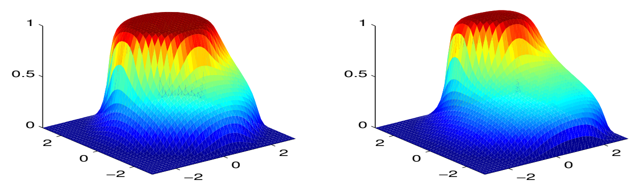


Figure 2: Example 3 (two-dimensional Buckley–Leverett-type equation). Left: numerical solution at $T = 0.5$ with $N = 40$, Right: numerical solution at $T = 1.0$ with $N = 40$.

accurate in the smooth regions, while on the nonsmooth regions absolute difference with analytical form solution of Barenblatt is of order 10^{-3} and 10^{-2} for $m = 2$ and $m = 4$, respectively.

Example 3. This example solves the Buckley–Leverett-type problem [3, 22]

$$u_t + f(u)_x + g(u)_y = \epsilon(u_{xx} + u_{yy})$$

with $\epsilon = 0.1$, the flux function of the form

$$g(u) = \frac{u^2}{u^2 + (1-u)^2},$$

$$f(u) = g(u)(1 - 5(1-u)^2).$$

The initial function for this example is

$$u(x, y, 0) = \begin{cases} 1, & (x - 0.25)^2 + (y - 0.25)^2 < 5, \\ 0, & otherwise. \end{cases}$$

This example includes gravitational effects in the x-direction. The numerical results at $T = 0.5$ and 1.0 are shown in Figure 2, with the computational domain $[-3, 3]^2$ divided into 40×40 uniform cells. The results compare well with those reported in [3]. Table 5 gives the respective errors and convergence rates by CWENO5. The numerical solutions compared with a "reference solution", obtained by WENO6 [26] with $N = 400$.

Example 4. This example solves the following initial-boundary-value problem for a scalar reaction-diffusion equation [7], where $\mathbf{x} = (x, y)$,

$$u_t = \Delta b(u) + S(\mathbf{x}, t, u), \tag{18}$$

$$u(\mathbf{x}, 0) = 0.5(1 + \sin(1.1(x - \cos(0.7y)))) \cos(0.5(y - \sin(1.3x))),$$

$$\nabla b(u) \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega \times [0, T]. \tag{19}$$

Table 5: Errors and orders of convergence for Example 3 at T=0.5

N	L_∞ error	L_∞ order	L_1 error	L_1 order	CPU
40	0.228(-3)	-	0.716(-4)	-	33.63
80	0.680(-5)	4.987	0.248(-5)	4.852	58.12
160	0.215(-6)	4.998	0.799(-7)	4.955	109.87
320	0.671(-8)	4.992	0.239(-8)	5.060	219.23

This problem may serve as a scalar prototype degenerate reaction-diffusion model. Here, the zero-flux boundary condition (19) implies that the reaction-diffusion domain is isolated from the external environment. For $S(\mathbf{x}, t, u) = S(u)$, (18) appears in [28] in an ecological setting, where u denotes the population density of a species, and $S(u)$ is its dynamics, where it is assumed that $S(0) = 0$ and $S'(0) \neq 0$. For example, $S(u) = u(1-u) - u^2/(1+u^2)$ corresponds to the population dynamics of the spruce band-worm [28] and models the growth of the population by a logistic expression and the rate of mortality due to predation by other animals. Authors of [7] modified this expression by a radial spatial factor, and used

$$
\begin{aligned}
&S(\mathbf{x}, t, u) = 10(\exp(-5r)u(1-u) + (\exp(-5r) - 1)\frac{u^2}{1+u^2}),\\
&r = \sqrt{(x-0.5)^2 + (y-0.5)^2},
\end{aligned}
$$

which means that the birth of individuals is concentrated near the centre $(0.5, 0.5)$, and mortality increases with increasing distance from the centre. Most standard spatial models of population dynamics simply assume that $b(u) = Du$, where the constant diffusion coefficient $D > 0$ measures the dispersal efficiency of the species under consideration. In this example, the strongly degenerate diffusion coefficient [7, 37] is used

$$
b(u) = \begin{cases} 0, & u \leq u_c, \\ D(u - u_c), & \textit{otherwise}, \end{cases}
$$

where $u_c > 0$ is an assumed critical value of u beyond which diffusion will take place. Here, we set $D = 1$ and $u_c = \frac{1}{2}$. The difficulty in the well-posedness analysis of the problem (18) lies in the boundary condition (19) when b is strongly degenerate. It is quite difficult to give a correct formulation of the zero-flux boundary conditions. The computational domain $[0, 1]^2$ is divided into 80×80 uniform cells. Figure 3 shows the numerical approximation at different time levels. The results compare well with those reported in [7].

Example 5. In this example, we focus on strongly degenerate parabolic problem. Consider the Burgers-like equation

$$
u_t + (u^2)_x + (u^2)_y = \epsilon(\nu(u)u_x)_x + \epsilon(\nu(u)u_y)_y.
$$

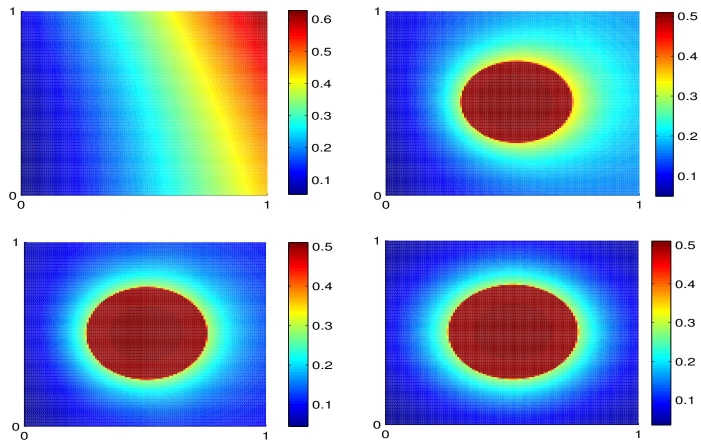


Figure 3: Example 4 (single-species reaction-diffusion model). From top left to bottom right, it shows the numerical solution at times $T = 0.0,\ 0.5,\ 1.0,\ 3.0$.

Table 6: Errors and orders of convergence for example 5 at T=0.5

N	L_∞ error	L_∞ order	L_1 error	L_1 order	CPU
40	0.223(-4)	-	0.118(-4)	-	30.73
80	0.715(-6)	4.963	0.362(-6)	5.000	53.98
160	0.214(-7)	5.059	0.112(-7)	5.014	100.11
320	0.599(-9)	5.162	0.334(-9)	5.054	213.23

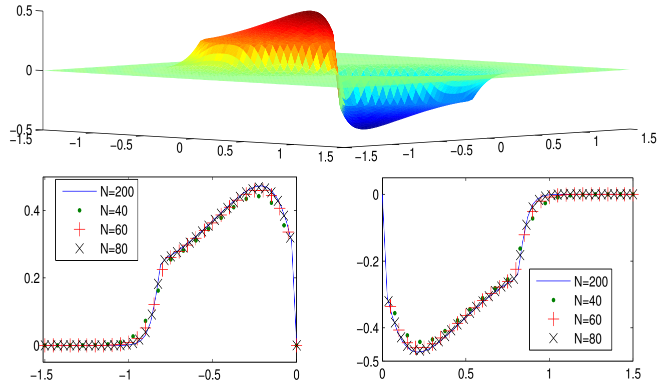


Figure 4: Example 5 (two-dimensional strongly degenerate parabolic problem). Top: numerical solution at $T = 0.5$ with $N = 80$, Bottom: one-dimensional cut along the y axis with different mesh sizes.

Here, $\epsilon = 0.1$ is considered and

$$\nu(u) = \begin{cases} 0, & |u| \leq 0.25, \\ 1, & |u| > 0.25. \end{cases}$$

In this example, zero boundary conditions are applied. The initial condition is equal to -1 and 1 inside two circles of radius 0.4 centred at $(0.5, 0.5)$ and $(-0.5, -0.5)$, respectively, and is zero elsewhere inside the square $[-1.5, 1.5]^2$. The L_∞ and L_1 errors of CWENO5 scheme are given in Table 6 for different mesh sizes. The computational domain is divided into 80×80 uniform cells. The numerical results obtained by CWENO5 scheme are presented in Figure 4, which compare well with those reported in [23,26]. The "reference solution" was obtained by WENO6 [26] scheme with $N = 400$.

Example 6. Turing [34] wrote his paper "The chemical basis of morphogenesis" in which he suggested an new theory. He hypothesized that the patterns, observed during embryonic development, occur in reaction to a spatial prepattern in biochemicals, which he termed morphogens. Cells would then answer to this prepattern by differentiating in a threshold-dependent way. Thus, Turing hypothesized that the patterns, observed in nature, such as pigmentation in animals, branching in trees, and skeletal structures are reflections of inhomogeneities in underlying biochemical signalling.

The system he considered took the form

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{D}\nabla^2\mathbf{u} + \mathbf{f}(\mathbf{u}).$$

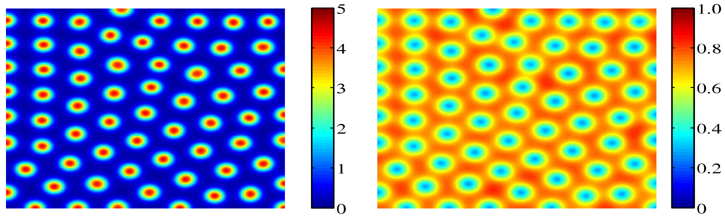


Figure 5: Example 6 (Turing's model for biological pattern formation). Left: u after 1113 simulated time units, Right: v after 1113 simulated time units.

Here, $\mathbf{u}$ is a vector of chemical condensations, $\mathbf{D} = \begin{bmatrix} D_u & 0 \\ 0 & D_v \end{bmatrix}$ is a matrix constant diffusion coefficients, and the nonlinear function $\mathbf{f}(\mathbf{u})$ is the reaction kinetics. Turing analysed the special case of equation (3) when there are two interacting morphogens, $\mathbf{u} = (u, v)^T$. They spread at rates D_u and D_v, respectively, and they also undergo reaction kinetics defined by $\mathbf{f}(\mathbf{u}) = (f(u, v), g(u, v))^T$. There are different forms of the reaction kinetics such as Gierer–Meinhardt [14] model

$$f(u, v) = c_1 - c_2 u + c_3 \frac{u^2}{(1 + ku^2)v}, \qquad g(u, v) = c_4 u^2 - c_5 v$$

and the Schnakenberg [31] model

$$f(u, v) = c_1 - c_{-1} u + c_3 u^2 v, \qquad g(u, v) = c_2 - c_3 u^2 v.$$

The parameters $\{c_{-1}, c_1, c_2, c_3\}$ represent the deterministic rates of the reactions. Now, consider the Schnakenberg model on the unit square domain $\Omega = [0, 1] \times [0, 1]$ with zero-flux boundary conditions. Let the parameter values be determined by $D_u = 1,\ D_v = 40,\ c_1 = 0.1,\ c_2 = 0.9, c_{-1} = 1$, and $c_3 = 1$. The initial conditions are random perturbations about the homogeneous steady state. This system has a uniform positive steady state (u^0, v^0), where $u^0 = c_1 + c_2$ and $v^0 = c_2/(c_1 + c_2)^2$. Figure 5 shows the formation of spot patterns. In this figure, the concentration values of u and v are illustrated which compare well with those reported in [27].

4 Conclusion

When applying the standard WENO idea, as Liu, Shu, and Zhang did in [26], in some cases (odd numbers of nodes in small stencils) the linear weights don't exist, and when they do exist, at least one of them is negative. In this research, another strategy was used to circumvent these problems. This strategy consists in applying the idea of the local discontinuous Galerkin method. First, the degenerate parabolic equation was considered as a nonlinear system of first order equations, and then this system was solved using a fifth-order finite difference weighted essentially nonoscillatory (WENO) method for conservation laws. The Symmetrical WENO procedure of [1] for solving this system was used. It was observed that this procedure generates a fifth-order method in smooth regions and remains essentially nonoscillatory near discontinuities. Numerical examples showed that the new scheme can obtain fifth-order accuracy in smooth regions and prevent the appearance of spurious solutions close to discontinuities.

Acknowledgement

The author is very thankful to reviewers for carefully reading the paper, their comments and suggestions have improved the quality of the paper.

References

1. Abedian, R., Adibi, H., and Dehghan, M. *A high-order symmetrical weighted hybrid ENO-flux limiter scheme for hyperbolic conservation laws*, Comput. Phys. Commun. 185 (2014), no. 1, 106–127.

2. Abedian, R., Adibi, H., and Dehghan, M. *A high-order weighted essentially non-oscillatory (WENO) finite difference scheme for nonlinear degenerate parabolic equations*, Comput. Phys. Commun. 184 (2013), no. 8, 1874–1888.

3. Acosta, C. D. and Bürger, R. *Difference schemes stabilized by discrete mollification for degenerate parabolic equations in two space dimensions*, IMA J. Numer. Anal. 32 (2012), no, 4, 1509–1540.

4. Aregba-Driollet, D., Natalini, R., and Tang, S. *Explicit diffusive kinetic schemes for nonlinear degenerate parabolic systems*, Math. Comp. 73 (2004), no. 245. 63–94.

5. Barenblatt, G. I. *On some unsteady motions of a liquid or a gas in a porous medium*, (Russian) Akad. Nauk SSSR. Prikl. Mat. Meh. 16, (1952). 67–78.

6. Barenblatt, G. I., Bertsch, M., Chertock, A. E., and Prostokishin, V. M. *Self-similar intermediate asymptotics for a degenerate parabolic filtration-absorption equation*, Proc. Natl. Acad. Sci. U.S.A. 97 (2000), no. 18, 9844–9848.

7. Bendahmane, M., Bürger, R., Ruiz-Baier, R., and Schneider, K. *Adaptive multiresolution schemes with local time stepping for two-dimensional degenerate reaction-diffusion systems*, Appl. Numer. Math. 59 (2009), no. 7, 1668–1692.

8. Capdeville, G. *A central WENO scheme for solving hyperbolic conservation laws on non-uniform meshes*, J. Comput. Phys, 227 (2008), no. 5, 2977–3014.

9. Cavalli, F., Naldi, G., Puppo, G., and Semplice, M. *High order relaxation schemes for nonlinear degenerate diffusion problems*, SIAM J. Numer. Anal. 45 (2007), no. 5, 2098–2119.

10. Cockburn, B. and Shu, C. -W. *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal. 35 (1998), no. 6, 2440–2463.

11. Davidovitch, B., Moro, E., and Stone, A. H. *Spreading of viscous fluid drops on a solid substrate assisted by thermal fluctuations*, Phys. Rev. Lett. 95 (2005), no. 24, 244–505.

12. Fedkiw, R., Merriman, B., and Osher, S. *Simplified discretization of systems of hyperbolic conservation laws containing advection equations*, J. Comput. Phys. 157 (2000), no. 1, 302–326.

13. Friedrichs, K. O. and Lax, P. D. *Systems of conservation equations with a convex extension*, Proc. Natl. Acad. Sci. U.S.A. 68 (1971), 1686–1688.

14. Gierer, A. and Meinhardt, H. *A theory of biological pattern formation*, Biol. Cybern. 12 (1972), no. 1, 30–39.

15. Hajipour, M. and Malek, A. *High accurate NRK and MWENO scheme for nonlinear degenerate parabolic PDEs*, Appl. Math. Model. 36 (2012), no. 9, 4439–4451.

16. Harley, C. *Asymptotic and dynamical analyses of heat transfer through a rectangular longitudinal fin*, J. Appl. Math. 2013, Art. ID 987327, 1–8.

17. Harten, A., Engquist, B., Osher, S., and Chakravarthy, S. *Uniformly high-order essentially nonoscillatory schemes, III*, J. Comput. Phys. 71 (1987), no. 2, 231–303.

18. Harten, A. and Osher, S. *Uniformly high-order accurate nonoscillatory schemes I*, SIAM J. Numer. Anal. 24 (1987), no. 2, 279–309.

19. Henrick, A. K., Aslam, T. D., and Powers, J. M. *Mapped weighted essentially nonoscillatory schemes: Achieving optimal order near critical points*, J. Comput. Phys. 207 (2005), no. 2, 542–567.

20. Holden, H., Karlsen, K. H., and Lie, K. -A. *Operator splitting methods for degenerate convection-diffusion equations II: Numerical examples with emphasis on reservoir simulation and sedimentation*, Comput. Geosci. 4 (2000), 287–322.

21. Jiang, G. S. and Shu, C. -W. *Efficient implementation of weighted ENO schemes*, J. Comput. Phys. 126 (1996), no. 1, 202–228.

22. Karlsen, K. H. and Risebro, N. H. *An operator splitting method for nonlinear convection-diffusion equations*, Numer. Math. 77 (1997), no. 3, 365–382.

23. Kurganov, A. and Tadmor, E. *New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations*, J. Comput. Phys. 160 (2000), no. 1, 241–282.

24. Levy, D., Puppo, G., and Russo, G. *Compact central WENO schemes for multidimensional conservation laws*, SIAM J. Sci. Comput. 22 (2000), no. 2, 656–672.

25. Liu, X. -D., Osher, S., and Chan, T. *Weighted essentially non-oscillatory schemes*, J. Comput. Phys. 115 (1994), no. 1, 200–212.

26. Liu, Y., Shu, C. -W., and Zhang, M. *High order finite difference WENO schemes for nonlinear degenerate parabolic equations*, SIAM J. Sci. Comput. 33 (2011), no. 2, 939–965.

27. Maini, P. K., Woolley, T. E., Baker, R. E., Gaffney, E. A., and Lee, S. S. *Turing's model for biological pattern formation and the robustness problem*, Interface focus 2 (2012), no. 4, 487–496.

28. Murray, J. D. *Mathematical Biology II: Spatial Models and Biomedical Applications*, Springer-Verlag, New York, 2003.

29. Nessyahu, H. and Tadmor, E. *nonoscillatory central differencing for hyperbolic conservation laws*, J. Comput. Phys. 87 (1990), no. 2, 408–463.

30. Peer, A. A. I., Dauhoo, M. Z., Gopaul, A., and Bhuruth, M. *A weighted ENO-flux limiter scheme for hyperbolic conservation laws*, Int. J. Comput. Math. 87 (2010), no. 15, 3467–3488.

31. Schnakenberg, J. *Simple chemical reaction systems with limit cycle behaviour*, J. Theoret. Biol. 81 (1979), no. 3, 389–400.

32. Shu, C. -W. *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, Advanced numerical approximation of nonlinear hyperbolic equations (Cetraro, 1997), 325–432, Lecture Notes in Math., 1697, Fond. CIME/CIME Found. Subser., Springer, Berlin, 1998.

33. Shu, C. -W. and Osher, S. *Efficient implementation of essentially nonoscillatory shock-capturing schemes*, J. Comput. Phys. 77 (1988), no. 2, 439–471.

34. Turing, M. A. *The chemical basis of morphogenesis*, Philos. Trans. Roy. Soc. London Ser. B 237 (1952), no. 641, 37–72.

35. van Leer, B. *Towards the ultimate conservative difference scheme V. A second-order sequel to Godunov's method*, J. Comput. Phys. 32 (1979), no. 1, 101–136.

36. Vázquez, J. L. *The Porous Medium Equation: Mathematical Theory*, Oxford University Press, USA, 2007.

37. Witelski, T. P. *Segregation and mixing in degenerate diffusion in population dynamics*, J. Math. Biol. 35 (1997), no. 6, 695–712.

38. Zel'dovich, Y. B. and Kompaneetz, A. S. *Towards a theory of heat conduction with thermal conductivity depending on the temperature*, in: Collection of Papers Dedicated to 70th Birthday of Academician A.F. Ioffe, Izd. Akad. Nauk SSSR, Moscow, 1950, pp. 61–71.

39. Zhang, Q. and Wu, Z. -L. *Numerical simulation for porous medium equation by local discontinuous Galerkin finite element method*. J. Sci. Comput. 38 (2009), no. 2, 127–148.

Iranian Journal of Numerical Analysis and Optimization
Vol. 8, No. 1, (2018), pp 63–80
DOI:10.22067/ijnao.v8i1.48147

A New Hybrid Method Based on Pseudo Differential Operators and Haar Wavelet to Solve ODEs

M.B. Ghaemi*, M. Jamalpour Birgani and E. Nabizadeh Morsalfard

Abstract

In this paper we present a new and efficient method by combining pseudo differential operators and Haar wavelet to solve the linear and nonlinear differential equations. The present method performs extremely well in terms of efficiency and simplicity.

Keywords: Haar wavelet; Fourier transform; Differential equations; Numerical solution; Pseudo differential operators.

1 Introduction

In this paper we have offered a method based on Haar wavelet in combination with pseudo differential operators to solve ODEs.

Recently, wavelet methods have proved useful as widely applied tools for the computation of numerical approximations. Different types of wavelets and approximating functions have been proposed for the numerical solution of differential equations using various methods [2,3,9,15–20], such as Galerkin and collocations methods, the filter bank methods [20], and the Kalman filters. In [8], the Haar wavelet method was applied for solving the boundary

*Corresponding author
Received 5 July 2015; revised 17 June 2017; accepted 11 November 2017
M.B. Ghaemi
Department of Mathematics, Iran University of Science and Technology, Tehran, Iran. e-mail: mghaemi@iust.ac.ir

M. Jamalpour Birgani
Department of Mathematics, Iran University of Science and Technology, Tehran, Iran. e-mail: m.j.birgani@gmail.com

E. Nabizadeh Morsalfard
Department of Mathematics, Iran University of Science and Technology, Tehran, Iran. e-mail: elmira_nabizadeh@yahoo.com

and initial value differential equations. Lepik in [12, 14] applied Haar wavelet to solve differential equations. Also, he used Haar wevelet for solving evolution and integral equations in [11, 13] and PDEs in [10].

pseudo differential operators are considered as natural extensions of linear partial differential operators. The study of pseudo differential operators grows out of research in 1960s on the singular integral operators. The pseudo differential operator is the generalization of linear partial differential operator, with roots entwined deep in solving differential equations in the classic and quantum mechanics [4, 5].

At first by applying Fourier inversion formula to the given differential equation we obtain its associated pseudo differential equation. Then in the obtained Pseudo differential operator we expand $\hat{f}$ (f is assumed to be the solution of ODE) into Haar series, and we employ the collocation method for the equation to get a numerical solution for ODE. By some numerical examples the advantages of our tool in comparison with Haar wavelet method are shown.

The paper is organized in 5 sections in the following way. In section 2 we recall some preliminaries. Some numerical solutions for some examples of linear and nonlinear differential equations by using our new method are presented in sections 3 and 4, respectively. Finally error analysis is given in section 5.

2 Preliminaries

2.1 Wavelets

Haar wavelets are the simplest wavelets among various types of wavelets. The Haar wavelet family on $[0,1]$ is defined as follows:

$$h_i(x) = \begin{cases} 1 & for\ x \in [\xi_1, \xi_2), \\ -1 & for\ x \in [\xi_2, \xi_3], \\ 0 & elsewhere, \end{cases}$$

where

$$\xi_1 = \frac{k}{m}, \quad \xi_2 = \frac{k+\frac{1}{2}}{m}, \quad \xi_3 = \frac{k+1}{m},$$

and $m = 2^j$, $j = 0, 1, \ldots, J$, indicate the level of the wavelet, and $k = 0, 1, \ldots, m-1$ is the translation parameter. The index i is calculated according to the formula $i = m+k+1$. The maximal value of i is $2M$ where $M = 2^J$. It is assumed that the value $i = 1$ corresponds to the scaling function of Haar wavelet $h_1(x) = 1$ for all $x \in [0,1]$.

2.2 Fourier Transform and Pseudo Differential Operators

One of the leading ideas in the theory of pseudo differential operators is to reduce the study of properties of a linear differential operator

$$P = \sum_{\alpha} c_\alpha(x)\partial_x^\alpha, \tag{1}$$

which is a polynomial in the derivatives ∂_x with constants c_α depending on x, to its *symbol*

$$p(x,\xi) = \sum_{\alpha} c_\alpha(x)(i\xi)^\alpha,$$

which is a polynomial in the phase variable $\xi \in \mathbb{R}$ with constants depending on the space variable x.

Fourier transform is the main tool in the theory of pseudo differential operators. Let f be integrable function; the Fourier transform and the inverse Fourier transform of f are defined as follows:

$$\hat{f}(\xi) := \mathfrak{F}[f](\xi) := \int_R e^{-ix\xi} f(x)dx$$

and

$$\mathfrak{F}^{-1}[\hat{f}](x) := \frac{1}{2\pi}\int_R e^{ix.\xi}\hat{f}(\xi)d\xi, \tag{2}$$

where $x, \xi \in R$. According to the inversion formula we have

$$\mathfrak{F}^{-1}[\mathfrak{F}(f)] \equiv \mathfrak{F}^{-1}[\hat{f}] = f. \tag{3}$$

Let P be the linear differential operator in (1) using the inversion formula,

$$\begin{aligned} Pf &= \sum_{\alpha} c_\alpha(x)\partial_x^\alpha \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix.\xi}\hat{f}(\xi)d\xi \\ &= \sum_{\alpha} c_\alpha(x)\frac{1}{2\pi}\int_{\mathbb{R}} (i\xi)^\alpha e^{ix.\xi}\hat{f}(\xi)d\xi \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix.\xi} p(x,\xi)\hat{f}(\xi)d\xi. \end{aligned}$$

Here p is considered as a symbol, and

$$(P(x,D_x)f)(x) := \int_{R^n} e^{ix.\xi} p(x,\xi)\hat{f}(\xi)d\xi$$

defines the associated *pseudo differential operator*. For more details one can refer to [1].

2.3 Combination of Haar Wavelet and Pseudo Differential Operators

Let $f \in L^2[0,1]$; then by (2) and (3), it can be approximated by

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ix\xi}\hat{f}(\xi)d\xi, \tag{4}$$

since $f \in L^2$ then $\hat{f} \in L^2$, and for level J we can write

$$\hat{f}(\xi) = \sum_{p=1}^{2M} a_p h_p(\xi),$$

where $a_p = \int_0^1 h_p(\xi)\hat{f}(\xi)d\xi,\ p = 1,\dots,2M$. By replacing this value in (4) we obtain

$$f(x) = \sum_{p=1}^{2M} a_p \int_0^1 e^{ix\xi}h_p(\xi)d\xi. \tag{5}$$

With differentiation of (5) we have

$$f'(x) = \sum_{p=1}^{2M} a_p \int_0^1 i\xi e^{ix\xi}h_p(\xi)d\xi, \qquad f''(x) = \sum_{p=1}^{2M} a_p \int_0^1 (i\xi)^2 e^{ix\xi}h_p(\xi)d\xi,$$

in other words

$$f^{(n)}(x) = \sum_{p=1}^{2M} a_p \int_0^1 (i\xi)^n e^{ixy}h_p(\xi)d\xi, \qquad for\ all\ \ n \in N. \tag{6}$$

Now consider the linear differential equation $\sum_{\alpha=0}^{n} A_\alpha(x)f^{(n)}(x)$, by replacing the obtained value for $f, f', \dots, f^{(n)}$ in this equation we obtain

$$\sum_{\alpha=0}^{n} A_\alpha(x)f^{(\alpha)}(x) = \sum_{p=1}^{2M} a_p \int_0^1 p(x,\xi)e^{ix\xi}h_p(\xi)d\xi,$$

where $p(x,\xi) = \sum_{\alpha=0}^{n} A_\alpha(x)(i\xi)^\alpha$. This new form of pseudo differential operators is our main tool in order to obtain numerical solution for linear and nonlinear differential equations in the next sections.

2.4 Solving Linear Differential Equations

Consider the differential equation of order n

$$\sum_{m=0}^{n} A_m(x)y^{(m)}(x) = f(x), \tag{7}$$

where $x \in [0,1]$ and $y^{(n-1)}(0) = c_n, \ldots, y(0) = c_0$. The interval $[0,1]$ is divided into $2M$ subintervals with the length $\Delta x = \frac{1}{2M}$.
According to the equality (6) we have

$$y^{(n)}(x) = \sum_{p=1}^{2M} a_p \int_0^1 (i\xi)^n e^{ixy} h_p(\xi) d\xi, \tag{8}$$

where a_p, with $p = 1, \ldots, 2M$ are unknown. Performing integration on (8) from 0 to x, yields

$$y^{(n-1)}(x) = \sum_{p=1}^{2M} a_p[\int_0^1 (i\xi)^{n-1} e^{ix\xi} h_p(\xi) d\xi - \int_0^1 (i\xi)^{n-1} h_p(\xi) d\xi] + c_n \tag{9}$$

continuing the integration of (9), we get

$$\begin{aligned} y^{(m)}(x) = & \sum_{p=1}^{2M} a_p[\int_0^1 (i\xi)^m e^{ixy} h_p(\xi) d\xi - x^{n-m-1} \int_0^1 (i\xi)^m h_p(\xi) d\xi] \\ & + \sum_{\alpha=0}^{n-m-1} \frac{1}{\alpha!} x^\alpha y^{(\alpha+m)}(0) \end{aligned} \tag{10}$$

for $m = 0, 1, \ldots, n-1$. Putting the values obtained for $y^{(n)}$, $y^{(n-1)}, \ldots, y$ and the collocation points

$$x_l = \frac{(l-0.5)}{2M}, \qquad l = 1, 2, \ldots, 2M, \tag{11}$$

in equation (7), the system below is obtained.

$$\begin{aligned} & \sum_{p=1}^{2M} a_p \sum_{m=0}^{n} S_m(x) = f(x_l) \\ & \qquad + \sum_{\nu=0}^{n-1} \sum_{\alpha=0}^{n-\nu-1} \frac{1}{\alpha!} (x_l)^\alpha y^{(\alpha+\nu)}(0) A_\nu(x_l), \end{aligned} \tag{12}$$

for $l = 1, \ldots, 2M$, where

$$S_m(x) = A_m(x)[\int_0^1 (i\xi)^m e^{ix\xi} h_p(\xi)d\xi - x^{n-m-1}\int_0^1 (i\xi)^m h_p(\xi)d\xi].$$

By calculating the coefficients a_p, $p = 1, \ldots, 2M$, and replacing them in (10), for $m = 0$, an approximation numerical solution is obtained. The matrix form of the system (12) is

$$a(\sum_{m=0}^{n} W_m) = F,$$

where W_m signifies $2M \times 2M$ matrices as follows:

$$W_m(p,l) = A_m(x_l)S_m(p,l), \qquad m = 0, 1, \ldots, n,$$

such that

$$S_m(p,l) = S_m(x_l),$$
$$S_n(p,l) = \int_0^1 (i\xi)^n e^{ix_l\xi} h_p(\xi)d\xi,$$

and F and a are $2M-$dimensional vectors

$$F(l) = f(x_l) + \sum_{\nu=0}^{n-1}\sum_{\alpha=0}^{n-\nu-1} \frac{1}{\alpha!}(x_l)^\alpha y^{(\alpha+\nu)}(0)A_\nu(x_l),$$
$$a = (a_1, \ldots, a_{2M}).$$

Also, the matrix representation of (1), for $m = 0$, is

$$y = aS_0 + T,$$

where the vector T is

$$T(l) = \sum_{\alpha=0}^{n-1} \frac{1}{\alpha!}(x_l)^\alpha y^{(\alpha)}(0), \qquad l = 1, \ldots, 2M.$$

Remark. For calculating the integrals, for example,

$$\int_0^1 (i\xi)^k e^{ix_l\xi} h_p(\xi)d\xi, \quad k = 0, 1, \ldots, n,$$

we can use of the points $\tau_l = \frac{(l-1)}{2M}$, $\quad l = 1, \ldots, 2M+1$. In other words,

$$\int_0^1 (i\xi)^k e^{ix_l\xi} h_p(\xi) = \sum_{s=1}^{2M} \int_{\tau_s}^{\tau_{s+1}} (i\xi)^k e^{ix_l\xi} h_p(\xi)d\xi.$$

Since for $\xi \in (\tau_s, \tau_{s+1})$ we have $h_p(\xi) = h_p(x_s)$; then

$$\int_0^1 (i\xi)^k e^{ix_l\xi} h_p(\xi) = \sum_{s=1}^{2M} h_p(x_s) \int_{\tau_s}^{\tau_{s+1}} (i\xi)^k e^{ix_l\xi} d\xi.$$

3 The Numerical Solution of Linear Differential Equations

In this section, some examples of linear differential equations are solved by using the new method. To get the error estimates, problems with the known exact solution $y_{ex}(x)$ are considered.

Example 1. Consider the differential equation

$$(x-1)y'(x) + (x^2+3)y(x) = x^4 + 7x^2 - 2x + 6, \tag{13}$$

where $y(0) = 0$. This equation has the exact solution $y_{ex}(x) = 2 + x^2$.
Now, we seek solutions for the equation (13) based on our hybrid method. In this equation

$$W_0(p,l) = (x_l^2 + 3)S_0(p,l),$$

$$W_1(p,l) = (x_l - 1)S_1(p,l),$$

$$F(l) = x_l^4 + 7x_l^2 - 2x_l + 6.$$

The obtained results and absolute errors, for some values of $J = 2$, are shown in Table 1 and Figure 1, respectively.

Example 2. Consider the differential equation

$$\begin{aligned} &(x-1)y''(x) + (x+1)y'(x) + y(x) = (2x^2 + 4x - 1)e^x, \\ &y(0) = 0, \qquad y'(0) = 1. \end{aligned} \tag{14}$$

This equation has the exact solution $y_{ex}(x) = xe^x$. In this equation

$$W_0(p,l) = S_0(p,l),$$

$$W_1(p,l) = (x_l + 1)S_1(p,l),$$

$$W_2(p,l) = (x_l - 1)S_2(p,l),$$

$$F(l) = (2x_l^2 + 4x_l - 1)e^{x_l} - (1 + 2x_l).$$

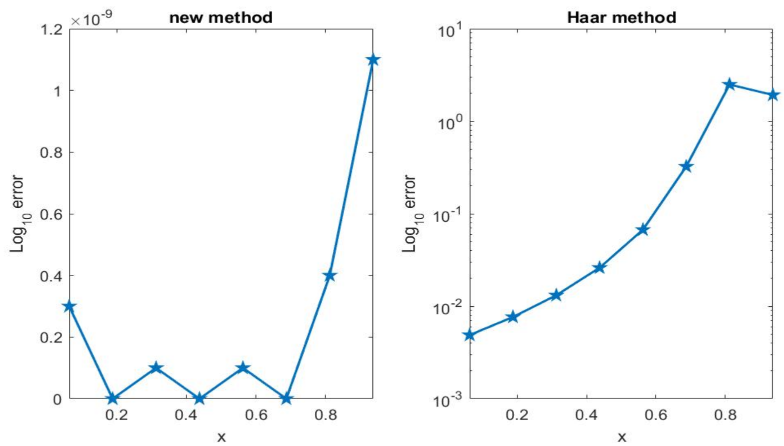


Figure 1: The absolute errors for example 1.

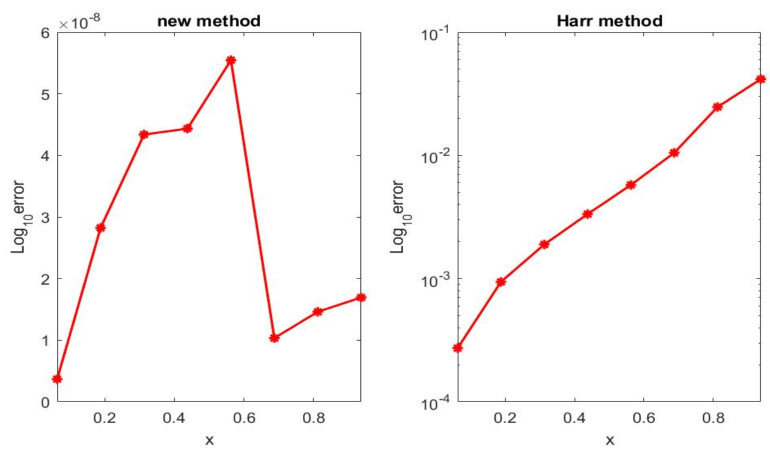


Figure 2: The absolute errors for example 2.

Table 1: Haar solution, Exact solution, The new method solution for equation (13)

$(x/16)$	Haar solution	Exact solution	The new method solution
1	2.0087906524	2.0039062500	2.0039062497
3	2.04280445967	2.0351562500	2.0351562500
5	2.11078811462	2.0976562500	2.0976562499
7	2.21748292022	2.1914062500	2.1914062500
9	2.38353209120	2.3164062500	2.3164062499
11	2.79651329536	2.4726562500	2.4726562500
13	0.16626536933	2.6601562500	2.6601562496
15	4.80205593409	2.8789062500	2.8789062511

The obtained results and absolute errors, for some values of J, are shown in Table 2 and Figure 2, respectively.

4 Nonlinear Differential Equations

In this section we adopt the new approach for nonlinear differential equations. A numerical solution of nonlinear differential equations can be obtained by using the new hybrid method. We consider a nonlinear form of differential equations as

$$K(x, u(x), u'(x), \ldots + u^{(n)}(x)) = f(x) \tag{15}$$

such that $x \in [0,1]$ and $u^{(n-1)}(0) = c_n, \ldots, u(0) = c_0$. K is nonlinear respect to one of the functions u, u', ... $u^{(n)}$ and the function f is given. Replacing x by x_l in the equation (15), for $l = 1, 2, \ldots, 2M$, gives

$$K(x_l, u(x_l), u'(x_l), \ldots + u^{(n)}(x_l)) = f(x_l),$$

where x_1, x_2, ..., and x_{2M} are the collocation points introduced in (11). Now, by applying (8) and (9) for the functions $u, u', \ldots, u^{(n)}$, we get a system with $2M$ equations for calculating the coefficients a_p. In general, this system is nonlinear, and we use the following approach to solve it.

Table 2: Haar solution, Exact solution, The new method solution for equation (14)

$(x/16)$	Haar solution	Exact solution	The new method solution
1	0.0668034	0.0665309	0.0665309
3	0.2271089	0.2261681	0.2261682
5	0.4290282	0.4271368	0.4271369
7	0.6809506	0.6776132	0.6776133
9	0.9929719	0.9872182	0.9872183
11	1.3777507	1.3672570	1.3672570
13	1.8556966	1.8309970	1.8309970
15	2.4355610	2.3939901	2.3939901

We assume that (15) is previously solved for the level $J-1$, where the number of collocation points is $M = 2^J$. For the next level, the number of collocation points is doubled (the value J is increased by one). The new values for $u(x), u'(x), \ldots, u^{(n)}(x)$ and a_p are estimated as

$$\hat{u}^{(n)(J)}(x) = \sum_{p=1}^{2M} \hat{a}_p^{(J)} \int_0^1 (i\xi)^n e^{ix\xi} h_p(\xi) d\xi, \tag{16}$$

and for $m = 0, \ldots, n-1$

$$\begin{aligned}\hat{u}^{(m)(J)}(x) = \sum_{p=1}^{2M} \hat{a}_p^{(J)} [\int_0^1 (i\xi)^m e^{ix\xi} h_p(\xi) d\xi - x^{n-m-1} \int_0^1 (i\xi)^m h_p(\xi) d\xi] \\ + \sum_{\alpha=0}^{n-m-1} \frac{1}{\alpha!} x^\alpha u^{(\alpha+m)}(0),\end{aligned} \tag{17}$$

where

$$\hat{a}_p^{(J)} = \begin{cases} \hat{a}_p^{(J-1)} & \text{for } p = 1, \ldots, M, \\ 0 & \text{for } p = M+1, \ldots, 2M. \end{cases} \tag{18}$$

These estimates are corrected by Newton method; the application of which for $l = 1, \ldots, 2M$ leads to the equation

$$\sum_{p=1}^{2M}(\frac{\partial K}{\partial a_p})\Delta a_p = -K(x_l, \hat{u}^{(J)}(x_l), \hat{u}'^{(J)}(x_l), \ldots, \hat{u}^{(n)(J)}(x_l)) + f(x_l), \quad (19)$$

where

$$\frac{\partial K}{\partial a_p} = \frac{\partial K}{\partial u}\frac{\partial u}{\partial a_p} + \frac{\partial K}{\partial u'}\frac{\partial u'}{\partial a_p} + \cdots + \frac{\partial K}{\partial u^{(n)}}\frac{\partial u^{(n)}}{\partial a_p}.$$

By solving (19), Δa_p, $p = 1, \ldots, 2M$, are calculated and equations (16), (17), and (18) are corrected as follows:

$$a_p^{(J)} = \hat{a}_p^{(J)} + \Delta a_p,$$

$$u^{(n)(J)}(x) = \sum_{p=1}^{2M} a_p^{(J)} \int_0^1 (i\xi)^n e^{ix\xi} h_p(\xi) d\xi, \quad (20)$$

and

$$u^{(m)(J)}(x) = \sum_{p=1}^{2M} a_p^{(J)}[\int_0^1 (i\xi)^m e^{ixy} h_p(\xi)d\xi - x^{n-m-1}\int_0^1 (i\xi)^m h_p(\xi)d\xi] + \sum_{\alpha=0}^{n-m-1}\frac{1}{\alpha!}x^\alpha u^{(\alpha+m)}(0), \qquad m = 0, \ldots, n-1.$$

We start the procedure by taking an initial solution of $\hat{a}_1^{(0)} = 1$ and $\hat{a}_2^{(0)} = 0$. These estimates are corrected by solving (19) for $J = 0$.

Example 3. Consider the the differential equation

$$u(x) + x^2 u(x) u'(x)^{\frac{1}{2}} = x^2(1 + \sqrt{2}x^{\frac{5}{2}}), \quad (21)$$

which has the exact solution $u_{ex}(x) = x^2$.
The obtained results for $J = 2$ are shown in Table 3 and Figures 3 and 4.

Example 4. Consider the the differential equation

$$x \cos x u'(x) + (x+3)u(x) + xu(x)^2 = x(1 + \sin x) + 3\sin x, u(0) = 0, \quad (22)$$

which has the exact solution $u_{ex}(x) = \sin x$. The obtained results for $J = 2$ are shown in Table 4 and Figures 5 and 6.

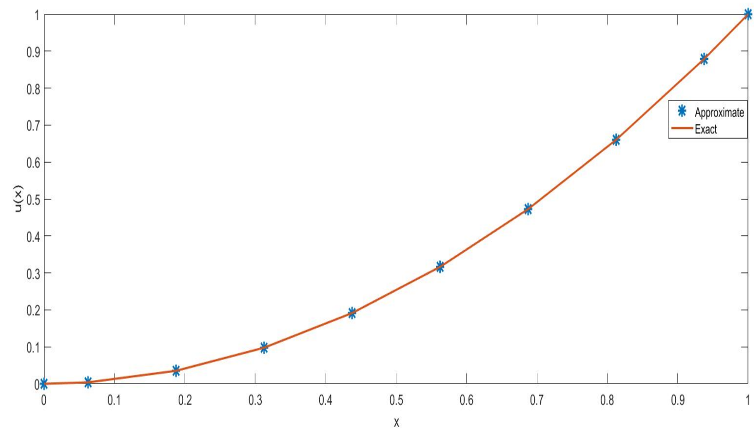


Figure 3: The comparison of the numerical solution and exact solution for example 3.

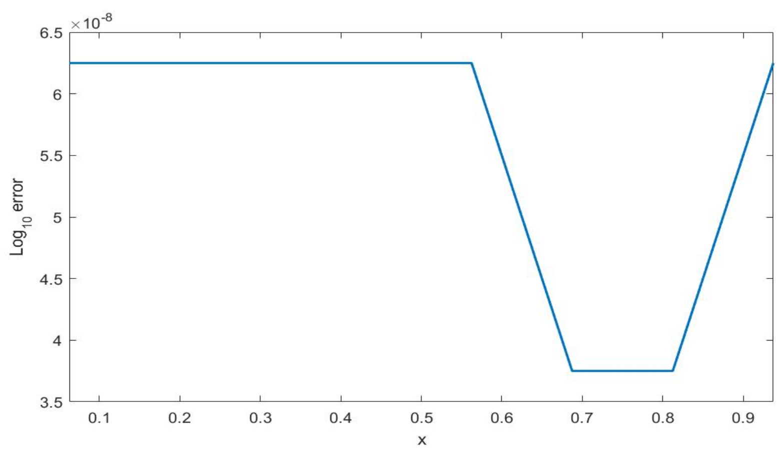


Figure 4: The absolute errors for example 3.

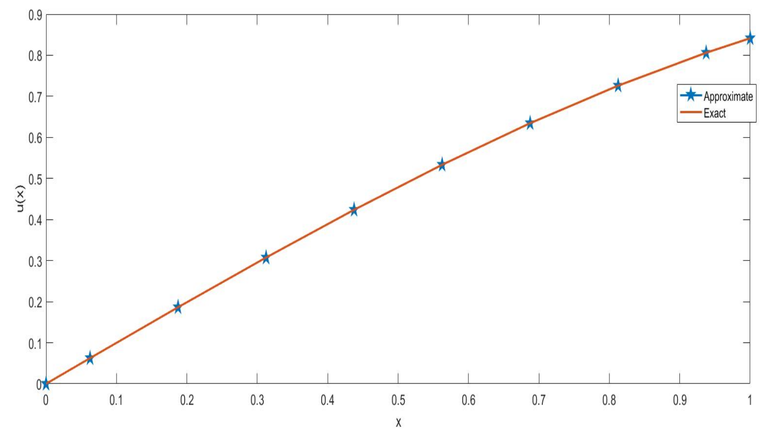


Figure 5: The comparison of the numerical solution and exact solution for Example 4.

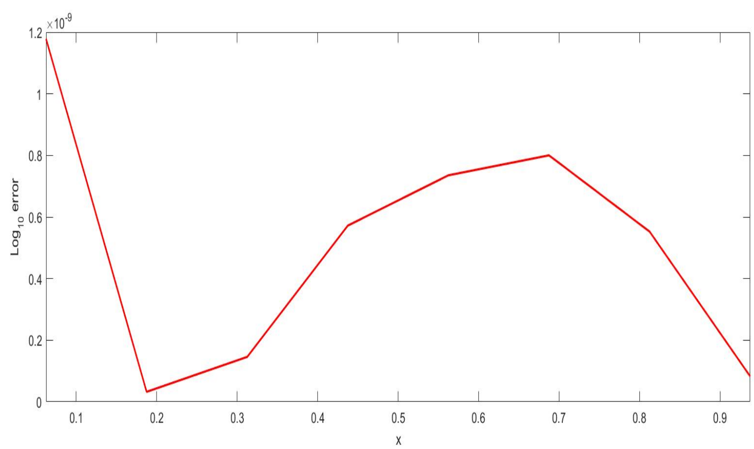


Figure 6: The absolute errors for example 4.

Table 3: Exact solution and The new method solution for equation (21)

$(x/16)$	Exact solution	The new method solution
1	0.00390	0.00390
3	0.03515	0.03515
5	0.09765	0.09765
7	0.19140	0.19140
9	0.31640	0.31640
11	0.47265	0.47266
13	0.66015	0.66016
15	0.87890	0.87890

5 Error analysis

In this section, the error analysis for the new method has been offered.

Lemma *Let $f \in L^2(R)$ be a continuous function defined in $(0,1)$. Then the error norm at Jth level satisfies the following inequalities*

$$E_J \leq \frac{3K^2}{16} 2^{-3J},$$

where $f(x) \leq K$, for all $x \in (0,1)$ and $K > 0$, and M is a positive number related to the Jth level resolution of the wavelet given by $M = 2^J$

Proof.

$$|E_J| = |u(x) - u_J(x)| = |\sum_{p=2M+1}^{\infty} a_p \int_0^1 e^{ix\xi} h_p(\xi) d\xi|,$$

where

$$u_J(x) = \sum_{p=1}^{2M} a_p \int_0^1 e^{ix\xi} h_p(\xi) d\xi$$

Table 4: Exact solution, The new method solution for equation (22)

$(x/16)$	Exact solution	The new method solution
1	0.0624593	0.0624592
3	0.1864032	0.1864033
5	0.3074385	0.3074385
7	0.4236762	0.4236762
9	0.5333026	0.5333026
11	0.6346070	0.6346070
13	0.7260086	0.7260086
15	0.8060811	0.8060811

$$\|E_J\|^2 = \int_0^1 (\sum_{p=2M+1}^{\infty} a_p \int_0^1 e^{ix\xi} h_p(\xi)d\xi)\overline{(\sum_{p=2M+1}^{\infty} a_p \int_0^1 e^{ix\xi} h_p(\xi)d\xi)}dx$$

$$= \sum_{p=2M+1}^{\infty} \sum_{k=2M+1}^{\infty} a_p\overline{a_k} \int_0^1 (\int_0^1 e^{ix\xi} h_p(\xi)d\xi)\overline{(\int_0^1 e^{ix\xi} h_k(\xi)d\xi)}dx$$

$$= \sum_{p=2M+1}^{\infty} \sum_{k=2M+1}^{\infty} a_p\overline{a_k} \int_0^1 (\sum_{s=1}^{2M} h_p(x_s) \int_{\tau_s}^{\tau_{s+1}} e^{ix\xi} d\xi)\overline{(\sum_{s=1}^{2M} h_k(x_s) \int_{\tau_s}^{\tau_{s+1}} e^{ix\xi} d\xi)}dx.$$

Applying mean value theorem

$$\sum_{p=2M+1}^{\infty} \sum_{k=2M+1}^{\infty} a_p\overline{a_k} \int_0^1 A_p(x)B_k(x)dx,$$

where

$$A_p(x) = \sum_{s=1}^{2M} \frac{1}{2M} h_p(x_s)(cos(x\eta_s) + isin(x\alpha_s))$$

and

$$B_k(x) = \sum_{k=1}^{2M} \frac{1}{2M} h_k(x_k)(cos(x\eta_k) - isin(x\alpha_k))$$

where $\eta_k, \alpha_k \in [\tau_k, \tau_{k+1}]$. Therefore

$$\begin{aligned}\|E_J\|^2 &\leq \sum_{p=2M+1}^{\infty}\sum_{k=2M+1}^{\infty}\frac{a_p\overline{a_k}}{2M^2}\int_0^1\sum_{s=1}^{2M}\sum_{s=1}^{2M}h_p(x_s)h_k(x_s)dx \\ &\leq \sum_{p=2M+1}^{\infty}\frac{|a_p|^2}{M}.\end{aligned}$$

Now

$$a_p = \int_0^1 2^{j/2}h_p(\xi)\hat{f}(\xi)d\xi = \int_0^1 2^{j/2}h_p(\xi)Re(\hat{f}(\xi))d\xi + i\int_0^1 2^{j/2}h_p(\xi)Im(\hat{f}(\xi))d\xi.$$

By applying the mean value theorem for $\beta_1 \in (\xi_1, \xi_2)$ and $\beta_2 \in (\xi_2, \xi_3)$

$$\begin{aligned}\int_0^1 2^{j/2}h_p(\xi)Re(\hat{f}(\xi))d\xi &= 2^{j/2}[\int_{\xi_1}^{\xi_2} Re(\hat{f}(\xi))d\xi - \int_{\xi_2}^{\xi_3} Re(\hat{f}(\xi))d\xi] \\ &= 2^{j/2}[(\xi_2-\xi_1)Re(\hat{f}(\beta_1)) - (\xi_3-\xi_2)Re(\hat{f}(\beta_2))].\end{aligned}$$

Applying the mean value theorem

$$\int_0^1 2^{j/2}h_p(\xi)Re(\hat{f}(\xi))d\xi = 2^{-j/2-1}(\beta_1-\beta_2)(Re\hat{f}(\lambda_1))',$$

where $\lambda_1 \in (\beta_1, \beta_2)$ and also

$$\int_0^1 2^{j/2}h_p(\xi)Im(\hat{f}(\xi))d\xi = 2^{-j/2-1}(\beta_3-\beta_4)(Im\hat{f}(\lambda_2))',$$

where $\lambda_2 \in (\beta_3, \beta_4)$, $\beta_3 \in (\xi_1, \xi_2)$, and $\beta_4 \in (\xi_2, \xi_3)$.
Therefore

$$a_p = \int_0^1 2^{j/2}h_p(\xi)\hat{f}(\xi)d\xi = 2^{-j/2-1}((\beta_1-\beta_2)Re\hat{f}(\lambda_1) + i(\beta_3-\beta_4)Im\hat{f}(\lambda_2))'.$$

This implies that

$$|a_p|^2 \leq K^2 2^{-3j}.$$

Therefore,

$$\|E_J\|^2 \leq \sum_{i=2M+1}^{\infty}\frac{|a_i|^2}{M} \leq \sum_{i=2M+1}^{\infty}\frac{K^2 2^{-3j}}{M}$$

$$
\begin{aligned}
&\leq \frac{K^2}{M} \sum_{j=J+1}^{\infty} \sum_{i=2^j+1}^{2^{j+1}} 2^{-3j} \leq \frac{K^2}{M} \sum_{j=J+1}^{\infty} 2^{-3j}(2^{j+1} - 1 - 2^j + 1) \\
&\leq \frac{K^2}{M} \sum_{j=J+1}^{\infty} 2^{-2j} \leq \frac{K^2}{M} (\frac{2^{-2(J+1)}}{1 - \frac{1}{4}}) \\
&\leq \frac{3K^2}{16} 2^{-3J}.
\end{aligned}
$$

□

It is obvious that the error bound is inversely proportional to the level of resolution J of Haar wavelet. Hence, the accuracy in our new method improves as we increase the level of resolution J.

Conclusions In this paper a new method for numerical solution of ordinary differential equations based on pseudo differential operators and Haar wavelets is proposed. Its applicability and efficiency are checked on some problems. It is shown that with a small number of collocation points, more accurate solutions in comparison with Haar wavelet method can be obtained.

Acknowledgements

Authors are grateful to anonymous referees and editor for their constructive comments.

References

1. Abel, H. *Pseudodifferential and singular integral operators: an introduction with applications*, De Gruyter, 2011.

2. Arora, S., Singh Brar, Y., and Kumar S. *Haar Wavelet Matrices for the Numerical Solutions of Differential Equations*, Int. J. Comput. Appl. (0975–8887) 97 (2014), no. 18, 33–36.

3. Arora, S., Singh, I., and Singh Brar, Y. *Comparative study of Haar wavelet with numerical methods for partial differential equations*, International Journal of Pure and Applied Mathematics, 101 (2015), no. 4, 489–503.

4. Boggiiatto, P. and Rodino, L. *Quantization and pseudo-differential operators*, Cubo. Math. Educ. 5 (2003), no. 1, 237–272.

5. Cerejeiras, P. and Vieira, N. *Regularization of the non-stationary Schrodinger operator*, Math. Methods Appl. Sci. 32 (2009), no. 5, 535–555.

6. Chen, C. F. and Hsiao, C. H. *Wavelet approach to optimising dynamic systems*, IEE Proc. Control Theory Appl. 146 (1999), no. 2, 213–219.

7. Chen, C. F. and Hsiao, C. H. *Haar wavelet method for solving lumped and distributed-parameter systems*, IEE Proc. Control Theory Appl. 144 (1997), no. 1, 87–94.

8. Haq, F. I. Hussain, I., and Ali, A. *A haar wavelets based numerical method for third-order-boundary and initial value problems*, World Appl. Sci. J. 13 (2011), no. 10, 2244–2251.

9. Lazaar, S., Ponenti, P., Liandrat, J., and Tchamitchian, P. *Wavelet algorithms for numerical resolution of partial differential equations*, Comput. Methods Appl. Mech. Engrg. 116 (1994), no. 1-4, 309–314.

10. Lepik, Ü. *Solving PDEs with the aid of two-dimensional Haar wavelets*, Comput. Math. Appl. 61 (2011), no. 7, 1873–1879.

11. Lepik, Ü. *Solving integral and differential equations by the aid of nonuniform Haar wavelets*, Aapl. Math. Comput. 198 (2008), no. 1, 326–332.

12. Lepik, Ü. *Haar wavelet method for solving higher order differential equations* , Int. J. Math. Comput. 1 (2008), no. 8, 84–94.

13. Lepik, Ü. *Numerical solutions of evolution equations by the Haar wavelet method*, Appl. Math. Comput. 185 (2007), no. 1, 695–704.

14. Lepik, Ü. *Numerical solutions of differential equations using Haar wavelets*, Math. Comput. Simulatio 68 (2005), no. 2, 127–143.

15. Luo, G. Y., Osypow, D., and Irle, M. *Vibration modeling with fast Gaussian wavelet algorithm*, Adv. Eng. Software 33 (2002), no. 4, 191–197.

16. Pandit, S. and Kumar, M. *Haar Wavelet approach for numerical solution of two parameters singularly perturbed boundary value problems*, Appl. Math. Inf. Sci. 8(2014), no. 6, 2965–2974.

17. Singh, I., Arora. S., and Kumar, S. *Numerical solution of wave equation using Haar wavelet*, Int. J. Pure Appl. Math. 98 (2015), no. 4, 457–469.

18. Tangborn, A. and Zhang, S. Q. *Wavelet transform adapted to an approximate Kalman filter system*, Appl. Numer. Math. 33 (2000), no. 1-4, 307–316.

19. Vasilyev, O. V. Paolucci, S., and Sen, M. *A multilevel wavelet collocation method for solving partial differential equations in a finite domain*, J. Comput. Phys. 120 (1995), no. 1, 33–47.

20. Walden, J. *Filter bank methods for hyperbolic PDEs*, SIAM J. Numer. Anal. 36 (1999), no. 4, 1183–1233.

Iranian Journal of Numerical Analysis and Optimization
Vol. 8, No. 1, (2018), pp 81–109
DOI:10.22067/ijnao.v8i1.55385

Constrained Bimatrix Games with Fuzzy Goals and its Application in Nuclear Negotiations

H. Bigdeli*, H. Hassanpour, and J. Tayyebi

Abstract

Solving constrained bimatrix games in the fuzzy environment is the aim of this research. This class of two-person nonzero-sum games is considered with finite strategies and fuzzy goals when some additional linear constraints are imposed on the strategies. We consider constrained two-person nonzero-sum games with single and multiple payoffs. It is shown that an equilibrium solution of single-objective case can be characterized by solving a quadratic programming problem with linear constraints. Some mathematical programming problems are also introduced to obtain the equilibrium points in multi-objective case with crisp and fuzzy constraints. Finally, a political application of such games is presented which is about nuclear negotiations between two countries.

Keywords: Multiobjective game; Constrained game; Fuzzy constrained game; Nuclear negotiations.

1 Introduction

Many real-world problems can be modeled as game problems. When the game theory is used to analysis some conflict problems in real situations, impreci-

*Corresponding author

Received 23 April 2016; revised 18 February 2017; accepted 22 November 2017

H. Bigdeli
Researcher, Institute for the Study of War, Command and Staff University, Tehran, I.R. Iran army. e-mail: hamidbigdeli92@gmail.com

H. Hassanpour
Department of Mathematics, Faculty of Mathematical Science and Statistics, University of Birjand, Birjand, I.R. Iran. e-mail: hhassanpour@birjand.ac.ir

J. Tayyebi
Department of Industrial Engineering, Faculty of Industrial and Computer Engineering, Birjand University of Technology, Birjand, I.R. Iran. e-mail: javadtayyebi@birjandut.ac.ir

sion or fuzziness is inherent in human judgments. Two types of inaccuracies of human judgments should be incorporated in the games; the players' ambiguous understanding of the payoffs and the fuzzy goals of the players for each of the objectives. Fuzzy set theory is well known for its ability to model decision making problems involving vagueness due to the lack of information and/or imprecision of the available information on the problem set-up [21, 24]. This ability has been successfully exploited for modeling different problems in various disciplines. In the field of fuzzy games, a considerable number of studies have been made (see, for example, [1,11,13,17]). Two-person nonzero-sum finite games are a kind of noncooperative games [19]. Two-person nonzero-sum games are also referred to as bimatrix games because they can be expressed by a pair of payoff matrices. These kind of games contain two-person zero-sum games as a special case. In real-world decision making problems facing humans today, people want to attain simultaneous goals; that is, they have multiple objectives. Hence, it seems natural that the game theoretic approaches to conflict resolution require to handle multiple objectives simultaneously. Wierzbicki [23] defined equilibrium solutions of multiobjective game based on order relations, using several preference cones and optimality criteria such as Pareto optimality for multiobjective noncooperative n-person games with nonlinear payoff functions. Corley [7] defined equilibrium solutions for multiobjective two-person nonzero-sum games and developed a method to compute them. Authors of [5] defined a proxy single-objective game with payoffs corresponding to a scalarizing function with weighting coefficients in multiobjective two-person nonzero-sum games and discussed the existence of equilibrium solutions for the original multiobjective two-person nonzero-sum games through the existence of the equilibrium solutions for the single objective proxy game. Fahem and Radjef [9] investigated the concept of properly efficient equilibrium for a multicriteria noncooperative strategic game. Nishizaki and Sakawa [17] studied two-person nonzero-sum game incorporating fuzzy goals in single and multiobjective environments. They defined an equilibrium solution with respect to the degree of attainment of the fuzzy goals in two-person nonzero-sum games and showed that an equilibrium obtained from solving several mathematical programming problems.

In some real-life game problems, choice of strategies for players is constrained due to some practical reason why this should be; that is, not all mixed strategies in a game are permitted for each player. These decision problems give rise to constrained games. Constrained matrix games initially were studied by Charnes [6] and then in a more general case by Kawaguchi and Maruyama [10]. Dresher [8] gave a real example of the constrained matrix game. Li and Cheng [11] presented a method based on the multiobjective programming to solve constrained matrix games with fuzzy numbers. Li and Hong [13] proposed a method for solving constrained matrix games with payoffs of triangular fuzzy numbers. They introduced the concepts of Alpha-constrained matrix games for the constrained matrix games with payoffs of triangular fuzzy numbers. Also, they [12] proposed Alpha-cut based linear programming methodology for the constrained matrix games with payoffs of trapezoidal fuzzy numbers.

The authors used multiobjective optimization in multiobjective zero-sum fuzzy matrix games in other researches [3, 4]. In this paper, we consider the case of constrained bimatrix games with single and multiple payoffs in which fuzzy goals are specified for the objectives by the players.

The remainder of the paper is organized as follows. In section 2, some preliminaries and necessary definitions about fuzzy sets and bimatrix games are presented. In section 3, constrained bimatrix games with fuzzy goals are introduced. It is shown that the equilibrium solution of this class of nonzero-sum games can be characterized by solving a quadratic programming problem. Then, multiobjective constrained bimatrix games with fuzzy goals is considered, and a mathematical programming problem to solve such games is presented. Also, for the case that additional linear constraints are imposed on the strategies are fuzzy constraints, a mathematical programming problem is introduced to obtain the equilibrium point of such games. In section 4, a political application of multiobjective constrained bimatrix games with fuzzy goals is presented in which nuclear negotiations between two countries have been discussed. Finally, conclusion is made in section 5.

2 Preliminaries

In this section, we provide some definitions and preliminaries of fuzzy sets and bimatrix games according to [1, 20].

Let X denote a universal set. A fuzzy subset $\tilde{A}$ of X is defined by its membership function $\mu_{\tilde{A}} : X \to [0, 1]$ which assigns to each element $x \in X$ a real number $\mu_{\tilde{A}}(x)$ in the interval $[0, 1]$. The value of $\mu_{\tilde{A}}(x)$ represents the grade of membership of x in $\tilde{A}$. The fuzzy subset $\tilde{A}$ is denoted by a set of ordered pairs of elements x and their grades $\mu_{\tilde{A}}(x)$; that is, $\tilde{A} = \{(x, \mu_{\tilde{A}}(x))|x \in X\}$.

A fuzzy decision making problem is characterized by a set X of possible alternatives and a set of fuzzy goals G_i, $i = 1, \ldots, p$, as well as a set of fuzzy constraints C_j, $j = 1, \ldots, n$, each of which is expressed by a fuzzy set on X. For such a decision making problem, Bellman and Zadeh [2] proposed that a fuzzy decision is determined by an appropriate aggregation of the fuzzy sets G_i, $i = 1, \ldots, p$, and C_j, $j = 1, \ldots, n$. Realizing that both the fuzzy goal and the fuzzy constraint are desired to be satisfied simultaneously, they suggested the aggregation operator to be the fuzzy intersection. Thus a fuzzy decision D could be defined as the fuzzy set $D = (G_1 \cap \cdots \cap G_p) \cap (C_1 \cap \cdots \cap C_n)$; that is, $\mu_D(x) : X \to [0, 1]$ given by $\mu_D(x) = \min(\mu_{G_1}(x), \ldots, \mu_{G_p}(x), \mu_{C_1}(x), \mu_{C_n}(x))$. Once the fuzzy decision D is known, we can define $x^* \in X$ to be an optimal decision if $\mu_D(x^*) = \max_x \mu_D(x)$.

Let two players in a two-person bimatrix game be denoted by Players I and II. In a two-person bimatrix game, each player has finite number of strategies. The payoff function is determined by two matrices $A, B \in \mathbb{R}^{m\times n}$. When Player I chooses his ith strategy and Player II his jth strategy, then a_{ij} and b_{ij} are the payoffs of Players I and II, respectively. Thus, a bimatrix game is determined by a pair of matrices (A,B). Rational behavior of players is assumed; that is, each of them attempts to maximize his reward. Mixed strategy is a probability vector over the set of pure strategies. So, bimatrix game can be represented as $BG : (X, Y, A, B)$, where

$$X = \{x \in \mathbb{R}^m | \sum_{i=1}^{m} x_i = 1, x_i \geq 0 \qquad i = 1, \dots, m\}$$

and

$$Y = \{y \in \mathbb{R}^n | \sum_{j=1}^{n} y_j = 1, y_j \geq 0,\ j = 1, \dots, n\}$$

are called the mixed strategy spaces for Players I and II, and A and B are called the payoff matrices for Players I and II, respectively.

Definition 1. A pair $(x^*, y^*) \in X \times Y$ is said to be an equilibrium solution of the bimatrix game BG if

$$x^T Ay^* \leq x^{*T} Ay^* \qquad \forall x \in X$$

and

$$x^{*T} By \leq x^{*T} By^* \qquad \forall y \in Y.$$

In other words, no player has a motivation to change his strategy.

The following theorem due to Nash guarantees the existence of an equilibrium solution of the bimatrix game.

Theorem 1. [20]*Every bimatrix game $BG : (X, Y, A, B)$ has at least one equilibrium solution.*

A Nash equilibrium solution of the bimatrix game BG can be obtained by solving an appropriate quadratic programming problem as discussed below.

Theorem 2. [15]*Let $BG : (X, Y, A, B)$ be a given bimatrix game. The pair (x^*, y^*) is an equilibrium solution of BG if and only if it is an optimal solution to the following quadratic programming problem:*

$$\begin{array}{ll} \max\limits_{x,y,\alpha,\beta} & x^T(A+B)y-\alpha-\beta \\ s.t. & Ay-\alpha e_m \leq 0, \\ & B^T x-\beta e_n \leq 0, \\ & x\in X,\, y\in Y, \end{array} \tag{1}$$

where α and β are scalars; e_m and e_n are, respectively, m- and n-dimensional column vectors whose elements are all ones. The optimal values of α and β are the expected payoffs to Players I and II, respectively. Furthermore, the optimal objective value of the problem (1) equals zero.

We will use Pareto optimality concept of multiobjective optimization in solution concept of multiobjective two-person nonzero-sum games. Thus, we review some of solutions concepts in multiobjective mathematical programming. For convenience, let us introduce the following notation: for any vectors $z, z^{'} \in \mathbb{R}^N$

$$\begin{array}{l} z = z^{'} \Longleftrightarrow z_i = z_i^{'} \quad i=1,\ldots,N; \\ z \leqq z^{'} \Longleftrightarrow z_i \leqq z_i^{'} \quad i=1,\ldots,N; \\ z < z^{'} \Longleftrightarrow z_i < z_i^{'} \quad i=1,\ldots,N; \\ z \leq z^{'} \Longleftrightarrow z \leqq z^{'} \quad \text{and} \quad z \neq z^{'}. \end{array}$$

A multiobjective mathematical programming problem can be written as

$$\begin{array}{l} \min f(z) = (f_1(z),\ldots,f_l(z)) \\ \text{s.t.} \\ \qquad z \in Z = \{z \in \mathbb{R}^N | g(z) \leqq 0,\, h(z) = 0\}, \end{array}$$

where $g(z) = (g_1(z),\ldots,g_{m_1}(z))$, $h(z) = (h_1(z),\ldots,h_{m_2}(z))$, and 0 is zero vector with the same dimension as the left hand side vectors.
There does not generally exist a solution minimizing all of the objectives simultaneously. Therefore, Pareto optimal (efficient) solutions are introduced as follows.

Definition 2. [22] $z^* \in Z$ is said to be a Pareto optimal solution if there does not exist another $z \in Z$ such that $f(z) \leq f(z^*)$.

As a slightly weaker solution concept than Pareto optimality, weak Pareto optimal solutions are defined by replacing $\leq$ with $<$ in Definition 2.

3 Constrained bimatrix games with fuzzy goals

Nishizaki and Sakawa [18] studied two-person nonzero-sum games incorporating fuzzy goals in multiobjective environment. They defined an equilibrium solution with respect to the degree of attainment of the fuzzy goals and proved that an equilibrium solution can be obtained by solving several certain mathematical programming problems. Let us consider a case of BG in which all mixed strategies x and y must be chosen from some sets $\hat{X}$ and $\hat{Y}$ determined by some linear constraints. A constrained bimatrix game is denoted by $CBG : (\hat{X}, \hat{Y}, A, B)$ in which

$$\hat{X} = \left\{ x \in \mathbb{R}^m \left| e_m^T x = 1, Dx \leq c \; , \; x \geq 0 \right. \right\}, \tag{2}$$

and

$$\hat{Y} = \left\{ y \in \mathbb{R}^n \left| e_n^T y = 1, Ey \leq d \; , \; y \geq 0 \right. \right\}, \tag{3}$$

are called the mixed strategy spaces for Players I and II in the constrained game, respectively. In the above sets $c \in \mathbb{R}^p$, $d \in \mathbb{R}^q$, $D \in \mathbb{R}^{p \times m}$, and $E \in \mathbb{R}^{q \times n}$. e_m and e_n are m- and n-dimensional vectors of ones.
When Player I chooses a mixed strategy $x \in \hat{X}$ and Player II chooses a mixed strategy $y \in \hat{Y}$, the values $x^T Ay$ and $x^T By$ are the expected payoffs for Players I and II, respectively.

Definition 3. A pair $(x^*, y^*) \in \hat{X} \times \hat{Y}$ is said to be an equilibrium solution of the constrained bimatrix game CBG if

$$\begin{array}{ll} x^T Ay^* \leq x^{*T} Ay^* & \quad \forall x \in \hat{X}, \\ x^{*T} By \leq x^{*T} By^* & \quad \forall y \in \hat{Y}. \end{array}$$

In this paper, we incorporate fuzzy goals for objectives and introduce equilibrium solutions in terms of maximizing the degree of attainment for fuzzy goals similar to the work of [18]. In the following, we define fuzzy goals for Players I and II.

Definition 4. Let the expected payoffs of Player I be $D_1 = \{x^T Ay | \; x \in \hat{X}, y \in \hat{Y}\}$. A fuzzy goal for Player I is a fuzzy set $\tilde{G}_1$ represented by an increasing membership function $\mu_1 : D_1 \to [0, 1]$.

Definition 5. Let the expected payoffs of Player II be $D_2 = \{x^T By | \; x \in \hat{X}, y \in \hat{Y}\}$. A fuzzy goal for Player II is a fuzzy set $\tilde{G}_2$ with an increasing membership function $\mu_2 : D_2 \to [0, 1]$.

Note that μ_1 (μ_2) assigns to each expected payoff $x^T Ay \in D_1$ ($x^T By \in D_2$) a real number $\mu_1(x^T Ay)$ ($\mu_2(x^T By)$) which denotes the degree of satisfaction of Player I (Player II) from the payoff $x^T Ay$ ($x^T By$).

An equilibrium solution of a constrained bimatrix game with fuzzy goals is defined with respect to the degree of attainment of the fuzzy goals.

Definition 6. (Equilibrium solution). A pair $(x^*, y^*) \in \hat{X} \times \hat{Y}$ is said to be an equilibrium solution of the constrained bimatrix game $CBG : (\hat{X}, \hat{Y}, A, B)$ with fuzzy goals if

$$\mu_1(x^{*T}Ay^*) \geq \mu_1(x^TAy^*) \qquad \forall x \in \hat{X},$$
$$\mu_2(x^{*T}By^*) \geq \mu_2(x^{*T}By) \qquad \forall y \in \hat{Y}.$$

The membership function μ_1 (μ_2) in the above definition can be interpreted as Player I's (Player II's) satisfaction function of payoff. Thus the game with fuzzy goals can be reduced to an ordinary two-person constrained nonzero-sum game whose payoff function is its satisfaction function. Let us consider the fuzzy goal of Player I (Player II) as "the expected payoff should be substantially greater than or equal to $\bar{a}$ ($\bar{b}$)". Here, we use linear membership functions for these fuzzy goals as follows (see Figure 1 for μ_1):

$$\mu_1(x^TAy) = \begin{cases} 0, & x^TAy \leq \underline{a}, \\ 1 - \frac{\bar{a}-x^TAy}{\bar{a}-\underline{a}}, & \underline{a} \leq x^TAy \leq \bar{a}, \\ 1, & x^TAy \geq \bar{a}, \end{cases} \tag{4}$$

and

$$\mu_2(x^TBy) = \begin{cases} 0, & x^TBy \leq \underline{b}, \\ 1 - \frac{\bar{b}-x^TBy}{\bar{b}-\underline{b}}, & \underline{b} \leq x^TBy \leq \bar{b}, \\ 1, & x^TBy \geq \bar{b}. \end{cases} \tag{5}$$

In (4) and (5), $\bar{a}$ and $\bar{b}$ are the desirable values of expected payoffs for Players I and II, respectively. Also, $\underline{a}$ and $\underline{b}$ are the least allowable values for their expected payoffs. Thus $\bar{a} - \underline{a}$ ($\bar{b} - \underline{b}$) is the most allowable value for violation from $\bar{a}$ ($\bar{b}$) for Player I (Player II). Notice that $\underline{a}$, $\bar{a}$, $\underline{b}$, and $\bar{b}$ could be any values with $\bar{a} > \underline{a}$ and $\bar{b} > \underline{b}$ which is given by players. In the case that they can not give these values; we can set $\underline{a} = \min_i \min_j a_{ij}$, $\bar{a} = \max_i \max_j a_{ij}$, $\underline{b} = \min_i \min_j b_{ij}$, and $\bar{b} = \max_i \max_j b_{ij}$, in which $\underline{a}$ ($\underline{b}$) gives the worst degree of satisfaction and $\bar{a}$ ($\bar{b}$) gives the best degree of satisfaction for Player I (Player II).
Letting

$$\hat{A} = \frac{A}{\bar{a}-\underline{a}}, \hat{B} = \frac{B}{\bar{b}-\underline{b}}, c_1 = -\frac{\underline{a}}{\bar{a}-\underline{a}}, \text{ and } c_2 = -\frac{\underline{b}}{\bar{b}-\underline{b}};$$

the membership functions $\mu_1(x^TAy)$ and $\mu_2(x^TBy)$ can be rewritten as

$$\mu_1(x^TAy) = \begin{cases} 0, & x^TAy \leq \underline{a}, \\ c_1 + x^T\hat{A}y, & \underline{a} \leq x^TAy \leq \bar{a}, \\ 1, & x^TAy \geq \bar{a}, \end{cases}$$

and

$$\mu_2(x^TBy) = \begin{cases} 0, & x^TBy \leq \underline{b}, \\ c_2 + x^T\hat{B}y, & \underline{b} \leq x^TBy \leq \overline{b}, \\ 1, & x^TBy \geq \overline{b}. \end{cases}$$

respectively.

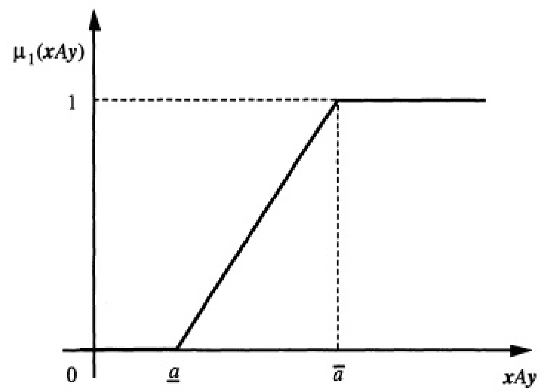


Figure 1: A linear membership function for fuzzy goal of Player I.

The following theorem states that an equilibrium solution of $CBG : (\hat{X}, \hat{Y}, \hat{A}, \hat{B})$ is equal to the equilibrium solution of $CBG : (\hat{X}, \hat{Y}, A, B)$. Moreover, it is also an equilibrium solution with respect to the degree of attainment of the fuzzy goals for $CBG : (\hat{X}, \hat{Y}, A, B)$ with fuzzy goals.

Theorem 3. *A pair of strategies (x^*, y^*) satisfies the conditions*

$$\begin{array}{ll} x^{*T}\hat{A}y^* \geq x^T\hat{A}y^* & \forall x \in \hat{X} \\ x^{*T}\hat{B}y^* \geq x^{*T}\hat{B}y & \forall y \in \hat{Y} \end{array}$$

if and only if (x^*, y^*) satisfies the following conditions

$$\begin{array}{ll} x^{*T}Ay^* \geq x^TAy^* & \forall x \in \hat{X} \\ x^{*T}By^* \geq x^{*T}By & \forall y \in \hat{Y} \end{array}$$

furthermore, when the membership functions of the fuzzy goals are linear functions, (x^*, y^*) satisfies the following conditions

$$\begin{array}{ll} \mu_1(x^{*T}Ay^*) \geq \mu_1(x^TAy^*) & \forall x \in \hat{X} \\ \mu_2(x^{*T}By^*) \geq \mu_1(x^{*T}By) & \forall y \in \hat{Y} \end{array}$$

Proof. The presented proof in [1] (Theorems 9.2.1 and 9.2.2) in the case of unconstrained bimatrix games is also valid for our problem. $\square$

The following example shows that the converse of the second part of Theorem 3 is not established.

Example 1. Set

$$X = \{x|x_1 + x_2 = 1, x_1, x_2 \geq 0\},\ Y = \{y|y_1 + y_2 = 1, y_1, y_2 \geq 0\},$$

$$A = B = \begin{bmatrix} 1 & 0.75 \\ 0.5 & 0.25 \end{bmatrix},$$

$$\mu_1(x^TAy) = \begin{cases} 0, & x^TAy \leq 0.25, \\ 1 - \frac{0.5 - x^TAy}{0.25}, & 0.25 \leq x^TAy \leq 0.5, \\ 1, & x^TAy \geq 0.5, \end{cases}$$

and

$$\mu_2(x^TBy) = \begin{cases} 0, & x^TBy \leq 0.25, \\ 1 - \frac{0.5 - x^TBy}{0.25}, & 0.25 \leq x^TBy \leq 0.5, \\ 1, & x^TBy \geq 0.5. \end{cases}$$

For $x^* = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$ and $y^* = \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix}$, we have

$$\mu_1\left(x^{*T}Ay^*\right) \geq \mu_1\left(x^TAy^*\right) \qquad \forall x \in X$$

and

$$\mu_2\left(x^{*T}By^*\right) \geq \mu_2\left(x^{*T}By\right) \qquad \forall y \in Y.$$

Because

$$x^{*T}Ay^* = (0.2, 0.8)\begin{bmatrix} 1 & 0.75 \\ 0.5 & 0.25 \end{bmatrix}\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = [0.6 \quad 0.35]\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = 0.525$$

and

$$\mu_1\left(x^{*T}Ay^*\right) = \mu_1\left(0.525\right) = 1$$

which imply that

$$1 = \mu_1\left(x^{*T}Ay^*\right) \geq \mu_1\left(x^TAy^*\right) \qquad \forall x \in X.$$

Similarly, $1 = \mu_1\left(x^{*T}By^*\right) \geq \mu_1\left(x^{*T}By\right)\ \forall y \in Y$.

On the other hand for $x = \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$, we have

$$x^TAy^* = (0.4, 0.6)\begin{bmatrix} 1 & 0.75 \\ 0.5 & 0.25 \end{bmatrix}\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = [0.7 \ \ 0.45]\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = 0.625$$

$$\mu_1\left(x^TAy^*\right) = \mu_1\left(0.625\right) = 1$$

We see that, $x^{*T}Ay^* = 0.525 < x^TAy^* = 0.625$. Therefore (x^*, y^*) is not an equilibrium point for bimatrix game (X, Y, A, B). Hence the converse of Theorem 3 is not established.

Now, we have the following theorem which provides a quadratic programming problem to find an equilibrium solution of the $CBG : (\hat{X}, \hat{Y}, A, B)$ with fuzzy goals.

Theorem 4. *(Equivalence theorem) Let $(x^*, y^*, \alpha^*, \beta^*, u^*, v^*)$ be an optimal solution of the following quadratic programming problem:*

$$\begin{array}{ll} \max\limits_{x,y,u,v,\alpha,\beta} & x^T(\hat{A}+\hat{B})y - cu - dv - \alpha - \beta \\ s.t. & \hat{A}y \leq cue_m + \alpha e_m, \\ & \hat{B}^T x \leq dve_n + \beta e_n, \\ & D^T x - c \leq 0, \\ & E^T y - d \leq 0, \\ & e_m^T x = 1, \\ & e_n^T y = 1, \\ & x, y, u, v \geq 0. \end{array} \tag{6}$$

Then (x^*, y^*) is an equilibrium solution of the $CBG : (\hat{X}, \hat{Y}, A, B)$ with fuzzy goals.

Proof. First, we show that the optimal value of objective function is zero. The constraints of the problem evidently imply that

$$x^T\hat{A}y + x^T\hat{B}y - cu - dv - \alpha - \beta \leq 0.$$

Thus the objective optimal value of the problem (6) is nonpositive. Consider the point $(\bar{x}, \bar{y}, \bar{\alpha}, \bar{\beta}, \bar{u}, \bar{v})$, where $\bar{x} = x^* \in \hat{X}, \bar{y} = y^* \in \hat{Y}, \bar{\alpha} = x^{*T}\hat{A}y^* - cu^*, \bar{\beta} = x^{*T}\hat{B}y^* - dv^*, \bar{u} = u^*$, and $\bar{v} = v^*$. We first prove that it is feasible to the problem (6). Assume that it is not feasible. Obviously, the solution does not satisfy at least one of the two first constraints of the problem (6). We assume that the first constraint does not hold. The proof of the other case is similar. Thus,

$$\hat{A}y^* > cu^*e_m + \alpha^* e_m. \tag{7}$$

Multiplying both sides of (7) $x^* \in \hat{X}$, implies that

$$x^{*T}\hat{A}y^* > cu^* + \alpha^* = x^{*T}\hat{A}y^*$$

which is incorrect.

On the other hand, it is clear that the objective value of $(\bar{x}, \bar{y}, \bar{\alpha}, \bar{\beta}, \bar{u}, \bar{v})$ is zero. Thus it means that the optimal value of objective (6) is zero.
Now, let $(x^*, y^*, \alpha^*, \beta^*, u^*, v^*)$ be an optimal solution of the problem (6); thus

$$x^{*T}\hat{A}y^* + x^{*T}\hat{B}y^* - cu^* - dv^* - \alpha^* - \beta^* = 0. \tag{8}$$

From the first and second constraints, for any $x \in \hat{X}$ and $y \in \hat{Y}$, we have

$$\begin{aligned} x^T \hat{A} y^* &\leq cu^* + \alpha^* \\ \text{and } x^{*T} \hat{B} y &\leq dv^* + \beta^*, \end{aligned} \tag{9}$$

which imply

$$x^T \hat{A} y^* + x^{*T} \hat{B} y \leq cu^* + \alpha^* + dv^* + \beta^*.$$

Also, the equation (9) implies in particular that

$$\begin{aligned} x^{*T} \hat{A} y^* &\leq cu^* + \alpha^* \\ \text{and } x^{*T} \hat{B} y^* &\leq dv^* + \beta^*. \end{aligned}$$

On the other hand, from (8) we have

$$x^{*T} \hat{A} y^* + x^{*T} \hat{B} y^* = cu^* + \alpha^* + dv^* + \beta^*.$$

Thus

$$\begin{aligned} x^{*T} \hat{A} y^* &= cu^* + \alpha^* \\ \text{and } x^{*T} \hat{B} y^* &= dv^* + \beta^*. \end{aligned}$$

Therefore, (9) implies that

$$\begin{aligned} x^{*T} \hat{A} y^* &\geq x^T \hat{A} y^* \qquad \forall x \in \hat{X} \\ \text{and } x^{*T} \hat{B} y^* &\geq x^{*T} \hat{B} y \qquad \forall y \in \hat{Y} \end{aligned}$$

This proves that (x^*, y^*) is an equilibrium pair of $CBG : (\hat{X}, \hat{Y}, \hat{A}, \hat{B})$. But, by Theorem 3, this means that (x^*, y^*) is an equilibrium solution of the $CBG : (\hat{X}, \hat{Y}, A, B)$ with fuzzy goals. □

Remark 1. When an optimal solution $(x^*, y^*, \alpha^*, \beta^*, u^*, v^*)$ of (6) has been obtained, (x^*, y^*) gives an equilibrium solution of the constrained bimatrix game $CBG : (\hat{X}, \hat{Y}, A, B)$ with fuzzy goals. The degree of attainment of fuzzy goal $\tilde{G}_1$ and $\tilde{G}_2$ can then be determined by $\mu_1(x^{*T} A y^*)$ (or $\mu_1(x^{*T} \hat{A} y^*)$) and $\mu_2(x^{*T} B y^*)$ (or $\mu_2(x^{*T} \hat{B} y^*)$).

3.1 Multiobjective constrained bimatrix games with fuzzy goals

Multiobjective two-person nonzero-sum games can be expressed as multiple $m \times n$ matrices

$$A^1 = \begin{bmatrix} a_{11}^1 & \dots & a_{1n}^1 \\ \vdots & \ddots & \vdots \\ a_{m1}^1 & \dots & a_{mn}^1 \end{bmatrix}, \dots, A^r = \begin{bmatrix} a_{11}^r & \dots & a_{1n}^r \\ \vdots & \ddots & \vdots \\ a_{m1}^r & \dots & a_{mn}^r \end{bmatrix},$$

$$B^1 = \begin{bmatrix} b_{11}^1 & \dots & b_{1n}^1 \\ \vdots & \ddots & \vdots \\ b_{m1}^1 & \dots & b_{mn}^1 \end{bmatrix}, \dots, B^s = \begin{bmatrix} b_{11}^s & \dots & b_{1n}^s \\ \vdots & \ddots & \vdots \\ b_{m1}^s & \dots & b_{mn}^s \end{bmatrix},$$

where Player I has m pure strategies and r objectives and Player II has n pure strategies and s objectives. When Player I chooses the pure strategy i and Player II chooses the pure strategy j, they respectively receive the payoff vectors $(a_{ij}^1, \dots, a_{ij}^r)$ and $(b_{ij}^1, \dots, b_{ij}^s)$. The mixed strategy spaces for Players I and II are, respectively, given by (2) and (3).

When Player I chooses a constrained mixed strategy $x \in \hat{X}$ and Player II chooses a constrained mixed strategy $y \in \hat{Y}$, the k-th payoff of Player I is $f_1^k(x,y) = x^T A^k y$ and the l-th expected payoff of Player II is $f_2^l(x,y) = x^T B^l y$. Assume that for each objective $f_1^k(x,y) = x^T A^k y, k = 1, \dots, r$, Player I has a fuzzy goal such as "the k-th objective value of the game should be substantially more than or equal to some value $\bar{a}^k$". This statement can be quantified by eliciting a membership function. Assume that the corresponding linear membership function is defined as follows:

$$\mu_1^k(x^T A^k y) = \begin{cases} 0, & x^T A^k y \le \underline{a}^k, \\ 1 - \frac{\bar{a}^k - x^T A^k y}{\bar{a}^k - \underline{a}^k}, & \underline{a}^k \le x^T A^k y \le \bar{a}^k, \\ 1, & x^T A^k y \ge \bar{a}^k, \end{cases}$$

where $\underline{a}^k$ is the least allowable value of the k-th payoff.
Similarly, assume that for each objective $f_2^l(x,y) = x^T B^l y, l = 1, \dots, s$, Player II has a fuzzy goal such as "the l-th objective value of the game should be substantially more than or equal to some value $\bar{b}^l$." Then the corresponding linear membership function is defined as follows:

$$\mu_2^l(x^T B^l y) = \begin{cases} 0, & x^T B^l y \le \underline{b}^l, \\ 1 - \frac{\bar{b}^l - x^T B^l y}{\bar{b}^l - \underline{b}^l}, & \underline{b}^l \le x^T B^l y \le \bar{b}^l, \\ 1, & x^T B^l y \ge \bar{b}^l, \end{cases}$$

where $\underline{b}^l$ is the least allowable value of the l-th payoff.
Letting

$$\hat{A}^k = \frac{A^k}{\bar{a}^k - \underline{a}^k}, \hat{B}^l = \frac{B^l}{\bar{b}^l - \underline{b}^l}, c_1^k = -\frac{\underline{a}^k}{\bar{a}^k - \underline{a}^k}, \text{and } c_2^l = -\frac{\underline{b}^l}{\bar{b}^l - \underline{b}^l},$$

the membership functions $\mu_1^k(x^T A^k y)$ and $\mu_2^l(x^T B^l y)$ can be rewritten as

$$\mu_1^k(x^TA^ky) = \begin{cases} 0, & x^TA^ky \leq \underline{a}^k, \\ c_1^k + x^T\hat{A}^ky, & \underline{a}^k \leq x^TA^ky \leq \bar{a}^k, \\ 1, & x^TA^ky \geq \bar{a}^k, \end{cases}$$

and

$$\mu_2^l(x^TB^ly) = \begin{cases} 0, & x^TB^ly \leq \underline{b}^l, \\ c_2^l + x^T\hat{B}^ly, & \underline{b}^l \leq x^TB^ly \leq \bar{b}^l, \\ 1' & x^TB^ly \geq \bar{b}^l. \end{cases}$$

We consider the equilibrium solutions in terms of maximizing the degree of attainment of fuzzy goals. We use the method of Bellman and Zadeh [2] to aggregate the fuzzy goals.
The aggregated fuzzy goals of Players I and II are as follows, respectively:

$$\begin{array}{l} \mu_1(x,y) = \min_{k=1,\dots,r} \mu_1^k(x^TA^ky), \\ \mu_2(x,y) = \min_{l=1,\dots,s} \mu_2^l(x^TB^ly). \end{array} \tag{10}$$

Definition 7. A pair of strategies (x^*, y^*) is an equilibrium solution with respect to the degree of attainment of the fuzzy goals aggregated by the minimum component for the multiobjective $CBG : (\hat{X}, \hat{Y}, A, B)$ with fuzzy goals if for any other mixed strategies $x \in \hat{X}$ and $y \in \hat{Y}$

$$\mu_1(x^*, y^*) \geq \mu_1(x, y^*)$$

and

$$\mu_2(x^*, y^*) \geq \mu_2(x^*, y),$$

where μ_1 and μ_2 are given by (10).

Based on the above definition, an equilibrium solution is (x^*, y^*) in which, x^* and y^* are the optimal solutions to the following two mathematical programming problems:

$$\begin{array}{ll} \text{maximize} & \min\{\mu_1^1(x^TA^1y^*), \dots, \mu_1^r(x^TA^ry^*)\} \\ \text{s.t.} & e_m^Tx = 1, \\ & D^Tx \leq c, \\ & x \geq 0, \end{array}$$

and

$$\begin{array}{ll} \text{maximize} & \min\{\mu_2^1(x^{*T}B^1y), \dots, \mu_2^s(x^{*T}B^sy)\} \\ \text{s.t.} & e_n^Ty = 1, \\ & E^Ty \leq d, \\ & y \geq 0. \end{array}$$

Remark 2. Similar to the proposed theorem for unconstrained bimatrix games in [17] (Lemma 4.2.1), we can say that an equilibrium point is a pair of strategies (x^*, y^*) in which x^* and y^* are the optimal solutions of the following problems, respectively:

$$\begin{array}{ll} \text{maximize} & \min_{k=1,\dots,r}\{c_1^k + x^T \hat{A}^k y^*\} \\ \text{s.t.} & e_m^T x = 1, \\ & D^T x \le c, \\ & 0 \le c_1^k + x^T \hat{A}^k y^* \le 1, \quad k = 1,\dots,r, \\ & x \ge 0, \end{array} \tag{11}$$

and

$$\begin{array}{ll} \text{maximize} & \min_{l=1,\dots,s}\{c_2^l + x^{*T} \hat{B}^l y\} \\ \text{s.t.} & e_n^T y = 1, \\ & E^T y \le d, \\ & 0 \le c_2^l + x^{*T} \hat{B}^l y \le 1, \quad l = 1,\dots,s, \\ & y \ge 0. \end{array} \tag{12}$$

Since the constraints of the problems (11) and (12) are separable, these problems yield the following mathematical programming problem:

$$\begin{array}{ll} \text{maximize}(\min_{k=1,\dots,r}\{c_1^k + x^T \hat{A}^k y^*\} + \min_{l=1,\dots,s}\{c_2^l + x^{*T} \hat{B}^l y\}) & \\ \text{s.t.} \quad D^T x \le c, & \\ \quad\quad E^T y \le d, & \\ \quad\quad e_m^T x = 1, & \\ \quad\quad e_n^T y = 1, & \\ \quad\quad 0 \le c_1^k + x^T \hat{A}^k y^* \le 1, \quad k = 1,\dots,r, & \\ \quad\quad 0 \le c_2^l + x^{*T} \hat{B}^l y \le 1, \quad l = 1,\dots,s, & \\ \quad\quad x \ge 0, & \\ \quad\quad y \ge 0. & \end{array}$$

Now, we can present the following theorem.

Theorem 5. *Let $(x^*, y^*, \alpha^*, \beta^*, \lambda_1^*, \lambda_2^*, u^*, v^*)$ be an optimal solution to the following nonlinear programming problem*

(NLP1):

$$\begin{array}{lr} \max \lambda_1 + \lambda_2 - cu - dv - \alpha - \beta & \\ s.t. & \\ \quad \hat{A}^k y + c_1^k e_m \le cue_m + \alpha e_m \quad \textit{for some } k \in \{1,\dots,r\}, & (c1) \\ \quad (\hat{B}^l)^T x + c_2^l e_n \le dve_n + \beta e_n \quad \textit{for some } l \in \{1,\dots,s\}, & (c2) \\ \quad x^T \hat{A}^k y + c_1^k \ge \lambda_1, \quad k = 1,\dots,r, & (c3) \\ \quad x^T \hat{B}^l y + c_2^l \ge \lambda_2, \quad l = 1,\dots,s, & (c4) \\ \quad D^T x \le c, & (c5) \\ \quad E^T y \le d, & (c6) \end{array}$$

$$
\begin{array}{ll}
e_m^T x = 1, & (c7)\\
e_n^T y = 1, & (c8)\\
0 \leq c_1^k + x^T \hat{A}^k y \leq 1, & (c9)\\
0 \leq c_2^l + x^T \hat{B}^l y \leq 1, & (c10)\\
0 \leq \lambda_1, \lambda_2 \leq 1, & (c11)\\
x, y, u, v \geq 0. & (c12)
\end{array}
$$

Then (x^*, y^*) is an equilibrium solution with respect to the degree of attainment of the aggregated fuzzy goal.

Proof. First, we show that the optimal value of objective is zero. Let $(x, y, \alpha, \beta, \lambda_1, \lambda_2, u, v)$ be an arbitrary feasible solution of the problem (NLP1). From the constraints (c1) and (c7), it is easy to verify that

$$x^T \hat{A}^k y + c_1^k \leq cu + \alpha \qquad \text{for some}\, k \in \{1, \ldots, r\}.$$

Thus

$$\min_k \left\{ x^T \hat{A}^k y + c_1^k \right\} \leq cu + \alpha.$$

Similarly, from the constraints (c2) and (c8) we have

$$\min_l \left\{ x^T \hat{B}^l y + c_2^l, \right\} \leq dv + \beta.$$

From the constraints (c3) and (c4),

$$\min_k \left\{ x^T \hat{A}^k y + c_1^k \right\} \geq \lambda_1,$$

and

$$\min_l \left\{ x^T \hat{B}^l y + c_2^l \right\} \geq \lambda_2.$$

Therefore, we have

$$\lambda_1 + \lambda_2 - cu - dv - \alpha - \beta \leq 0.$$

In other words, the objective function value of (NLP1) is less than or equal to 0. We can verify that the solution $(\bar{x}, \bar{y}, \bar{\alpha}, \bar{\beta}, \bar{\lambda}_1, \bar{\lambda}_2, \bar{u}, \bar{v})$ in which

$$
\begin{array}{l}
\bar{x} = x^* \in \hat{X},\\
\bar{y} = y^* \in \hat{Y},\\
\bar{\alpha} = \min_k \left\{ x^{*T} \hat{A}^k y^* + c_1^k \right\} - cu^*,\\
\bar{\beta} = \min_l \left\{ x^{*T} \hat{B}^l y^* + c_2^l \right\} - dv^*,\\
\bar{\lambda}_1 = \min_k \left\{ x^{*T} \hat{A}^k y^* + c_1^k \right\},\\
\bar{\lambda}_2 = \min_l \left\{ x^{*T} \hat{B}^l y^* + c_2^l \right\},\\
\bar{u} = u^*, \text{and } \bar{v} = v^*,
\end{array}
$$

is feasible to the problem (NLP1). Assume that it is not feasible. So the solution does not satisfies at least one of the two constraints (c1) and (c2) of the problem (NLP1). We assume that the constraint (c1) does not hold. The proof of the other case is similar. Thus,

$$\hat{A}^k y^* + c_1^k e_m > cu^* e_m + \alpha^* e_m. \tag{13}$$

Multiplying $x^* \in \hat{X}$ both sides of (13), we have

$$x^{*T} \hat{A}^k y^* + c_1^k > cu^* + \alpha^*.$$

Thus,

$$\min_k \left\{ x^{*T} \hat{A}^k y^* + c_1^k \right\} > \min_k \left\{ x^{*T} \hat{A}^k y^* + c_1^k \right\}.$$

Also, its objective function value is zero. It means that the optimal value of objective of the nonlinear problem (NLP1) is zero.

Now, Let $(x^*, y^*, \alpha^*, \beta^*, \lambda_1^*, \lambda_2^*, u^*, v^*)$ be an optimal solution to the programming problem (NLP1). Thus

$$\lambda_1^* + \lambda_2^* - cu^* - dv^* - \alpha^* - \beta^* = 0. \tag{14}$$

From the constraints (c1) and (c2), for each feasible solution $(x, y, \alpha, \beta, \lambda_1, \lambda_2, u, v)$, we have

$$\begin{array}{ll} x^T \hat{A}^k y^* + c_1^k \leq cu^* + \alpha^* & \text{for some k} \in \{1, \dots, m\}, \\ x^{*T} \hat{B}^l y + c_2^l \leq dv^* + \beta^* & \text{for some l} \in \{1, \dots, n\}. \end{array}$$

Thus,

$$\begin{array}{l} \min_k \left\{ x^T \hat{A}^k y^* + c_1^k \right\} \leq cu^* + \alpha^*, \\ \min_l \left\{ x^{*T} \hat{B}^l y + c_1^l \right\} \leq dv^* + \beta^*. \end{array} \tag{15}$$

Also, according to the constraints (c3) and (c4),

$$\begin{array}{l} \min_k \left\{ x^{*T} \hat{A}^k y^* + c_1^k \right\} \geq \lambda_1^*, \\ \min_l \left\{ x^{*T} \hat{B}^l y^* + c_2^l \right\} \geq \lambda_2^*. \end{array} \tag{16}$$

Now, from (14) and (16) it follows that

$$\alpha^* + \beta^* + cu^* + dv^* = \lambda_1^* + \lambda_2^* \leq \min_k \{x^{*T} \hat{A}^k y^* + c_1^k\} + \min_l \{x^{*T} \hat{B}^l y^* + c_2^l\}.$$

Therefore,

$$\min_k \{x^{*T} \hat{A}^k y^* + c_1^k\} \geq - \min_l \{x^{*T} \hat{B}^l y^* + c_2^l\} + \alpha^* + \beta^* + cu^* + dv^*. \tag{17}$$

On the other hand, the first inequality of (15) for $x = x^*$ and (17) imply that

$$cu^* + \alpha^* \geq \min_k\{x^{*T}\hat{A}^k y^* + c_1^k\} \geq -\min_l\{x^{*T}\hat{B}^l y^* + c_2^l\} + \alpha^* + \beta^* + cu^* + dv^*.$$

It follows that

$$\min_l\{x^{*T}\hat{B}^l y^* + c_2^l\} \geq \beta^* + dv^*. \tag{18}$$

Similarly,

$$\min_k\{x^{*T}\hat{A}^k y^* + c_1^k\} \geq \alpha^* + cu^*. \tag{19}$$

Thus, from (15), (18), and (19), we have

$$\begin{aligned} \min\nolimits_k\{x^{*T}\hat{A}^1 y^* + c_1^1\} = \alpha^* + cu^*, \\ \min\nolimits_l\{x^{*T}\hat{B}^l y^* + c_2^l\} = \beta^* + dv^*. \end{aligned} \tag{20}$$

Consequently, (15) and (20) imply that

$$\min_k\{x^{*T}\hat{A}^k y^* + c_1^k\} \geq \min_k\{x^T\hat{A}^k y^* + c_1^k\}, \qquad \forall x \in \hat{X},$$

and

$$\min_l\{x^{*T}\hat{B}^l y^* + c_2^l\} \geq \min_l\{x^{*T}\hat{B}^l y + c_2^l\} \qquad \forall y \in \hat{Y}.$$

Hence, x^* and y^* are, respectively, optimal solutions to the problems (11) and (12). Therefore, by Remark 2 the pair (x^*, y^*) is an equilibrium solution with respect to the degree of attainment of the fuzzy goals aggregated by the minimum component for the multiobjective game. □

Using the techniques of solving nonlinear programming problems containing constraints such as (c1) and (c2), the nonlinear problem (NLP1) can be solved by solving the following mixed binary nonlinear programming problem.

$$\begin{aligned}
&\max\ \lambda_1 + \lambda_2 - cu - dv - \alpha - \beta \\
&\text{s.t.} \\
&\quad \hat{A}^1 y + c_1^1 e_m \leq cue_m + \alpha e_m + M(1-\delta_1)e_m, \\
&\quad \vdots \\
&\quad \hat{A}^r y + c_1^r e_m \leq cue_m + \alpha e_m + M(1-\delta_r)e_m, \\
&\quad (\hat{B}^1)^T x + c_2^1 e_n \leq dve_n + \beta e_n + M(1-\delta_1^{'})e_n, \\
&\quad \vdots \\
&\quad (\hat{B}^s)^T x + c_2^s e_n \leq dve_n + \beta e_n + M(1-\delta_s^{'})e_n, \\
&\quad x^T\hat{A}^k y + c_1^k \geq \lambda_1, \qquad k = 1, \ldots, r, \\
&\quad x^T\hat{B}^l y + c_2^l \geq \lambda_2, \qquad l = 1, \ldots, s,
\end{aligned}$$

$$
\begin{aligned}
& D^T x \leq c, \\
& E^T y \leq d, \\
& e_m^T x = 1, \\
& e_n^T y = 1, \\
& 0 \leq c_1^k + x^T \hat{A}^k y \leq 1, \\
& 0 \leq c_2^l + x^T \hat{B}^l y \leq 1, \\
& 0 \leq \lambda_1, \lambda_2 \leq 1, \\
& \sum_{k=1}^{r} \delta_k = 1, \\
& \sum_{l=1}^{s} \delta_l^{'} = 1, \\
& \delta_k = 0 \, or \, 1, \qquad k = 1, \ldots r, \\
& \delta_l^{'} = 0 \, or \, 1, \qquad l = 1, \ldots s, \\
& x, y, u, v \geq 0,
\end{aligned}
$$

where M is an enough large number. This problem can be solved by branch and bound methods of solving mixed zero-one nonlinear problems [14], or by using optimization softwares such as Lingo [25] and Gams [26].

3.2 Multiobjective constrained bimatrix games with fuzzy goals and fuzzy constraints

Assume that the constraints imposed on mixed strategies are fuzzy constraints. In other words, the mixed strategy spaces for Players I and II are, respectively, as follows:

$$\tilde{X} = \left\{ x \in \mathbb{R}^m \left| e_m^T x = 1, Dx \tilde{\geq} c \ , \ x \geq 0 \right. \right\},$$

$$\tilde{Y} = \left\{ y \in \mathbb{R}^n \left| e_n^T y = 1, Ey \tilde{\geq} d \ , \ y \geq 0 \right. \right\}.$$

The symbol "$\tilde{\geq}$" denotes a relaxed or fuzzy version of the ordinary inequality "$\geq$". (We consider only fuzzy version of $\geq$ in order to adopt to the considered fuzzy goals. Note that a $\leq$ constraint can be transformed to a $\geq$ one). More explicitly, these fuzzy inequalities mean that "for $t = 1, \ldots, p$ and $\tau = 1, \ldots, q$, the constraints $D_t x$ and $E_\tau y$ should be substantially more than or equal to c_t and d_τ, respectively ".

For treating the fuzzy inequality $D_t x \tilde{\geq} c_t$, we use the following linear membership function

$$\mu_1(D_t x) = \begin{cases} 0, & D_t x \leq \underline{c_t}, \\ 1 - \frac{\bar{c}_t - D_t x}{\bar{c}_t - \underline{c_t}}, & \underline{c_t} \leq D_t x \leq \bar{c}_t, \\ 1, & D_t x \geq \bar{c}_t, \end{cases} \tag{21}$$

where $\bar{c_t} = c_t$ and $\underline{c_t} = c_t - \delta_t$, in which δ_t is a subjectively chosen constant expressing the limit of the admissible violation of the inequality. Letting $\hat{D}_t = \frac{D_t}{\bar{c_t}-\underline{c_t}}$ and $c_{1t}^{'} = -\frac{c_t}{\bar{c_t}-\underline{c_t}}$, the above membership function can be rewritten as follows:

$$\mu_1(D_t x) = \begin{cases} 0, & D_t x \leq \underline{c_t}, \\ c_{1t}^{'} + \hat{D}_t x, & \underline{c_t} \leq D_t x \leq \bar{c_t}, \\ 1, & D_t x \geq \bar{c_t}. \end{cases} \tag{22}$$

Also, $\mu_2(E_\tau y)$ can be defined similar to (21) and rewritten as

$$\mu_2(E_\tau y) = \begin{cases} 0, & E_\tau y \leq \underline{d_\tau}, \\ c_{2\tau}^{'} + \hat{E}_\tau y, & \underline{d_\tau} \leq E_\tau y \leq \bar{d_\tau}, \\ 1, & E_\tau y \geq \bar{d_\tau}, \end{cases} \tag{23}$$

where $\hat{E}_\tau = \frac{E_\tau}{\bar{d_\tau}-\underline{d_\tau}}$, $c_{2\tau}^{'} = -\frac{d_\tau}{\bar{d_\tau}-\underline{d_\tau}}$, $\bar{d_\tau} = d_\tau$, and $\underline{d_\tau} = d_\tau - \delta_\tau^{'}$.

According to these membership functions, we define an equilibrium solution with respect to the degree of attainment of the aggregated fuzzy goal as follows.

Definition 8. A pair of strategies (x^*, y^*) is an equilibrium solution with respect to the degree of attainment of the fuzzy goals aggregated by the minimum component for the multiobjective $CBG : (\tilde{X}, \tilde{Y}, A, B)$ with fuzzy goals if for any other mixed strategies $x \in \tilde{X}$ and $y \in \tilde{Y}$

$$\min_{\substack{k=1,\ldots,r \\ t=1,\ldots,p}} \left\{\mu_1^k(x^{*T} A^k y^*), \mu_1(D_t x^*)\right\} \geq \min_{\substack{k=1,\ldots,r \\ t=1,\ldots,p}} \left\{\mu_1^k(x^T A^k y^*), \mu_1(D_t x)\right\},$$

and

$$\min_{\substack{l=1,\ldots,s \\ \tau=1,\ldots,q}} \left\{\mu_2^l(x^{*T} B^l y^*), \mu_2(E_\tau y^*)\right\} \geq \min_{\substack{l=1,\ldots,s \\ \tau=1,\ldots,q}} \left\{\mu_2^l(x^{*T} B^l y), \mu_2(E_\tau y)\right\}.$$

According to the membership functions (22) and (23), we can say that an equilibrium point is a pair of strategies (x^*, y^*) in which x^* and y^* are the optimal solutions of the following problems, respectively.

$$\begin{array}{ll} maximize & \min\{c_1^1 + x^T \hat{A}^1 y^*, \ldots, c_1^r + x^T \hat{A}^r y^*, c_{11}^{'} + \hat{D}_1 x, \ldots, c_{1p}^{'} + \hat{D}_p x\} \\ s.t. & e_m^T x = 1, \\ & 0 \leq c_1^k + x^T \hat{A}^k y^* \leq 1, \quad k = 1, \ldots, r, \\ & 0 \leq c_{1t}^{'} + \hat{D}_t x \leq 1, \qquad t = 1, \ldots, p, \\ & x \geq 0, \end{array} \tag{24}$$

and

$$
\begin{array}{lll}
maximize & \min\{c_2^1 + x^{*T}\hat{B}^1 y, \ldots, c_2^s + x^{*T}\hat{B}^s y, c_{21}^{'} + \hat{E}_1 y, \ldots, c_{2q}^{'} + \hat{E}_q y\} & \\
s.t. & e_n^T y = 1, & \\
 & 0 \leq c_2^l + x^{*T}\hat{B}^l y \leq 1, \quad l = 1, \ldots, s, & \\
 & 0 \leq c_{2\tau}^{'} + \hat{E}_\tau y \leq 1, \qquad \tau = 1, \ldots, q, & \\
 & y \geq 0. &
\end{array} \tag{25}
$$

Note that the problems (24) and (25) are similar to the problems presented in [17] for solving bimatrix games with fuzzy goals in which additional fuzzy constraints imposed on the strategies becomes additional fuzzy goals. Thus it is natural to extend the presented theorems in [17]. The following theorem expresses that an equilibrium point can be obtained by solving a mathematical programming problem.

Theorem 6. *(Equivalence theorem). Let $(x^*, y^*, \alpha^*, \beta^*, \lambda_1^*, \lambda_2^*)$ be an optimal solution to the following nonlinear programming problem:*

$$
\begin{array}{lll}
\textit{(NLP2):} & & \\
\max \lambda_1 + \lambda_2 - \alpha - \beta & & \\
s.t. & & \\
\begin{cases} \hat{A}^k y + c_1^k e_m - \alpha e_m \leq 0 & \textit{for some } k \in \{1, \ldots, r\}, \\ \\ \hat{D}_t + c_{1t}^{'} e_m \leq \alpha e_m & \textit{for some } t \in \{1, \ldots, p\}, \end{cases} & \\
\begin{cases} (\hat{B}^l)^T x + c_2^l e_n - \beta e_n \leq 0 & \textit{for some } l \in \{1, \ldots, s\}, \\ \\ \hat{E}_\tau + c_{2\tau}^{'} e_n \leq \beta e_n & \textit{for some } \tau \in \{1, \ldots, q\}, \end{cases} & \\
x\hat{A}^k y + c_1^k \geq \lambda_1, & k = 1, \ldots, r, \\
\hat{D_t}^T x + c_{1t}^{'} \geq \lambda_1, & t = 1, \ldots, p, \\
x\hat{B}^l y + c_2^l \geq \lambda_2, & l = 1, \ldots, s, \\
\hat{E_\tau}^T y + c_{2\tau}^{'} \geq \lambda_2, & \tau = 1, \ldots, q, \\
e_m^T x = 1, & \\
e_n^T y = 1, & \\
0 \leq c_1^k + x\hat{A}^k y, \leq 1 & k = 1, \ldots, r, \\
0 \leq c_{1t}^{'} + \hat{D}_t x \leq 1, & t = 1, \ldots, p, \\
0 \leq c_2^l + x^T\hat{B}^l y \leq 1, & l = 1, \ldots, s, \\
0 \leq c_{2\tau}^{'} + \hat{E}_\tau y \leq 1, & \tau = 1, \ldots, q, \\
0 \leq \lambda_1, \lambda_2 \leq 1, & \\
x, y \geq 0. &
\end{array}
$$

Then (x^*, y^*) is an equilibrium solution with respect to the degree of attainment of the aggregated fuzzy goal.

Proof. The proof is similar to the proof of Theorem 5. □

Note that if the fuzzy constraints imposed on mixed strategies are not considered, the problem (NLP2) is the same as the introduced problem by Nishizaki and Sakawa [17] to solve bimatrix game with fuzzy goals.

Similar to the problem with crisp constraints in subsection 3.1, to solve this problem, it is sufficient to solve the following mixed binary nonlinear programming.

$$
\begin{aligned}
&\max \lambda_1 + \lambda_2 - \alpha - \beta \\
&s.t. \\
&\quad \hat{A}^1 y + c_1^1 e_m \le \alpha e_m + M(1-\delta_1)e_m, \\
&\quad \vdots \\
&\quad \hat{A}^r y + c_1^r e_m \le \alpha e_m + M(1-\delta_r)e_{m,'} \\
&\quad \hat{D}_1 + c_{11}^{'} e_m \le \alpha e_m + M(1-\delta_{r+1})e_m, \\
&\quad \vdots \\
&\quad \hat{D}_p + c_{1p}^{'} e_m \le \alpha e_m + M(1-\delta_{r+p})e_m, \\
&\quad (\hat{B}^1)^T x + c_2^1 e_n \le \beta e_n + M(1-\delta_1^{'})e_n, \\
&\quad \vdots \\
&\quad (\hat{B}^s)^T x + c_2^s e_n \le \beta e_n + M(1-\delta_s^{'})e_n, \\
&\quad \hat{E}_1 + c_{21}^{'} e_n \le \beta e_n + M(1-\delta_{s+1}^{'})e_n, \\
&\quad \vdots \\
&\quad \hat{E}_q + c_{2q}^{'} e_n \le \beta e_n + M(1-\delta_{s+q}^{'})e_n, \\
&\quad x^T \hat{A}^k y + c_1^k \ge \lambda_1, \quad k = 1, \dots, r, \\
&\quad \hat{D}_t^T x + c_{1t}^{'} \ge \lambda_1, \quad t = 1, \dots, p, \\
&\quad x^T \hat{B}^l y + c_2^l \ge \lambda_2, \quad l = 1, \dots, s, \\
&\quad \hat{E}_\tau^T y + c_{2\tau}^{'} \ge \lambda_2, \quad \tau = 1, \dots, q, \\
&\quad e_m^T x = 1, \\
&\quad e_n^T y = 1, \\
&\quad 0 \le c_1^k + x\hat{A}^k y \le 1, \quad k = 1, \dots, r, \\
&\quad 0 \le c_{1t}^{'} + \hat{D}x \le 1, \quad t = 1, \dots, p, \\
&\quad 0 \le c_2^l + x^T \hat{B}^l y \le 1, \quad l = 1, \dots, s, \\
&\quad 0 \le c_{2\tau}^{'} + \hat{E}y \le 1, \quad \tau = 1, \dots, q, \\
&\quad 0 \le \lambda_1, \lambda_2 \le 1, \\
&\quad \sum_{k=1}^{r+p} \delta_k = 1, \\
&\quad \sum_{l=1}^{s+q} \delta_l^{'} = 1, \\
&\quad \delta_k = 0 \, or \, 1, \quad k = 1, \dots r+p, \\
&\quad \delta_l^{'} = 0 \, or \; 1, \quad l = 1, \dots s+q, \\
&\quad x, y, u, v \ge 0.
\end{aligned}
$$

Again, the above problem can be solved by branch and bound methods of solving mixed zero-one nonlinear problems [14], or by using optimization softwares such as Lingo [25] and Gams [26].

In this section, We use multiobjective optimization to study multiobjective two-person constrained bimatrix games. Thus, we expect to consider Pareto optimality concept for multiobjective two-person bimatrix games. Before defining weak Pareto optimal equilibrium solution, we recall the definition of the best reply strategies as follows by using the concept of Pareto optimality in multiobjective optimization.

Definition 9. [17]. Let a payoff vector of Player I be denoted by $p_1(x,y) \in \mathbb{R}^r$ when Player I chooses a mixed strategy $x \in \hat{X}$, and let Player II chooses a mixed strategy $y \in \hat{Y}$. Player I's preference cone is defined by

$$C_1 = \left\{ z = (z^1, \ldots, z^r) \in \mathbb{R}^r | z^k \geq 0,\, k = 1, \ldots, r \right\}.$$

Then, given Player II's strategy $\hat{y}$, the set of payoffs for the weak Pareto best reply strategies is defined by

$$p_1(\hat{y}) = \{p_1(x,\hat{y}) \in z_1(\hat{y}) | z_1(\hat{y}) \cap (p_1(x,\hat{y}) + intC_1) = \varnothing$$

$$for\, some\, strategy\, x \in \hat{X}\, of\, Player\, I\}$$

where $z_1(\hat{y})$ is the set of attainable payoffs of Player I against the strategy $\hat{y} \in \hat{Y}$ of Player II and $intC_1$ denotes the set of interior points of C_1. Similarly, let Player II's payoff vector be denoted by $p_2(x,y) \in \mathbb{R}^s$, and let Player II's preference cone be

$$C_2 = \left\{ z = (z^1, \ldots, z^s) \in \mathbb{R}^s | z^l \geq 0,\, l = 1, \ldots, s \right\}.$$

Then, given Player I's strategy $\hat{x} \in \hat{X}$, the set of payoffs for weak Pareto best reply strategies is defined by

$$p_2(\hat{x}) = \{p_2(\hat{x},y) \in z_2(\hat{x}) | z_2(\hat{x}) \cap (p_2(\hat{x},y) + intC_2) = \varnothing,$$

$$for\, some\, strategy\, y \in \hat{Y}\, of\, Player\, II\},$$

where $z_2(\hat{x})$ is the set of attainable payoffs of Player II against the strategy $\hat{x}$ of Player I.

Definition 10. [23]. Let the payoff vectors of Players I and II be $p_1(x,y) = (p_1^1(x,y), \ldots, p_1^r(x,y))$ and $p_2(x,y) = (p_2^1(x,y), \ldots, p_2^s(x,y))$, respectively. For any pair of strategies $x \in \hat{X}$ and $y \in \hat{Y}$, let Player I's set of payoff vectors for the weak Pareto best reply strategies and Player II's set of payoff vectors for the weak Pareto best reply strategies be denoted by $p_1(y)$ and $p_2(x)$, respectively. Then the set of the weak Pareto optimal equilibrium solutions is defined by

$$WPE = \{(x^*, y^*) | p_1^*(x^*, y^*) \in p_1(y^*),\, p_2^*(x^*, y^*) \in p_2(x^*)\}.$$

Pareto optimal equilibrium solution can be defined by replacing $intC_1$ and $intC_2$ with $C_1 \backslash \{0\}$ and $C_2 \backslash \{0\}$ in the above definitions, respectively.
Wierzbicki [23] explored in detail the relation between scalarizing functions and Pareto optimal equilibrium solutions. If scalarizing function is a strictly monotone function, then the obtained equilibrium point is Pareto optimal

equilibrium solution, and if it is monotone function which is not always strictly monotone, then obtained equilibrium point is weak Pareto optimal equilibrium solution. Also, suppose a player assesses that $\underline{a}^k$ or $\underline{b}^l$ is sufficiently small and assesses that $\bar{a}^k$ or $\bar{b}^l$ is sufficiently large. We have the following theorem.

Theorem 7 *An equilibrium solution with respect to the degree of attainment of the fuzzy goals aggregated by a minimum component is a weak Pareto optimal equilibrium solution.*

Proof. According to that the minimum component is not always strictly monotone, the proof is completed. □

4 A political application: investigating nuclear negotiations

The application of game theory and its importance to social science were quickly recognized. Now, game theory is used extensively in analyzing psychology, philosophy, sociology, politics, and economics. Most importantly, game theory is used to analyze international relations, specifically in countries with conflicting goals and interests. Game theory provides a logical analysis of situations of conflict and cooperation. A game is a situation in which

1) There are at least two players. A player may be an individual, but it may also be a more general entity like a company, a nation, or even a biological species.

2) Each player has a number of possible strategies, courses of action which he may choose to follow.

3) The strategies chosen by each player determine his outcome of the game.

4) Associated to each possible outcome of the game there is a collection of numerical payoffs, one to each player. These payoffs represent the value of the outcome to the different players.

Game theory can be divided into two categories: zero-sum games and nonzero-sum games. Zero-sum games are games in which one player wins and the other loses. The two players do not cooperate, and their interests are in total conflict. Each player chooses a certain set of strategies, and he does not know the choices of the other player. Nonzero-sum games, on the other hand, are games in which the interests of the players are not strictly opposed. The success of one player does not come as a result of the failure of the other player.

Because the interests of the players are not in total conflict and not strictly coincident, a nonzero-sum game allows the players to both compete and cooperate to achieve outcomes that are advantageous to both players. Nonzero-sum games can be played without communication, with communication before the game, and with cooperation. In international relations, countries compete and cooperate to maximize their national interests. These countries also have diplomatic relations, and therefore, they communicate with one another. Therefore, this study will focus on the nonzero-sum games in which the players can communicate before the game strategic moves.

It is important to note that the data in this discussion are bogus totally and is brainchild of the authors and have not specific explanation for what happened in the world of today between two countries.

Let us explain the details of the problem. Country I is sanctioned by Country II, and its allies due to nuclear activities. Now, Country I is not in a good economic position. To fix the economic problem, Country I decides to negotiate with Country II to lift nuclear sanctions and achieve the favorable economic situation in the region. However, it is possible to Country I find the weak political position among its allies and its people. It is possible to improve the political situation in the other countries due to negotiation but Country I is not optimistic to this improvement. Since Country I is interested in a win-win game, so we formulate our problem as a win-win game. In this negotiation it is assumed that after the establishment of results, the countries act on their decisions. During the negotiations, Country I's strategies are as follows:

(1) Do not reduce the number of centrifuges (even increase their numbers) and extend nuclear activities.

(2) Reduce of the number of centrifuges and limit the nuclear activities under strict inspections.

(3) Suspend the nuclear activities.

Also, Country II's strategies are as follows:

(1) Do not cancel any of sanctions.

(2) Cancel some natural or legal sanctions.

(3) Military attack.

Country II has stated that if Country I chooses the strategies (2) or (3), then he may choose second strategy. Both countries do not trust each other to do after agreement. So data should be chosen in such a way that the distrust between two countries be considered. Here we assume that both countries said that they would act after the agreement. Note that there are opponents of this agreement in congress of Country II. Thus, it is expected that after the agreement, by congressional pressure on the government some

Country II

Country I	(50,30)	(80,5)	(45,10)
	(20,70)	(20,50)	(30,15)
	(10,60)	(10,30)	(15,5)

Political position

Country II

Country I	(60,30)	(70,40)	(5,20)
	(1,50)	(50,50)	(0,10)
	(5,40)	(20,40)	(5,10)

Economic position

new sanctions are imposed against Country I way out of nuclear activities. Country II choosing the first strategy in addition to sanctions, concerned to military threats and create sedition in Country I. It is also possible, even move with second strategy to military threats and create sedition in this country. However, with Country II's commitment, Country I predicts these actions after the agreement. Although it is possible mathematical models of these issues are complex in real world, our aim of this discussion is only to show a part of applications of these games in the nuclear negotiations and the method of determining the best strategy.

Suppose that Country I will incur cost as much as 55, 70, and 80 unites if he chooses the strategies (1), (2), and (3), respectively. Country I does not want to spend more than 75 unites; that is, the mixed strategies of the country I must satisfy the constraint condition: $55x_1 + 70x_2 + 80x_3 \leq 75$.

Country II will incur cost as much as 30, 60, and 90 unites, by choosing the strategies (1), (2), and (3), respectively. However, Country II only provides 80 unites;that is, the mixed strategies of the country II must satisfy the constraint condition: $30y_1 + 60y_2 + 90y_3 \leq 80$.

Assume that the payoff matrices of this game in view of political and economic positions are given in Table 1. The numbers in these tables are the payoffs for the countries. Country I's payoffs are the first numbers, and Country II's payoffs are second numbers.

We ask with players the amounts which are satisfied with them and the limit of the admissible violation of the inequalities. Let fuzzy goals $\tilde{G}_1^1$ and $\tilde{G}_1^2$ of the Country I for the two objectives be represented by the following linear membership functions:

$$\mu_1(x^T A^1 y) = \begin{cases} 0, & x^T A^1 y \leq 20, \\ 1 - \frac{40 - xA^1 y}{40-20}, & 20 \leq x^T A^1 y \leq 40, \\ 1, & x^T A^1 y \geq 40, \end{cases}$$

and

$$\mu_1(x^TA^2y)=\begin{cases}0, & x^TA^2y\leq 15,\\ 1-\frac{50-xA^2y}{25-15}, & 15\leq x^TA^2y\leq 50,\\ 1, & x^TA^2y\geq 50,\end{cases}$$

Let fuzzy goals $\tilde{G}_2^1$ and $\tilde{G}_2^2$ of the Country II for the two objectives be represented by the following linear membership functions:

$$\mu_2(x^TB^1y)=\begin{cases}0, & x^TB^1y\leq 30,\\ 1-\frac{40-xB^1y}{40-30}, & 10\leq x^TB^1y\leq 40,\\ 1, & x^TB^1y\geq 40,\end{cases}$$

and

$$\mu_2(x^TB^2y)=\begin{cases}0, & x^TB^2y\leq 35,\\ 1-\frac{40-xB^2y}{45-35}, & 20\leq x^TB^2y\leq 40,\\ 1, & x^TB^2y\geq 40.\end{cases}$$

By solving the mixed binary mathematical programming problem (NLP1), the equilibrium point is obtained. Thus, equilibrium solutions are optimal solutions to the following problem.

$$\begin{aligned}
&\max\ \lambda_1+\lambda_2-75u-80v-\alpha-\beta\\
&\text{s.t.}\\
&\quad 2.5y_1+4y_2+2.25y_3-75u-\alpha\leq 1+M\delta_1,\\
&\quad y_1+y_2+1.5y_3-75u-\alpha\leq 1+M\delta_1,\\
&\quad 0.5y_1+0.5y_2+0.75y_3-75u-\alpha\leq 1+M\delta_1,\\
&\quad 12/7y_1+2y_2+1/7y_3-75u-\alpha\leq 3/7+M\delta_2,\\
&\quad 1/35y_1+10/7y_2+0y_3-75u-\alpha\leq 3/7+M\delta_2,\\
&\quad 1/7y_1+4/7y_2+1/7y_3-75u-\alpha\leq 3/7+M\delta_2,\\
&\quad 3x_1+7x_2+6x_3-80v-\beta\leq 3+M\delta_1^{'},\\
&\quad 0.5x_1+5x_2+3x_3-80v-\beta\leq 3+M\delta_1^{'},\\
&\quad 1x_1+1.5x_2+0.5x_3-80v-\beta\leq 3+M\delta_1^{'},\\
&\quad 6x_1+10x_2+8x_3-80v-\beta\leq 7+M\delta_2^{'},\\
&\quad 8x_1+10x_2+8x_3-80v-\beta\leq 7+M\delta_2^{'},\\
&\quad 4x_1+4x_2+2x_3-80v-\beta\leq 7+M\delta_2^{'},\\
&\quad 4/7x_1y_1+1/35x_2y_1+1/7x_3y_1+2x_1y_2+10/7x_2y_2+12/7x_3y_2\\
&\quad\ +1/7x_1y_3+0x_2y_3+1/7x_3y_3-\lambda_1\geq 3/7,\\
&\quad 6x_1y_1+10x_2y_1+8x_3y_1+8x_1y_2+10x_2y_2+8x_3y_2\\
&\quad\ +4x_1y_3+2x_2y_3+2x_3y_3-\lambda_2\geq 7,\\
&\quad 1.5x_1y_1+0.5x_2y_1+0.25x_3y_1+4x_1y_2+0.5x_2y_2+0.5x_3y_2\\
&\quad\ +0.5x_1y_3+1.5x_2y_3+0.25x_3y_3-\lambda_1\geq 1,
\end{aligned}$$

$$
\begin{aligned}
&3x_1y_1 + 7x_2y_1 + 6x_3y_1 + 0.5x_1y_2 + 5x_2y_2 + 3x_3y_2\\
&\ +1x_1y_3 + 1.5x_2y_3 + 0.5x_3y_3 - \lambda_2 \geq 3,\\
&55x_1 + 70x_2 + 80x_3 \leq 75,\\
&30y_1 + 60y_2 + 90y_3 \leq 80,\\
&x_1 + x_2 + x_3 = 1,\\
&y_1 + y_2 + y_3 = 1,\\
&12/7x_1y_1 + 1/35x_2y_1 + 1/7x_3y_1 + 2x_1y_2 + 10/7x_2y_2 + 4/7x_3y_2\\
&\ +1/7x_1y_3 + 0x_2y_3 + 1/7x_3y_3 - \lambda_1 \geq 3/7,\\
&2.5x_1y_1 + 1x_2y_1 + 0.5x_3y_1 + 4x_1y_2 + 0.5x_2y_2 + 0.5x_3y_2\\
&\ +2.25x_1y_3 + 1.5x_2y_3 + 0.5x_3y_3 - \lambda_1 \geq 1,\\
&3x_1y_1 + 7x_2y_1 + 6x_3y_1 + 0.5x_1y_2 + 5x_2y_2 + 3x_3y_2\\
&\ +1x_1y_3 + 1.5x_2y_3 + 0.5x_3y_3 - \lambda_1 \geq 3,\\
&6x_1y_1 + 10x_2y_1 + 8x_3y_1 + 8x_1y_2 + 10x_2y_2 + 8x_3y_2\\
&\ +4x_1y_3 + 2x_2y_3 + 2x_3y_3 - \lambda_1 \geq 7,\\
&\delta_1 + \delta_2 = 1,\\
&\delta_1^{'} + \delta_2^{'} = 1,\\
&\delta_1, \delta_2 = 0\, or\, 1,\\
&\delta_1^{'}, \delta_2^{'} = 0\, or\, 1,\\
&x_1, x_2, x_3 \geq 0,\\
&y_1, y_2, y_3 \geq 0,\\
&u, v \geq 0.
\end{aligned}
$$

This problem is solved by Lingo software. The obtained strategies are as $(x_1^*, x_2^*, x_3^*) = (0.56, 0.38, 0.06)$ for Country I and $(y_1^*, y_2^*, y_3^*) = (0.69, 0.31, 0)$ for Country II. This means that according to the conditions of the problem the chance of success in negotiations between the two countries is that both countries choose their first strategy.

5 Conclusion

In this paper, constrained bimatrix games with single and multiple payoffs are studied, which incorporate fuzzy goals for objectives. Such fuzzy games has not been considered in previous researches, based on the best knowledge of the authors. A programming problem with linear constraints and a quadratic objective function was presented to obtain the equilibrium solution of single objective problem of such games. In multiobjective case, equilibrium points was defined and the mathematical programming problems was introduced to obtain the equilibrium points of this class of games with fuzzy goals and the fuzzy/crisp constraints. We proposed two mixed binary nonlinear programming problems to solve these mathematical programming problems. Weak Pareto optimal equilibrium solution was defined and showed that obtained equilibrium points of these problems are weak Pareto optimal equilibrium solutions. A political application of these games was presented which investigate nuclear negotiations between two countries. The equilibrium point of this game was obtained by the proposed method. Solving constrained multi-objective bimatrix games with fuzzy payoffs and fuzzy nonlinear goals needs more researches, which can be our future work.

References

1. Bector, C. R. and Chandra, S. *Fuzzy mathematical programming and fuzzy matrix games*, Springer-Verlag, Berlin Heidelberg, 2005.

2. Bellman, R. E. and Zadeh, L. A. *Decision making in a fuzzy environment*, Management Sci. 17 (1970/71), 141–164.

3. Bigdeli, H. and Hassanpour, H. *A satisfactory strategy of multiobjective two person matrix games with fuzzy payoffs*, Iranian Journal of Fuzzy Systems, 13 (2016), 17–33.

4. Bigdeli, H., Hassanpour, H., and Tayyebi, J. *The optimistic and pessimistic solutions of single and multiobjective matrix games with fuzzy payoffs and analysis of some of military problems.* Passive Defence Sci. & Tech. 2 (2017), 133–145.

5. Borm, P. E. M., Tijs, S. H., and van den Aarssen, J. C. M. *Pareto equilibria in multiobjective games*, XIII Symposium on Operations Research (Paderborn, 1988), 303–312, Methods Oper. Res., 60, Hain, Frankfurt am Main, 1990.

6. Charnes, A. *Constrained games and linear programming*, Proc. Nat. Acad. Sci. USA 39 (1953), 639–641.

7. Corley, H. W. *Games with vector payoffs*, J. Optim. Theory Appl. 47 (1985), no. 4, 491–498.

8. Dresher, M. *Games of strategy theory and applications*, Prentice-Hall Applied Mathematics Series Prentice-Hall, Inc., Englewood Cliffs, N.J. 1961.

9. Fahem, K. and Radjef, M. S. *Properly efficient nash equilibrium in multicriteria non-cooperative games*, Math. Methods Oper. Res. 82(2015), no. 2, 175–193.

10. Kawaguchi, T. and Maruyama, Y. *A note on minmax (maxmin) programming*, Management Science, 22 (1976), 670–676.

11. Li, D. F. and Cheng, C. *Fuzzy multiobjective programming methods for fuzzy constrained matrix games with fuzzy numbers*, Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 10 (2002), no. 4, 385–400.

12. Li, D. F. and Hong, F. X. *Alfa-cut based linear programming methodology for constrained matrix games with payoffs of trapezoidal fuzzy numbers*, Fuzzy Optim. Decis. Mak. 12 (2013), no. 2, 191–213.

13. Li, D. F. and Hong, F. X. *Solving constrained matrix games with payoffs of triangular fuzzy numbers*, Comput. Math. Appl. 64 (2012), no. 4, 432–446.

14. Li, D. and Sun, X. *Nonlinear integer programming*, Springer Science, New York, 2006.

15. Mangasarian, O. L. and Stone, H. *Two-person nonzero-sum games and quadratic programming*, J. Math. Anal. Appl. 9 (1964), 348–355.

16. Neumann, J. V. and Morgenstern, O. *Theory of games and economic behavior*, Wiley, New York, 1944,

17. Nishizaki, I. and Sakawa, M. *Fuzzy and multiobjective games for conflict resolution*, Springer-Verlag, Berlin Heidelberg, 2001.

18. Nishizaki, I. and Sakawa, M. *Equilibrium solutions for multiobjective bimatrix games incorporating fuzzy goals*, J. Optim. Theory Appl. 86 (1995), no. 2, 433–458.

19. Owen, G. *Game theory*, Academic Press, San Diego, Second Edition 1982, Third Edition, 1995.

20. Parthasarathy, T. and Raghavan, T. E. S. *Some topics in two-Person games*, No.22 American Elsevier Publishing Company, New York, 1971.

21. Sakawa, M. *Fuzzy sets and interactive multiobjective optimization*, Plenum press, New York, 1993.

22. Steuer, R. *Multiple criteria optimization:theory, computation, and application*, John Wiley & Sons, New York, 1986.

23. Wierzbicki, A. P. *Multiple criteria solutions in noncooperative game-theory part III: theoretical foundations.* Kyoto Institute of Economic Research Discussion paper, No. 288, 1990.

24. Zadeh, L. A. *The concept of a linguistic variable and its application to approximate reasoning*, Inform. Sci. 8 (1975), 199–249.

25. http:// www.lindo.com

26. http:// www.gams.com

Iranian Journal of Numerical Analysis and Optimization
Vol. 8, No. 1, (2018), pp 111–128
DOI:10.22067/ijnao.v8i1.57845

Some efficient Nordsieck integration methods for IVPs

N. Barghi Oskouie, A. Abdi* and G. Hojjati

Abstract

In this paper, in continuation of the construction of efficient numerical methods for stiff IVPs, we construct type two Nordsieck second derivative general linear methods with order $p = s$, where s is the number of internal stages, and stage order $q = p$. Implementation of the constructed methods with fixed and variable stepsize is discussed which verifies their efficiency.

Keywords: Stiff differential equations; Nordsieck second derivative general linear methods; A- and L-stability; Variable stepsize implementation.

1 Introduction

Second derivative general linear methods (SGLMs) for the numerical solution of autonomous ordinary differential equations (ODEs) with initial value problem

$$\begin{aligned} &y'(x) = f(y(x)), \qquad x \in [x_0, \overline{x}], \\ &y(x_0) = y_0, \end{aligned} \tag{1}$$

*Corresponding author
Received 22 January 2017; revised 29 March 2017; accepted 23 April 2017
N. Barghi Oskouie
Faculty of Mathematical Sciences, University of Tabriz, Tabriz, Iran. e-mail: n_barghi@tabrizu.ac.ir

A. Abdi
Faculty of Mathematical Sciences, University of Tabriz, Tabriz, Iran. e-mail: a_abdi@tabrizu.ac.ir

G. Hojjati
Faculty of Mathematical Sciences, University of Tabriz, Tabriz, Iran. e-mail: ghojjati@tabrizu.ac.ir

where $f : \mathbb{R}^m \to \mathbb{R}^m$ and m is the dimensionality of the system, have been studied in recent years. These methods were introduced by Butcher and Hojjati in [11] and investigated more in [2–6] by Abdi and Hojjati. In the construction of SGLMs which are extension of general linear methods (GLMs) [10, 13–15, 17]), there are a lot of free parameters which allow us to construct high order methods with a small number of internal stages to reduce computational cost [1].

We recall that SGLMs are characterized by four integers, (p, q, r, s) where p and q are, respectively, order and stage order of the method, r is the number of input and output approximations, and s is the number of internal stages. Let $Y^{[n]} = [Y_i^{[n]}]_{i=1}^s$ be an approximation of stage order q to the vector $y(x_{n-1} + ch) = [y(x_{n-1} + c_i h)]_{i=1}^s$, and let the vectors $f(Y^{[n]}) = [f(Y_i^{[n]})]_{i=1}^s$ and $g(Y^{[n]}) = [g(Y_i^{[n]})]_{i=1}^s$ denote the stage first and second derivative values, where $g(\cdot) = f'(\cdot)f(\cdot)$, respectively. If $r = p + 1$, we can assume the input and output vectors at the step number n, $y^{[n-1]}$ and $y^{[n]}$, are approximations of order p to the Nordsieck vectors

$$\begin{bmatrix} y(x_{n-1}) \\ hy'(x_{n-1}) \\ \vdots \\ h^p y^{(p)}(x_{n-1}) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} y(x_n) \\ hy'(x_n) \\ \vdots \\ h^p y^{(p)}(x_n) \end{bmatrix},$$

respectively. In an SGLM used for the numerical solution of (1), these values are related by

$$\begin{aligned} Y^{[n]} &= h(A \otimes I_m)f(Y^{[n]}) + h^2(\overline{A} \otimes I_m)g(Y^{[n]}) + (U \otimes I_m)y^{[n-1]}, \\ y^{[n]} &= h(B \otimes I_m)f(Y^{[n]}) + h^2(\overline{B} \otimes I_m)g(Y^{[n]}) + (V \otimes I_m)y^{[n-1]}, \end{aligned} \tag{2}$$

where $n = 1, 2, \ldots, N$, $Nh = \overline{x} - x_0$, h is the stepsize, and $\otimes$ is the Kronecker product of two matrices. Here $A, \overline{A} \in \mathbb{R}^{s\times s}$, $U \in \mathbb{R}^{s\times r}$, $B, \overline{B} \in \mathbb{R}^{r\times s}$, and $V \in \mathbb{R}^{r\times r}$. The coefficients matrix V in the Nordsieck SGLM (2) has the form

$$V = \left[\begin{array}{c|c} 1 & v^T \\ \hline 0 & \dot{V} \end{array}\right],$$

$v = [v_1 \;\; v_2 \;\; \cdots \;\; v_{r-1}]^T$, $\dot{V} \in \mathbb{R}^{(r-1)\times(r-1)}$. In this paper, the methods will be restricted to the case where $p = q = r - 1 = s$ and the eigenvalues of $\dot{V}$ are zeros. The latter condition ensures zero-stability of the method.

An SGLM in Nordsieck form has order p and stage order $q = p$ if and only if [11]

$$U = C - ACK - \overline{A}CK^2,$$

$$V = E - BCK - \overline{B}CK^2,$$

where the matrices $C \in \mathbb{R}^{s\times(p+1)}$, $K \in \mathbb{R}^{(p+1)\times(p+1)}$, and $E \in \mathbb{R}^{(p+1)\times(p+1)}$ are defined by

$$C := \begin{bmatrix} 1 & \frac{c}{1!} & \frac{c^2}{2!} & \cdots & \frac{c^p}{p!} \end{bmatrix}, \qquad K := [0 \quad e_1 \quad e_2 \quad \cdots \quad e_p\,],$$

and

$$E := \exp(K) = \begin{bmatrix} 1 & \frac{1}{1!} & \frac{1}{2!} & \cdots & \frac{1}{p!} \\ 0 & 1 & \frac{1}{1!} & \cdots & \frac{1}{(p-1)!} \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \frac{1}{1!} \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

with e_j as the jth vector of canonical basis in $\mathbb{R}^{p+1}$, respectively.

The stability behavior of SGLMs is defined using the standard test problem of Dahlquist $y' = \xi y$, where ξ is a complex number. If method (2) is applied to this problem, then the stability matrix is

$$M(z) = V + \big(zB + z^2\overline{B}\big)\big(I - zA - z^2\overline{A}\big)^{-1}U,$$

where $z = h\xi$ and the stability function of the method is defined as the characteristic polynomial of $M(z)$; that is,

$$p(w,z) = \det\big(wI - M(z)\big).$$

If $p(w,z) = w^{r-1}\big(w - R(z)\big)$, the method is said to possess "Runge–Kutta stability (RKS)". GLMs and SGLMs with RKS property were studied by Butcher and Jackiewicz in [9, 12–14] and by Abdi and Hojjati in [2–7], respectively.

For economical implementation, it is assumed that the matrices A and $\overline{A}$ have a lower triangular form with the same diagonal entries λ and μ, respectively. SGLMs are divided into four types, depending on the nature of the differential system to be solved and the computer architecture that is used to implement these methods. Types 1 and 2 are those with arbitrary a_{ij} and $\overline{a}_{ij}$, where $\lambda = \mu = 0$ and $\lambda > 0$, $\mu < 0$, respectively. Such methods are appropriate, respectively, for nonstiff and stiff differential systems in a sequential computing environment. Requiring $a_{ij} = \overline{a}_{ij} = 0$, cases $\lambda = \mu = 0$ and $\lambda > 0$, $\mu < 0$ lead, respectively, to types 3 and 4 methods which can be useful, respectively, for nonstiff and stiff systems in a parallel computing environment.

The construction and implementation of type 2 SGLMs with $p = s + 1$ were discussed in [4]. Also, the construction of parallel Nordsieck SGLMs with $p = s$ and their order barriers were studied in [5]. These order barriers were obtained under the assumption of RKS property. In this paper, we are going to construct type 2 Nordsieck SGLMs with $p = s$, which are efficient methods for stiff systems. Also, efficiency of the constructed methods are shown by their implementation in a variable stepsize environment.

Next sections of this paper are organized as follows: In section 2, we construct A- and L-stable SGLMs in the Nordsieck form with RKS property of orders 2, 3, and 4. Considering Nordsieck SGLMs in the variable stepsize mode, implementation issues including local error estimation, and stepsize control are discussed in section 3. Finally, in section 4, some results of numerical experiments on some stiff test systems are presented and compared with those obtained by Nordsieck SGLMs of the same order.

2 Construction of type 2 methods

In this section, the construction of type 2 Nordsieck SGLMs of order p and stage order $q = p$ with some desired stability properties is explained.

2.1 Methods with $p = q = s = r - 1 = 2$

We construct methods with $p = q = s = r - 1 = 2$ and RKS property. We look for methods which their stability function has the form [5]

$$R(z) = \frac{1 + n_1 z + n_2 z^2 + n_3 z^3}{(1 - \lambda z - \mu z^2)^2},$$

where

$$1 + \sum_{k=1}^{3} n_k z^k = \exp(z)(1 - \lambda z - \mu z^2)^2 - \mathcal{C} z^3 + O(z^4)$$

with $\mathcal{C}$ as the error constant of the method. For this method to be A-stable, it is necessary and sufficient that $\lambda > 0$, $\mu < 0$, and so that the $E(y)$ is non-negative for all real y, where the E-polynomial is defined by

$$E(y) = \left|1 - \lambda i y + \mu y^2\right|^4 - \left|1 + n_1 i y - n_2 y^2 - n_3 i y^3\right|^2.$$

By choosing $\mathcal{C} = 10^{-4}$, a detailed calculation shows that

$$E(y) = y^4\big(E_0 + E_1 y^2 + E_2 y^4\big),$$

where

$$\begin{aligned}
E_0 &= \frac{1247}{15000} + 2\mu^2 + \lambda^2 - 2\mu - \frac{4997}{7500}\lambda + 4\lambda\mu,\\
E_1 &= -\frac{24970009}{900000000} - \lambda^4 + 2\lambda^3 - \frac{19997}{15000}\lambda^2 - 4\mu^2 + \frac{4997}{7500}\mu + \frac{4997}{15000}\lambda\\
&\quad +4\mu^3 - 2\lambda^2\mu^2 - 4\lambda^3\mu + 8\lambda\mu^2 + 8\lambda^2\mu - \frac{34997}{7500}\lambda\mu,\\
E_2 &= 7\mu^4.
\end{aligned}$$

Pairs of (λ, μ) with values in domain $[0,2] \times [-2,0]$ giving L-stability are shown in Figure 1.

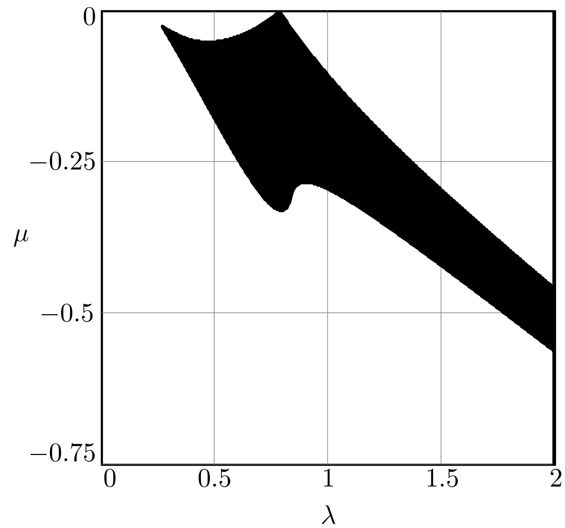


Figure 1: L-stable choices of (λ, μ) for $p = s = 2$ corresponding to $\mathcal{C} = 10^{-4}$.

We select a single example, characterized by $\lambda = \frac{4}{5}$, $\mu = -\frac{1}{5}$ and $c = [\frac{1}{2} \;\; 1]^T$. The coefficients of the method are

$$\left[\begin{array}{c|c|c} A & \overline{A} & U \\ \hline B & \overline{B} & V \end{array}\right] = \left[\begin{array}{cc|cc|ccc}
\frac{4}{5} & 0 & -\frac{1}{5} & 0 & 1 & -\frac{3}{10} & -\frac{3}{40}\\
-\frac{967}{18750} & \frac{4}{5} & \frac{494}{3375} & -\frac{1}{5} & 1 & \frac{4717}{18750} & -\frac{253}{12500}\\
\hline
-\frac{967}{18750} & \frac{4}{5} & \frac{494}{3375} & -\frac{1}{5} & 1 & \frac{4717}{18750} & -\frac{253}{12500}\\
0 & 1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0 & 0
\end{array}\right].$$

2.2 Methods with $p = q = s = r - 1 = 3$

We construct methods with $p = q = s = r - 1 = 3$ and RKS property. We look for methods which their stability function has the form [5]

$$R(z) = \frac{1 + n_1 z + n_2 z^2 + n_3 z^3 + n_4 z^4 + n_5 z^5}{(1 - \lambda z - \mu z^2)^3},$$

where

$$1 + \sum_{k=1}^{5} n_k z^k = \exp(z)(1 - \lambda z - \mu z^2)^3 - \mathcal{C}_1 z^4 - \mathcal{C}_2 z^5 + O(z^6).$$

Here $\mathcal{C}_1$ is the error constant of the method and $\mathcal{C}_2$ is an arbitrary number. The E-polynomial has the form

$$E(y) = y^6 \big(E_0 + E_1 y^2 + E_2 y^4 + E_3 y^6 \big).$$

For these methods to be A-stable, it is necessary and sufficient that $\lambda > 0$, $\mu < 0$, and so that $E_0 + E_1 x + E_2 x^2 + E_3 x^3 + E_4 x^4$ is non-negative for all positive real numbers x, where E_0, E_1, E_2, E_3, and E_4 are complicated expressions in terms of λ and μ. By choosing $\mathcal{C}_1 = \mathcal{C}_2 = 10^{-4}$, pairs of (λ, μ) with values in domain $[0, 2] \times [-2, 0]$ giving L-stability are shown in Figure 2.

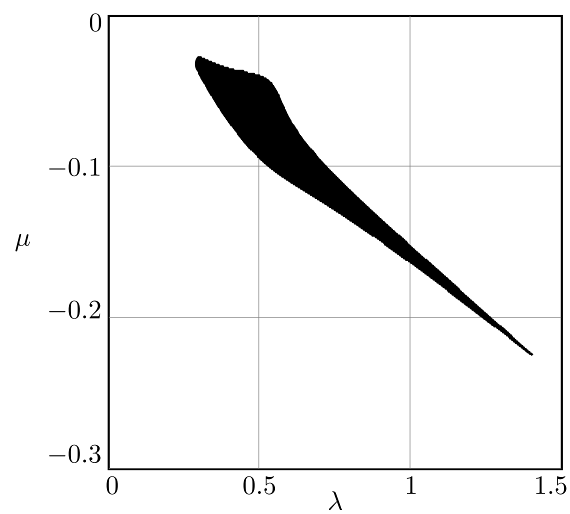


Figure 2: L-stable choices of (λ, μ) for $p = s = 3$ corresponding to $\mathcal{C}_1 = \mathcal{C}_2 = 10^{-4}$.

Here, we represent an example, characterized by

$$\lambda = \frac{1}{2}, \qquad \mu = -\frac{1}{15}, \qquad c = [\frac{1}{3} \quad \frac{2}{3} \quad 1]^T.$$

The coefficients of the method are

$$A = \begin{bmatrix} 0.5000000000000000 & 0 & 0 \\ 1.4279081052775164 & 0.5000000000000000 & 0 \\ 1.0000000000000000 & -0.3168631901664915 & 0.5000000000000000 \end{bmatrix},$$

$$\overline{A} = \begin{bmatrix} -0.0666666666666667 & 0 & 0 \\ -0.3067166674763493 & -0.0666666666666667 & 0 \\ -0.0602082721233515 & 0.0288951398441268 & -0.0666666666666667 \end{bmatrix},$$

$$B = \begin{bmatrix} 1.0000000000000000 & -0.3168631901664915 & 0.5000000000000000 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 84.1340111524194390 & -15.9895442199910120 & -37.9511333307057703 \end{bmatrix},$$

$$\overline{B} = \begin{bmatrix} -0.0602082721233515 & 0.0288951398441268 & -0.0666666666666667 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1.7866934603873189 & 20.0458159571414841 \end{bmatrix},$$

$$U = \begin{bmatrix} 1.0000000000000000 & -0.1666666666666667 & -0.0444444444444444 & 0.0006172839506173 \\ 1.0000000000000000 & -1.2612414386108497 & -0.2136971453939340 & 0.0056267104705261 \\ 1.0000000000000000 & -0.1831368098335086 & -0.0241114076097811 & -0.0010021824846360 \end{bmatrix},$$

$$V = \begin{bmatrix} 1.0000000000000000 & -0.1831368098335086 & -0.0241114076097811 & -0.0010021824846360 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -30.1933336017226565 & 2.3070365964725901 & 0 \end{bmatrix}.$$

2.3 Methods with $p = q = s = r - 1 = 4$

In this part, we construct methods of the Nordsieck SGLMs of type 2 with $p = q = s = r - 1 = 4$ and $c = [0 \;\; 0 \;\; 0 \;\; 1]^T$. Setting some free parameters in order to make calculation easier, RKS conditions make the coefficients matrices of the method to take the following forms

$$\left[\begin{array}{cccc|cccc|ccccc}
\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{12} & 0 & 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{12} & 0 & 0\\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & -\frac{1}{4} & -\frac{1}{12} & 0 & 0 & 1 & -1 & \frac{1}{3} & 0 & 0\\
\frac{1}{2} & 1 & \frac{1}{2} & 0 & -\frac{1}{4} & 1 & -\frac{1}{12} & 0 & 1 & -2 & -\frac{2}{3} & 0 & 0\\
\frac{1}{2} & -1 & 1 & \frac{1}{2} & -\frac{1}{4} & 1 & -1 & -\frac{1}{12} & 1 & 0 & \frac{1}{3} & 0 & 0\\
\hline
\frac{1}{2} & -1 & 1 & \frac{1}{2} & -\frac{1}{4} & 1 & -1 & -\frac{1}{12} & 1 & 0 & \frac{1}{3} & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\
6 & 0 & 0 & -6 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0\\
12 & 0 & 0 & -12 & 7 & -1 & 0 & 6 & 0 & 0 & 0 & 0 & 0
\end{array}\right].$$

For this method the only nonzero eigenvalue of $M(z)$ is

$$R(z) = \frac{1 + \frac{z}{2} + \frac{z^2}{12}}{1 - \frac{z}{2} + \frac{z^2}{12}},$$

consequently, by the Ehle conjecture [19], the method is A-stable. The error constant of the method is $\mathcal{C} = \frac{1}{720}$.

3 Implementation of the methods in a variable stepsize mode

In this section, we recall some implementation strategies given in [4], to apply the constructed methods in variable stepsize environment.

A Nordsieck SGLM in the variable stepsize mode takes the form

$$\begin{aligned}
Y^{[n]} &= h_n(A \otimes I_m)f(Y^{[n]}) + h_n^2(\overline{A} \otimes I_m)g(Y^{[n]}) + (UD(\delta_n) \otimes I_m)y^{[n-1]},\\
y^{[n]} &= h_n(B \otimes I_m)f(Y^{[n]}) + h_n^2(\overline{B} \otimes I_m)g(Y^{[n]}) + (VD(\delta_n) \otimes I_m)y^{[n-1]},
\end{aligned} \tag{3}$$

where

$$D(\delta_n) := \mathrm{diag}(1, \delta_n, \delta_n^2, ..., \delta_n^p), \quad \delta_n = h_n/h_{n-1}.$$

Due to the structure of the matrices V and $D(\delta_n)$, the matrix $VD(\delta_n)$ has one eigenvalue with magnitude one and a zero eigenvalue with multiplicity p, for any value of δ_n.

To achieve a suitable choice of stepsize for the next step, we first need to estimate the leading term in the local truncation error. To do this, we approximate $h^{p+1}y^{(p+1)}(x_n)$; so that $\mathrm{LTE}(x_n) = \mathcal{C}_p h^{p+1}y^{(p+1)}(x_n)$ can be calculated as an approximation to the local truncation error.

For the methods with $p = s = 2$ and abscissas $c = [\frac{1}{2} \;\; 1]^T$, we use the linear combination of the form

$$\text{est}(x_n) := \mathcal{C}_p\Big(\alpha_1 hf(Y_1) + \alpha_2 hf(Y_2) + \beta h^2 g(Y_1)\Big),$$

with

$$\alpha_1 = -8, \quad \alpha_2 = 8, \quad \beta = -4.$$

For the methods with $p = s = 3$ and abscissas $c = [\frac{1}{3} \;\; \frac{2}{3} \;\; 1]^T$, we use the linear combination of the form

$$\text{est}(x_n) := \mathcal{C}_p\Big(\alpha_1 hf(Y_1) + \alpha_2 hf(Y_2) + \alpha_3 hf(Y_3) + \beta h^2 g(Y_1)\Big),$$

with

$$\alpha_1 = \frac{243}{2}, \quad \alpha_2 = -162, \quad \alpha_3 = \frac{81}{2}, \quad \beta = 27.$$

For the methods with $p = s = 4$ and abscissas $c = [0 \;\; 0 \;\; 0 \;\; 1]^T$, we use the linear combination of the form

$$\text{est}(x_n) := \mathcal{C}_p\Big(\alpha_1 hf(Y_3) + \alpha_2 hf(Y_4) + \beta_1 h^2 g(Y_3) + \beta_2 h^2 g(Y_4) + \gamma y_4^{[n-1]}\Big),$$

with

$$\alpha_1 = 72, \quad \alpha_2 = -72, \quad \beta_1 = 48, \quad \beta_2 = 24, \quad \gamma = 12.$$

To control the stepsize, we use the following strategy

$$\text{est}(x_n) \le Rtol \cdot \max\{\|y_n\|, \|y_{n+1}\|\} + Atol, \tag{4}$$

where $Atol$ and $Rtol$ are given absolute and relative tolerances, respectively. If the control (4) is not satisfied, the current step is repeated with a halved stepsize. Otherwise, the current step is accepted and we carry out the next step with the new stepsize defined as

$$h_{n+1} = \delta_{n+1} h_n,$$

where

$$\delta_{n+1} = \min\Big\{facmax, \Big(\frac{fac \cdot tol}{\|\text{est}(x_n)\|_\infty}\Big)^{\frac{1}{p+1}}\Big\}.$$

Here, $facmax$ and fac are safety factors built into a code to prevent the step from increasing too rapidly and to avoid an excessive number of rejected steps. We have chosen $Atol = Rtol = tol$, $facmax = 2$, and $fac = 0.9$.

4 Numerical result

In this section we present the results of numerical experiments to show efficiency of the constructed methods in section 2 for fixed and variable stepsize mode using the provided techniques in section 3. We consider the following test problems:

- Problem 1. The nonlinear stiff system of ODEs

$$\begin{cases} y_1'(x) = -1002y_1(x) + 1000y_2^2(x), \\ y_2'(x) = y_1(x) - y_2(x)(1 + y_2(x)), \end{cases}$$

  whose exact solution is $[y_1(x),\ y_2(x)]^T = [\exp(-2x),\ \exp(-x)]^T$ and $x \in [0, 2]$. This problem is stiff with an approximate stiffness ratio of 10^3 near to $x = 0$.

- Problem 2. A system of differential equation, is called CUSP, resulting from discretization of the diffusion terms by the method of line the periodic boundary-value problem [16], is as below

$$\begin{cases} \frac{\partial y}{\partial t} = -\frac{1}{\varepsilon}(y^3 + ay + b) + \sigma \frac{\partial^2 y}{\partial x^2}, \\ \frac{\partial a}{\partial t} = b + 0.07\nu + \sigma \frac{\partial^2 a}{\partial x^2}, \\ \frac{\partial a}{\partial t} = (1 - a^2)b - a - 0.4y + 0.035\nu + \sigma \frac{\partial^2 b}{\partial x^2}, \end{cases}$$

  where

$$\nu = \frac{u}{0.1 + u}, \qquad u = (y - 0.7)(y - 1.3).$$

  This problem takes the form

$$\begin{cases} \dot{y}_i = -\varepsilon^{-1}(y_i^3 + a_i y_i + b_i) + D(y_{i-1} - 2y_i + y_{i+1}), \\ \dot{a}_i = b_i + 0.07\nu_i + D(a_{i-1} - 2a_i + a_{i+1}), \qquad i = 1, 2, \ldots, N, \\ \dot{b}_i = (1 - a_i^2)b_i - a_i - 0.4y_i + 0.035\nu_i + D(b_{i-1} - 2b_i + b_{i+1}), \end{cases}$$

  where

$$\nu_i = \frac{u_i}{0.1 + u_i}, \qquad u_i = (y_i - 0.7)(y_i - 1.3), \qquad D = \sigma N^2,$$

  with periodic boundary condition

$$\begin{array}{lll} y_0 := y_N, & a_0 := a_N, & b_0 := b_N, \\ y_{N+1} := y_1, & a_{N+1} := a_1, & b_{N+1} := b_1. \end{array}$$

  We take $\sigma = \frac{1}{144}, \epsilon = 10^{-4}, N = 32$, and the initial values as

$$y_i(0) = 0, \quad a_i(0) = -2\cos(\frac{2i\pi}{N}), \quad a_i(0) = 2\sin(\frac{2i\pi}{N}), \quad i = 1, 2, \ldots, N,$$

  with $t_{out} = 1.1$.

- Problem 3. For the ODE case, the Ring modulator problem originates from electrical circuit analysis is of the form [18]

$$
\begin{cases}
y_1' = C^{-1}(y_8 - 0.5y_{10} + 0.5y_{11} + y_{14} - R^{-1}y_1),\\
y_2' = C^{-1}(y_9 - 0.5y_{12} + 0.5y_{13} + y_{15} - R^{-1}y_2),\\
y_3' = C_s^{-1}\Big(y_{10} - q(U_{D1}) + q(U_{D4})\Big),\\
y_4' = C_s^{-1}\Big(-y_{11} + q(U_{D2}) - q(U_{D3})\Big),\\
y_5' = C_s^{-1}\Big(y_{12} + q(U_{D1}) - q(U_{D3})\Big),\\
y_6' = C_s^{-1}\Big(-y_{13} - q(U_{D2}) + q(U_{D4})\Big),\\
y_7' = C_p^{-1}\Big(R_p^{-1}y_7 + q(U_{D1}) + q(U_{D2}) - q(U_{D3}) - q(U_{D4})\Big),\\
y_8' = -L_h^{-1}y_1,\\
y_9' = -L_h^{-1}y_2,\\
y_{10}' = L_{s2}^{-1}(0.5y_1 - y_3 - R_{g2}y_{10}),\\
y_{11}' = L_{s2}^{-1}(-0.5y_1 + y_4 - R_{g3}y_{11}),\\
y_{12}' = L_{s3}^{-1}(0.5y_2 - y_5 - R_{g2}y_{12}),\\
y_{13}' = L_{s3}^{-1}(-0.5y_2 + y_6 - R_{g3}y_{13}),\\
y_{14}' = L_{s1}^{-1}\Big(-y_1 + U_{in1}(t) - (R_i + R_{g1})y_{14}\Big),\\
y_{15}' = L_{s1}^{-1}\Big(-y_2 - (R_c + R_{g1})y_{15}\Big).
\end{cases}
$$

The auxiliary functions $U_{D1}, U_{D2}, U_{D3}, U_{D4}, q, U_{in1}$, and U_{in2} are given by

$$
\begin{cases}
U_{D1} = y_3 - y_5 - y_7 - U_{in2}(t),\\
U_{D2} = -y_4 + y_6 - y_7 - U_{in2}(t),\\
U_{D3} = y_4 + y_5 + y_7 + U_{in2}(t),\\
U_{D4} = -y_3 - y_6 + y_7 + U_{in2}(t),\\
q(U) = \gamma(e^{\delta U} - 1),\\
U_{in1}(t) = 0.5\sin(2000\pi t),\\
U_{in2}(t) = 2\sin(20000\pi t).
\end{cases}
$$

The values of the parameters are

$$
\begin{aligned}
&C = 1.6\times 10^{-8}, \quad C_s = 2\times 10^{-12}, \quad C_p = 10^{-12}, \quad L_h = 4.45,\\
&L_{s1} = 0.002, \quad L_{s2} = 5\times 10^{-4}, \quad L_{s3} = 5\times 10^{-4},\\
&R = 25000, \quad R_p = 50, \quad R_{g1} = 36.3, \quad R_{g2} = 17.3,\\
&R_{g3} = 50, \quad R_i = 50, \quad R_c = 600,\\
&\delta = 17.74933332, \ \gamma = 40.67286402\times 10^{-9},
\end{aligned}
$$

with the initial vector $y_0 = (0,0,0,0,0,0,0,0,0,0,0,0,0,0)^T$ and $t \in [0, 10^{-3}]$.

4.1 Fixed stepsize experiments

We first present fixed stepsize numerical results in order to show accuracy of constructed Nordsieck SGLMs and validate the order of these methods in the integration of stiff differential systems. To do this, we have applied the methods of order 2, 3, and 4 to the Problem 1. We have implemented the methods with a fixed stepsize $h = 1/2^k$, with several integer values of k. The results of numerical experiments for type 2 Nordsieck SGLMs are shown in the Tables 1, 2, and 3. In these tables, we have listed the norm of error $\|e_h(\overline{x})\|$ at the endpoint of integration $\overline{x} = 2$ and numerical estimate to the order of convergence, p, computed by the formula

$$p = \frac{\log(\|e_h(\overline{x})\|/\|e_{h/2}(\overline{x})\|)}{\log(2)},$$

where $e_h(\overline{x})$ and $e_{h/2}(\overline{x})$ are errors corresponding to stepsizes h and $h/2$ for Nordsieck SGLMs. Also, to show that the constructed methods are competitive with the efficient existing methods, we have reported the numerical results of type 2 methods which have been constructed in [2].

Table 1: Numerical results for Nordsieck SGLMs of order $p = q = 2$.

k	Type 2 method of order 2 [2]		Nordsieck SGLM of order 2	
	$e_h(\overline{x})$	p	$e_h(\overline{x})$	p
10	1.78×10^{-11}		1.35×10^{-10}	
11	5.44×10^{-12}	1.71	3.31×10^{-11}	1.88
12	1.43×10^{-12}	1.93	8.18×10^{-12}	1.94
13	2.64×10^{-13}	2.44	2.03×10^{-12}	1.98

Table 2: Numerical results for Nordsieck SGLMs of order $p = q = 3$.

k	Type 2 method of order 3 [2]		Nordsieck SGLM of order 3	
	$e_h(\overline{x})$	p	$e_h(\overline{x})$	p
4	4.07×10^{-6}		6.58×10^{-8}	
5	5.66×10^{-7}	2.85	8.66×10^{-9}	2.93
6	7.43×10^{-8}	2.92	1.11×10^{-9}	2.96
7	9.50×10^{-9}	2.96	1.40×10^{-10}	2.99

Table 3: Numerical results for Nordsieck SGLMs of order $p = q = 4$.

k	Type 2 method of order 4 [2]		Nordsieck SGLM of order 4	
	$e_h(\overline{x})$	p	$e_h(\overline{x})$	p
4	6.22×10^{-8}		5.81×10^{-9}	
5	4.21×10^{-9}	3.88	3.63×10^{-10}	4.00
6	2.74×10^{-10}	3.94	2.27×10^{-11}	4.00
7	1.75×10^{-11}	3.97	1.42×10^{-12}	4.00

4.2 Variable stepsize experiments

We first investigate the potential for efficient implementation in a variable stepsize environment by the reliability of the estimation error for Problem 1. The obtained results for the methods of order 2, 3, and 4 have plotted in Figures 3, 4, and 5, respectively. These figures confirm efficiency of the used estimation for the local truncation error. To compare, we also present the results of numerical experiments of the L-stable Nordsieck SGLM given in [5]. In the implementation of considered SGLMs, we apply the same introduced implementation strategies, including the starting procedures, stage predictors, local error estimation, and the changing stepsize. In our numerical results, we use the following abbreviations:

ns:	the number of steps
nrs:	the number of rejected steps
nfe:	the number of function evaluations
nJe:	the number of Jacobian evaluations
ge:	the global error
tol:	given tolerance
NSGLM2p:	type 2 Nordsieck SGLM of order p
NSGLM4p:	type 4 Nordsieck SGLM of order p

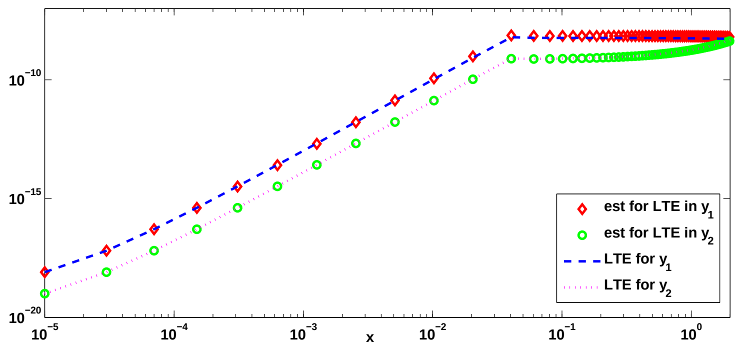


Figure 3: Local errors and local error estimates versus x of the method of order 2 for problem 1 with $h_0 = 10^{-5}$ and $tol = 10^{-8}$.

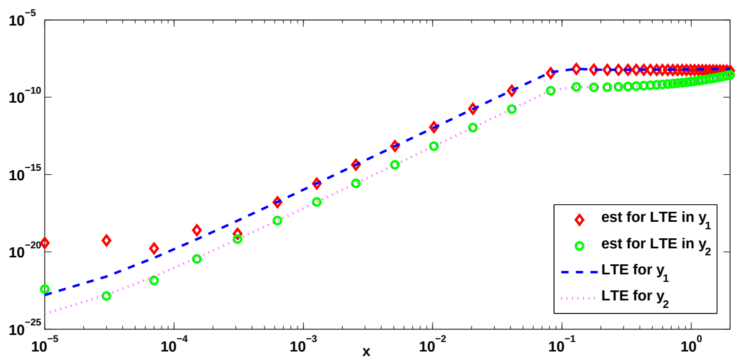


Figure 4: Local errors and local error estimates versus x of the method of order 3 for problem 1 with $h_0 = 10^{-5}$ and $tol = 10^{-8}$.

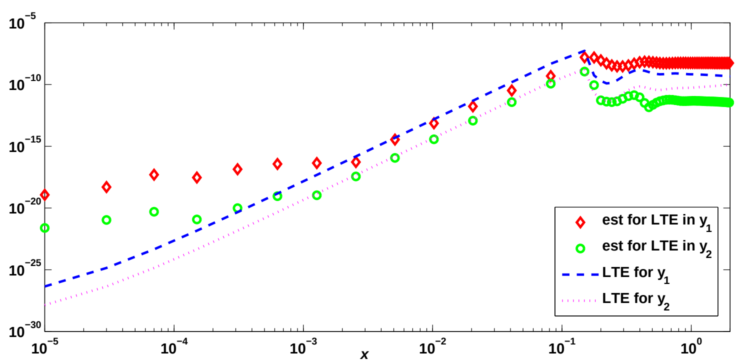


Figure 5: Local errors and local error estimates versus x of the method of order 4 for problem 1 with $h_0 = 10^{-5}$ and $tol = 10^{-8}$.

Table 4: Numerical results for Problem 2 solved by NSGLM2p and NSGLM4p of orders $2, 3$, and 4 with $h_0 = 10^{-3}$.

tol	Method	*ge*	*ns*	*nrs*	*nfe*	*nJe*
10^{-6}	NSGLM22	3.61×10^{-5}	169	26	1644	1256
	NSGLM42	2.02×10^{-4}	1313	11	6693	4047
10^{-8}	NSGLM22	1.07×10^{-6}	690	12	3718	2316
	NSGLM42	8.84×10^{-6}	5997	14	24187	12167
10^{-10}	NSGLM22	4.58×10^{-8}	3154	15	12847	6511
	NSGLM42	3.99×10^{-7}	27572	14	110499	55323
10^{-6}	NSGLM23	2.95×10^{-5}	178	36	2339	1700
	NSGLM43	5.80×10^{-5}	426	17	3248	1922
10^{-8}	NSGLM23	6.57×10^{-7}	474	14	4047	2586
	NSGLM43	1.28×10^{-6}	1341	23	8308	4219
10^{-10}	NSGLM23	1.47×10^{-8}	1450	18	8998	4597
	NSGLM43	5.59×10^{-9}	26879	646	165153	82581
10^{-6}	NSGLM24	9.16×10^{-5}	337	25	3308	1864
	NSGLM44	9.66×10^{-5}	313	14	3420	2116
10^{-8}	NSGLM24	2.14×10^{-6}	2040	50	16864	8508
	NSGLM44	2.19×10^{-6}	787	14	6599	2399
10^{-10}	NSGLM24	3.77×10^{-8}	2696	141	22752	11408
	NSGLM44	4.59×10^{-8}	1700	18	13934	7066

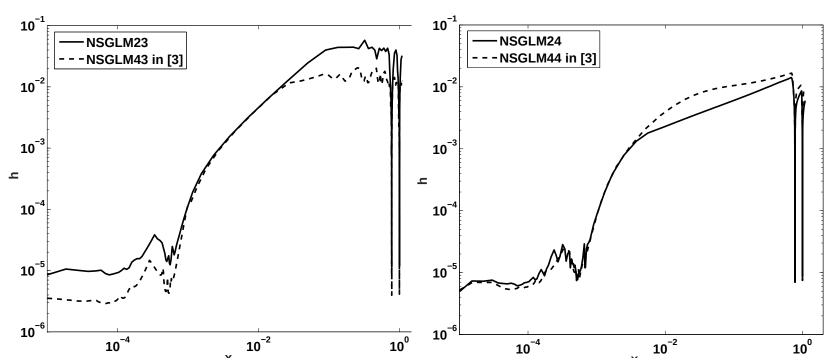
Some numerical results for Problem 2 and 3 demonstrating the computational cost of NSGLM2p are given in Tables 4 and 5 and compared with those in NSGLM4p for $p = 2, 3, 4$. Also, in Figures 6 and 7, we compare the accepted stepsizes of NSGLM23 with those in NSGLM24 and the accepted stepsizes of NSGLM43 with those in NSGLM44 through integration for Problem 2 and 3, recpectively. The numerical results show that the proposed methods are capable in solving stiff problems and competitive with the existing methods.

5 Conclusion

We constructed type 2 Nordsieck SGLMs of orders 2, 3, and 4 with RKS property. Order 2 and 3 methods are L-stable and order 4 method is A-stable. These methods have been equipped to the variable stepsize using Nordsieck technique. The capability of the proposed methods in solving stiff problems with long interval of integration and badly scaled solution have been validated by some numerical experiments and comparisons.

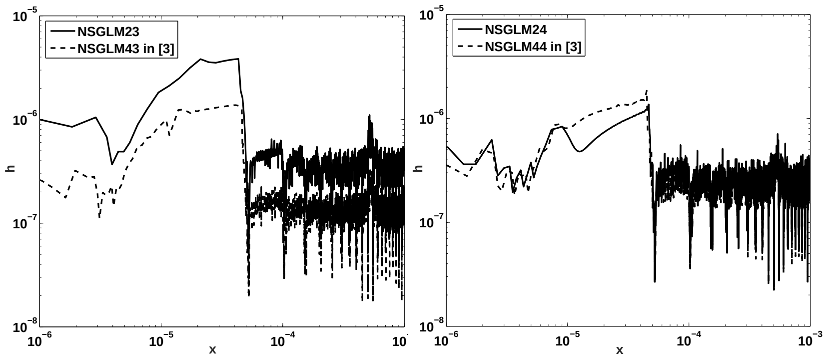
Table 5: Numerical results for Problem 3 solved by NSGLM2p and NSGLM4p of orders $2, 3$, and 4 with $h_0 = 10^{-6}$.

tol	Method	ge	ns	nrs	nfe	nJe
10^{-6}	NSGLM22	1.18×10^{-5}	3336	492	24233	16579
	NSGLM42	8.95×10^{-5}	27540	475	139942	83914
10^{-8}	NSGLM22	5.94×10^{-7}	14721	486	77469	47057
	NSGLM42	3.94×10^{-6}	126749	427	508703	254353
10^{-10}	NSGLM22	2.05×10^{-8}	67411	469	271519	135761
	NSGLM42	8.61×10^{-8}	272719	396	1092456	546228
10^{-6}	NSGLM23	1.71×10^{-5}	2790	476	30994	21199
	NSGLM43	1.68×10^{-5}	8243	499	61705	35482
10^{-8}	NSGLM23	1.46×10^{-7}	8729	705	76184	74885
	NSGLM43	5.31×10^{-7}	14444	450	155761	77884
10^{-10}	NSGLM23	5.59×10^{-9}	26879	646	165153	82581
	NSGLM43	1.61×10^{-8}	80230	414	483861	241932
10^{-6}	NSGLM24	1.75×10^{-3}	4070	642	42363	23519
	NSGLM44	2.11×10^{-3}	4563	591	51425	30813
10^{-8}	NSGLM24	4.86×10^{-5}	13440	1198	117655	59107
	NSGLM44	5.47×10^{-5}	10954	612	92614	46356
10^{-10}	NSGLM24	8.86×10^{-7}	24087	5512	236786	118394
	NSGLM44	1.35×10^{-6}	26896	602	219980	109992



(a) (b)

Figure 6: Accepted stepsizes versus x for Problem 2 with $h_0 = 10^{-3}$ and $tol = 10^{-6}$: (a) NSGLM23 and NSGLM43, (b) NSGLM24 and NSGLM44.



(a) (b)

Figure 7: Accepted stepsizes versus x for Problem 3 with $h_0 = 10^{-6}$ and $tol = 10^{-6}$: (a) NSGLM23 and NSGLM43, (b) NSGLM24 and NSGLM44.

References

1. Abdi, A. *Construction of high-order quadratically stable second-derivative general linear methods for the numerical integration of stiff ODEs*, J. Comput. Appl. Math. 303 (2016) 218–228.

2. Abdi, A., Braś, M., and Hojjati, G. *On the construction of second derivative diagonally implicit multistage integration methods*, Appl. Numer. Math. 76 (2014), 1–18.

3. Abdi, A. and Hojjati, G. *High order second derivative methods with Runge–Kutta stability for the numerical solution of stiff ODEs*, Iran. J. Numer. Anal. Optim. 5 (2015), 1–10.

4. Abdi, A. and Hojjati, G. *Implementation of Nordsieck second derivative methods for stiff ODEs*, Appl. Numer. Math. 94 (2015), 241–253.

5. Abdi, A. and Hojjati, G. *An extension of general linear methods*, Numer. Algor. 57 (2011), no. 2, 149–167.

6. Abdi, A. and Hojjati, G. *Maximal order for second derivative general linear methods with Runge–Kutta stability*, Appl. Numer. Math. 61 (2011), no. 10, 1046–1058.

7. Barghi Oskouie, N., Hojjati, G., and Abdi. A. *Efficient second derivative methods with extended stability regions for non-stiff IVPs.* Comp. Appl. Math. DOI: 10.1007/s40314-017-0436-y.

8. Butcher, J. C. *Numerical Methods for Ordinary Differential Equations*, Wiley, New York, 2008.

9. Butcher, J. C. *A transformation for the analysis of DIMSIMs*. BIT 34 (1994), no. 1, 25–32.

10. Butcher, J. C. and Chartier, P. *Parallel general linear methods for stiff ordinary differential and differential algebraic equations*, Appl. Numer. Math. 17 (1995), 213–222.

11. Butcher, J. C. and Hojjati, G. *Second derivative methods with RK stability*, Numer. Algor. 40 (2005), no. 4, 415–429.

12. Butcher, J. C. and Jackiewicz, Z. *Construction of high order diagonally implicit multistage integration methods for ordinary differential equations*. Appl. Numer. Math., 27 (1998), no. 1, 1–12.

13. Butcher, J. C. and Jackiewicz, Z. *Implementation of diagonally implicit multistage integration methods for ordinary differential equations*, SIAM J. Numer. Anal. 34 (1997), no. 6, 2119–2141.

14. Butcher, J. C. and Jackiewicz, Z. *Diagonally implicit general linear methods for ordinary differential equations*, BIT 33 (1993), 452–472.

15. Butcher, J. C. and Singh, A. D. *The choice of parameters in parallel general linear methods for stiff problems*, Appl. Numer. Math. 34 (2000), no. 1, 59–84.

16. Hairer, E. and Wanner, G. *Solving Ordinary Differential Equations II: Stiff and Differential–Algebraic Problems*, Springer, Berlin, 1996.

17. Jackiewicz, Z. *General linear methods for ordinary differential equations*, Wiley, New Jersey, 2009.

18. Mazzia, F. and Magherini, C. http://www.dm.uniba.it/ testset.

19. Wanner, G., Hairer, E., and Nørsett, S. P. *Order stars and stability theorems*, BIT, 18 (1978), no. 4, 475–489.

Iranian Journal of Numerical Analysis and Optimization
Vol. 8, No. 1, (2018), pp 129–141
DOI:10.22067/ijnao.v8i1.40776

A limited resource assignment problem with shortage in the fire department

E.S. Alavi and R. Ghanbari*

Abstract

Assigning available resources to fire stations is a main task of fire department's administrator in a city. The importance of this problem increases when the number of available resources are inadequate. In this situation, the goal is to assign the limited available resources to fire stations such that the associated penalties of the shortages are minimized. Here, we first give a mathematical approach to consider some penalties for the shortage. Next, we give an integer program to minimize the sum of associated penalties. The proposed model can be used in many other problems arisen from health services, emergency management, and so on. We also propose a heuristic to efficiently solve the problem in a reasonable time. Our proposed heuristic has two phases. In the first phase, using a greedy approach, our proposed heuristic constructs a proper feasible solution. Next, in the second phase, we propose a local search to improve the quality of the solution constructed in the first phase. To show the efficiency of our proposed heuristic, we compare our proposed heuristic with CPLEX based on the running time and the quality of obtained solutions on two groups of problems (real-word problems and randomly generated problems). The numerical results show that on the 80% of benchmark problems, the obtained solution is the same as CPLEX's solutions. Also, the running time of our proposed algorithm is almost 10 times better than CPLEX's running time, in average.

Keywords: Resource assignment problem; Fire stations; Shortage; Integer programming; Heuristics.

---

*Corresponding author
Received 22 January 2017; revised 29 March 2017; accepted 23 April 2017
E.S. Alavi
Department of Applied Mathematics, Faculty of Mathematical science, Ferdowsi University of Mashhad, Mashhad, Iran. e-mail: effatsadat.alavi@gmail.com

R. Ghanbari
Department of Applied Mathematics, Faculty of Mathematical science, Ferdowsi University of Mashhad, Mashhad, Iran. e-mail: rghanbari@um.ac.ir

1 Introduction

There are many decision making problems for a fire department's administrator or even for a fire district's administrator (see [5,9,13]; also, [1,3,7,11,12]). One of these decision problems is to assign available resources (e.g., equipments or fire engines etc.) to fire stations (see [12], page 524, to find some references). In the real world situations, it is possible that there are not sufficient resources to satisfy all the requirements of each fire station. Thus, some of fire stations may lack some resources to extinguish or to service the fire accidents or events in their regions (other nearly problems in emergency management are studied by researchers, e.g., [2, 6, 8]). To decrease the influences of the shortage for a type of resource, some penalties are considered, based on its importance in each fire station. In a fire station, to compute the importance of a required resource, we estimate the possibility of the occurring of each event in the fire station's region. Next, based on the predefined impact of the events and the estimated possibility for arising of the events, the corresponding penalties are considered. In addition, when fire incidents occur or rescue services are needed, the responsible fire station usually requests additional resources or even firefighters from its neighboring fire stations (neighboring services). However, the aim of the problem is to assign the available resources to fire stations such that the sum of the corresponding penalties is minimized.

Here, we give an integer linear formulation to model the problem. We also propose a heuristic to efficiently solve the problem. We finally compare the performance of our algorithms with ILOG OPL(CPLEX solver)'s result on some randomly generated test problem and some real world problems.

The remainder of the paper is organized as follows: In section 2, we give an integer program to model the problem. The proposed algorithm is presented in section 3. In section 4, we show the efficiency of our proposed algorithm in comparison with ILOG OPL on some randomly generated test problems and some real world problems. We conclude in section 5.

2 Mathematical model

We here consider all resources are independent. It is obvious that when two or more types of resources are dependent, one can consider them as a package. Therefore, we can assume the shortage exists for a resource. It is obvious that the model and related solution approach can be used as a procedure within a successive approach for solving the problem with shortage in two or more types of resources. As mentioned in section 1, before providing the integer program, we must give some necessary estimation or computation to determine the corresponding penalty of the recourse's shortage in a fire station. Table 1 presents some parameters or given data of the problem.

Table 1: Parameters or given data

T:	the availability of the resource.
q:	the number of events.
n:	the number of fire stations.
f_k:	the required number of the resource to service the event k.
p_{kj}:	the frequency of the event k in the region of station j.
pr_{kj}:	the possibility for occurring the event k in the region of station j.
w_k:	the importance (weight) of the event k.
$N(j)$:	the set of the fire station j's neighboring fire stations.
β_j:	the importance (weight) of the fire station j.

Note 1. In Table 1, $pr_{.,.}$, $w_.$, and $\beta_.$ are in $[0, 1]$.

Note that p_{kj} is obtained previously. In the other words, we can consider the frequency of the event k in region of the fire station j for example for a period of ten years. But, pr_{kj} is given by managers, and it shows the possibility of occurring the event k in the region of fire station j. To more clarify, we give an example. Consider, a school is recently founded in the region of the fire station j, and suppose there was not any school in this region. Therefore, some events have rarely occurred there. But, the possibility of rare occurring or never occurring events is now raised. Thus pr_{kj} and p_{kj} can be completely different. When all the situations are the same, β_j imposes the manager's decision to assign the resource to the fire station j. As an example, consider there are two fire stations with the same conditions, need a type of equipment. Also, consider one of them is in the center of the city, and the other is on the outskirts. Suppose that there is one equipment. In this situation, it is usually preferred to assign this equipment to the fire station that is in the center of the city. Therefore, the manager can consider the higher weight to the fire station that is in the center of the city to impose her/his preferences in the model.

To compute the associated penalty, we must do some necessary computations. We consider the relative frequency for the event k in the region of the fire station j as follows:

$$u_{kj} = \frac{p_{kj}}{\sum_{j=1}^{n} p_{kj}}. \tag{1}$$

Remark 1. Note that one can consider u_{kj} introduced in (1) as an estimation of the probability of occurring the event k in the region of the fire station j.

Definition 1. Let

$$v_{kj} = \max\{u_{kj}, pr_{kj}\}, \tag{2}$$

where pr_{kj} is given by the manager and u_{kj} is introduced in (1). Now, we define $v_{kj}w_k$ as the impact of the event k on the station j, for $j = 1, \ldots, n$ and $k = 1, \ldots, q$.

Note that we could not find any definition for v_{kj} in the literature since this model was proposed for a real problem that occurred in Mashhad. We did not have no more information in this real-world situation. Therefore, using experts in Fire Department of Mashhad, we proposed this formula. In this formula, we consider historical data p_{kj} and experts consideration pr_{kj}. Thus, $w_k v_{kj}$ shows how important the event k in the region of the station j.

Definition 2. We denote the possible events in the station j by Q_j defined as follows:

$$Q_j = \{k | f_k > 0, v_{kj} > 0\}. \tag{3}$$

where v_{kj} is defined in (2) and f_k is defined in Table 1.
We now introduce the decision variables in Table 2.

Table 2: Decision variables

x_j:	shows the number of the resources that are assigned to station j.
y_{jk}:	shows the shortage in the station j to service the event k.
h_{jk}:	shows the shortage in the station j to service the event k with considering the neighboring services.
d_{jk}:	if station j can not service the event k, takes 1, otherwise takes 0.

2.1 The objective function

The objective of the problem is to minimize the sum of the associated penalties. We here consider three penalties for servicing the event k in the station j based on the lack of ability for providing responsible service, the number of shortages and the number of shortages with considering neighboring services. Therefore, we consider the following penalty for the station j:

$$\sum_{k \in Q_j} (d_{jk} + y_{jk} + h_{jk})\, w_k v_{kj}, \tag{4}$$

where Q_j and v_{kj} are defined in (3) and (2), respectively, d_{jk}, y_{jk}, and h_{jk} are defined in Table 2, and w_k is defined in Table 1. To explain (4), we give two illustrative examples.

Illustrative Example 1. Consider there are five fire stations and only one resource exists. Also, the structure of the neighboring fire stations is given in Figure 1. Suppose that each of the stations 1, 3, 4, and 5 need one unit

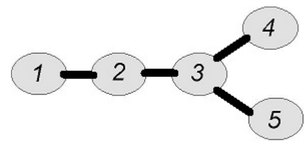


Figure 1: the graph of the neighborhoods in the illustrative Example 1.

of the resource and all situation are in the same. From Figure 1, the station 3 is the best candidate, because, the station 3 while servicing the occurred events in its own region, can participate to service the occurred events in the stations 4 and 5's regions. One can easily obtain the same result by using the sum of the penalties defined in (4) for each fire station.

Illustrative Example 2. Consider there are two fire stations and three possible events. Also, each fire station is without a neighbor (i.e. $N(1) = N(2) = \emptyset$). The other data are given as follows (see tables 1–2 and equations (2) and (3)):

$$T = 8, \quad Q_1 = \{2,3\}, \quad Q_2 = \{1,2,3\},$$
$$f = [4,\ 6,\ 2], \quad w = \frac{1}{0.7}[0.2,\ 0.3,\ 0.7],$$
$$v = \begin{bmatrix} 0.2 & 0.9 & 0.7 \\ 0.4 & 0.9 & 0.2 \end{bmatrix}$$

Now, we give the motivation for considering d_{kj} in our proposed model. If d_{kj} is removed from (4), we have the following penalty for the station j:

$$\sum_{k \in Q_j} (y_{jk} + h_{jk})\, w_k v_{kj}. \tag{5}$$

Since $N(1) = N(2) = \emptyset$, (5) can be considered as follows:

$$\sum_{k \in Q_j} (y_{jk})\, w_k v_{kj}. \tag{6}$$

Now, consider two feasible solutions, **(a):** $x_1 = x_2 = 4$, **(b):** $x_1 = 2$ and $x_2 = 6$. For **(a)**, we have

$$y_{12} = 2,\ y_{13} = 0,\ y_{21} = 0,\ y_{22} = 2,\ y_{23} = 0,$$

and based on (6), the sum of penalties is $\dfrac{2 \times 0.9 \times 0.3 + 2 \times 0.9 \times 0.3}{0.7} = \dfrac{1.08}{0.7}$. For **(b)**, we have

$$y_{12} = 4,\ y_{13} = 0,\ y_{21} = 0,\ y_{22} = 0,\ y_{23} = 0,$$

and based on (6), the sum of penalties is $\dfrac{4 \times 0.9 \times 0.3}{0.7} = \dfrac{1.08}{0.7}$. Thus, when we do not consider d_{kj} in the model, there is no difference between **(a)** and **(b)**. But, in **(b)**, only the event 2 in the station 1 is not covered. Whereas in **(a)**, the event 2 in stations 1 and 2 is not covered. Now, if we consider (4) as the associated penalty, then **(b)** is preferred to **(a)**.

2.2 The model

Now, the proposed integer program is given as follows.

$$\min \; z = \sum_{j=1}^{n} \beta_j \left\{ \sum_{k \in Q_j} (d_{jk} + y_{jk} + h_{jk})\, w_k v_{kj} \right\} \tag{7}$$

s.t.

$$\sum_{j=1}^{n} x_j \leq T. \tag{8}$$

$$y_{jk} \geq f_k - x_j \qquad \forall j \in \{1, \ldots, n\}\; k \in Q_j. \tag{9}$$

$$h_{jk} \geq f_k - x_j - \sum_{\bar{j} \in N(j)} x_{\bar{j}} \qquad \forall j \in \{1, \ldots, n\}\; k \in Q_j. \tag{10}$$

$$d_{jk} \leq y_{jk} \qquad \forall j \in \{1, \ldots, n\}\; k \in Q_j. \tag{11}$$

$$d_{jk} \geq \frac{y_{jk}}{M_f}, \qquad \forall j \in \{1, \ldots, n\}\; k \in Q_j, \text{where } M_f = \max_{k=1,\ldots,q} f_k. \tag{12}$$

$$y_{jk},\; h_{jk} \in \mathbb{N} \cup \{0\}, \text{and } d_{jk} \in \{0,1\} \qquad \forall j \in \{1, \ldots, n\}\; k \in Q_j.$$

$$x_j \in \mathbb{N} \cup \{0\} \qquad \forall j \in \{1, \ldots, n\}.$$

The constraint (8) shows that the number of assigned resources must be less than or equal to T. From Table 2, y_{kj} shows the shortage for servicing the event k in the station j. Thus, if we assign x_j unit of the resource to station j, then y_{kj} is $\max\{0, f_k - x_j\}$. Based on constraints (9), if $x_j \geq f_k$, then the corresponding constraint is redundant. Since the objective function is minimization, then y_{kj} takes zero. In the other case, if $x_j < f_k$, then the corresponding constraint ensures that $y_j \geq f_k - x_j > 0$. Similar to above, since the objective function is minimization, then y_{kj} takes $f_k - x_j$. Therefore, constraints (9) ensure that y_{kj} is $\max\{0, f_k - x_j\}$. Similar to constraints (9), constraints (10) ensure that $h_{jk} = \max\{0, f_k - x_j - \sum_{\bar{j} \in N(j)} x_{\bar{j}}\}$. Together constraints (11) and (12) guarantee, if y_{kj} is zero, then d_{kj} is zero and if $y_{kj} > 0$, then d_{kj} is 1.

Algorithm 1 Greedy phase

Step 1 {Initialization} Let $x^* = (0, \dots, 0)^T_{n\times 1}$, and let $z^* = z(x^*)$.

Step 2 {Main loop}
For $t = 1$ to T do:

- Let $z_* = +\infty$.
  For $j = 1$ to n do:
  - Let $\bar{x} = x^*$, $\bar{x_j} = \bar{x_j} + 1$ and $\bar{z} = z(\bar{x})$.
  - **If** $\bar{z} < z_*$, **then** let $z_* = \bar{z}$ and $\bar{j} = j$.

  endfor
- Let $x^*_{\bar{j}} = x^*_{\bar{j}} + 1$, and let $z^* = z_*$. **If** $z^* = 0$, then stop, the optimal solution is x^*.

endfor

Step 3 Return x^* as an initial solution.

3 A heuristic algorithm

In this section, we propose a heuristic algorithm to solve the problem. Our proposed algorithm has two phases. In the first phase, we propose a greedy procedure to construct a feasible solution (the greedy phase). Then, in the second phase, we apply a neighborhood search on the constructed solution (the improvement phase).

3.1 Greedy phase

In this phase, we construct a feasible solution. At first, we let $x_1 = \cdots = x_n = 0$. Then, in each iteration, we compute the corresponding difference of the objective function for each fire station, when the number of the assigned resources in a fire station is increased one unit. Then, we select the station j to increase x_j, when the corresponding difference of the objective function is the best.
To evaluate $x = (x_1, \dots, x_n)^T$, let $y_{kj} = \max\{0, f_k - x_j\}$, and let $h_{jk} = \max\{0, f_k - x_j - \sum_{\bar{j}\in N(j)} x_{\bar{j}}\}$. Also, if y_{kj} is zero, then let $d_{kj} = 0$ and otherwise let $d_{kj} = 1$. Now, the corresponding objective value of x can be computed by (7). We denote the corresponding objective function value of x by $z(x)$. The greedy phase is given in Algorithm 1.

Note 2. In Algorithm 1, if we can find a solution x^* with $z(x^*) = 0$, then we have found a solution without any penalty. Therefore, we terminate the algorithm, and we do not use the phase 2.

3.2 Improvement phase

Here, we describe the neighborhood search procedure. We use a simple local search to improve the quality of the solution constructed by Algorithm 1. We first give some definitions.

Definition 3. We denote the maximum number of requirements in the station j by b_j and define it as follows:

$$b_j = \max_{k \in Q_j} \{f_k\}, \qquad j = 1, \ldots, n, \tag{13}$$

where Q_j is defined by (3).

It is evident that if $x_j \geq b_j$, then we do not have any shortage in the station j.

Definition 4. Let

$$\Delta_j = \begin{cases} \min\{f_k - x_j | k \in Q_j, f_k > x_j\}, & x_j < b_j, \\ 0, & x_j \geq b_j, \end{cases} \tag{14}$$

where Q_j is defined by (3). Δ_j is the minimum number of the resources that if assigned to station j, an uncovered event (event that does not receive a response service) will be covered.

To explain our proposed local search, we need to define a move in the search space (see [10]). For the station j with $x_j < b_j$, we define

$$M_j = \left\{\bar{j} | x_{\bar{j}} \geq \Delta_j, \bar{j} \neq j\right\}, \tag{15}$$

where Δ_j is defined by (14). Note that, M_j is the set of stations that their current availability is more than the minimum requirement of the station j. Now, for the station j with $x_j < b_j$, a feasible solution can be obtained from the current feasible solution by $x_j = x_j + \Delta_j$ and $x_{\bar{j}} = x_{\bar{j}} - \Delta_j$, where $\bar{j} \in M_j$. We denote this move by $Mov(j, \bar{j}, \Delta_j)$.

Note 3. When a move is done, an effective procedure to compute the difference of the objective function value, can decrease the computational time. Consider x is a feasible solution and $\bar{x}$ is the obtained feasible solution by $Mov(j, \bar{j}, \Delta_j)$. Also, consider $z(x)$ and $z(\bar{x})$ are the objective function values of x and $\bar{x}$, respectively. We now define

Algorithm 2 Local search (the improvement phase).

Step 1 Let x^* is the solution constructed by Algorithm 1, and let $z^* = z(x^*)$.

Step 2 Let $S = \{j | x^*_j < b_j\}$.

Step 3 **For** $j = 1, \ldots, |S|$ do:

3-1 Compute Δ_j by (14).

3-2 **For** $\bar{j} \in M_j$, defined by (15), **if** $\Delta^z_{j,\bar{j},\Delta_j} < 0$, defined by (16), then $Mov(j, \bar{j}, \Delta_j)$, update x^* and z^*, and **goto Step 2**.

Step 4 **Return** x^* as an approximate solution.

$$\Delta^z_{j,\bar{j},\Delta_j} = z(\bar{x}) - z(x), \tag{16}$$

where $z(.)$ is defined by (4). It is obvious that $\Delta^z_{j,\bar{j},\Delta_j}$ can be computed by subtracting the sum of increased penalties of station $\bar{j}$ and its neighbors from the sum of the decreased penalties of station j and its neighboring stations. Therefore, the other stations don not play any role in the computation of $\Delta^z_{j,\bar{j},\Delta_j}$; thus in the implementation, the value of $\Delta^z_{j,\bar{j},\Delta_j}$ can be computed effectively. Also, One can use an approximation of $\Delta^z_{j,\bar{j},\Delta_j}$ in the implementation (similar to [4]).

We now give our proposed local search (as the improvement phase) in Algorithm 2.

Proposition 1. *Our proposed heuristic gives a feasible solution.*

Proof. In the greedy phase (Algorithm 1), our proposed heuristic begins from zero ($x = 0$), and then in each iteration of Step 2, only one unit is added to one entry of x. Thus, the return solution is a non-negative integer vector. The second phase starts by the feasible solution given by phase 1. In each iteration of Step 3 in Algorithm 2, each move $Mov(j, \bar{j}, \Delta_j)$ is done when $x_{\bar{j}} \geq \Delta_j$ (see (15)) and $\Delta_j > 0$ defined in (14) is an integer number. Thus, after each move the solution remains feasible. □

Note 4. There are some problems such that our proposed heuristic does not give the optimal solution (e.g., MASH2 in table 3). Also, we are not able to prove that our proposed heuristic is an ε-approximation algorithm [10]. However, the numerical results show that our proposed heuristic gives very good solution in an acceptable time for real-world problems (see Table 3).

4 Numerical experiments

To show the efficiency of our proposed algorithm, we compared the running time and the quality of the arrived solution of the proposed algorithm with ILOG OPL 6.3 (CPLEX 12.1) on two groups of test problems. In the first group, we used problems given by Fire Department of Masshad (MASH problems in Table 3). In the second group, we generated some test problems randomly (RAND problems in Table 3)[1]. We implemented our algorithm in Visual Studio C++ 2010. All implementations were done on a Notebook i5 2.5GHz with 4GB of RAM when only one core was used for our proposed algorithm and all cores were used for OPL. All results are given in Table 3.

Note 5. In some test problems, OPL terminated before convergence because the memory was not enough. In these situations, we reported the truncated results of OPL.

In Table 3, Gap is computed as follows:

$$Gap = \frac{z - \overline{z}}{\overline{z}},$$

where z is the objective value obtained from our algorithm and $\overline{z}$ is either the objective value of the best integer solution found by OPL or the objective value of the best node reported by OPL.

The mean of gaps of our proposed algorithm and the best integer solution found by OPL is only 3.7e-4. Despite OPL used all cores to perform its computations and the implementation of our algorithm used only one core, the running time of our proposed algorithm is very less than OPL's running time. From table 3, it is clear that, in 63% of problems, the given solution is equal to the best integer found by CPLEX. In 17% of problems, the given solution is better than CPLEX's solution. In these problems, CPLEX was terminated because the memory was not enough. In 20% of problems, the given solution is worse than CPLEX's solution. But in these problems the maximum gap from feasible solution given by CPLEX is only $9.6e-3$ and the running time of our proposed algorithm was 20 times better than CPLEX's running time, approximately. In some problems, CPLEX is not able to find the optimal solution. In these situations, the lower bound given by CPLEX is suitable for comparing the result. From Table 3, the maximum gap from the lower bound is $1.2e-2$. Thus we can say that the maximum error of our heuristic on problems is less than $2e-2$ when the running time is almost 10 times better than CPLEX's running time. Therefore, from Table 3, we can conclude that our proposed algorithm is applicable for solving the medium scale problems in a real time.

---

[1] For acquiring hold of the test problems, please contact the corresponding author (rghanbari@um.ac.ir) or download at http://profsite.um.ac.ir/∼rghanbari/IJNAO.rar

Table 3: Numerical Results

			Our Algorithm		OPL			Gap	
Problem	n	q	Best	Time (sec)	Best(integer)	Best(node)	Time(sec)	Integer	Node
MASH1	34	120	1883.5649	0.16	1883.5649	1883.5649	65.52	0	0
MASH2	34	120	3100.1002	1.03	3098.9425	3074.4241	1729.29	4e-4	8.4e-4
MASH3	34	120	1731.2355	0.67	1731.2355	1731.2355	267.40	0	0
MASH4	34	120	6538.2014	2.87	6536.7751	6509.6999	1012.93	2e-4	4.4e-4
MASH5	34	120	591.8589	0.83	591.8589	591.8589	141.85	0	0
MASH6	34	120	1630.4577	0.05	1614.9958	1614.9958	5.10	9.6e-3	9.6e-3
MASH7	34	120	0	0.00	0	0	0.08	0	0
MASH8	34	120	2164.0976	1.28	2164.0976	2147.2527	1285.28	0	7.8e-4
MASH9	34	120	530.2508	0.41	530.2508	530.2508	1.53	0	0
MASH10	34	120	1841.3675	2.87	1841.3675	1791.8676	873.59	0	2.76e-2
MASH11	34	120	222.7586	1.20	222.7586	222.7586	26.21	0	0
MASH12	34	120	7034.8948	3.62	7034.8948	6976.9721	689.24	0	8.3e-3
MASH13	34	120	929.6792	0.27	929.6792	929.6792	1.23	0	0
MASH14	34	120	1815.4193	1.57	1816.1045	1756.3611	1189.90	-4e-4	3.36e-2
RAND1	5	100	89.1726	0.00	89.1726	89.1726	0.03	0	0
RAND2	30	100	4766.9702	0.08	4766.9702	4766.9702	1.89	0	0
RAND3	60	100	9055.7206	11.79	9055.7206	9042.4482	698.93	0	1.5e-3
RAND4	100	100	7181.3464	87.25	7181.5357	7123.6819	895.88	-3e-4	8.1e-3
RAND5	5	150	2492.4775	0.00	2492.4775	2492.4775	4.35	0	0
RAND6	30	150	5116.5434	1.28	5118.1958	5075.0990	850.25	-1e-4	8.2e-3
RAND7	60	150	3993.0682	16.35	3993.0682	3948.5215	784.03	0	1.1e-2
RAND8	100	150	9308.8906	66.46	9309.9365	9047.4370	45.04	-1e-4	2.9e-2
RAND9	5	200	168.5741	0.00	168.5741	168.5741	0.11	0	0
RAND10	30	200	933.2573	0.66	933.2573	933.2573	154.83	0	0
RAND11	60	200	3437.8746	15.98	3439.9514	3389.9545	706.86	-6e-4	1.4e-2
RAND12	100	200	3186.5378	131.57	3185.4674	3159.9767	762.33	3e-4	8.4e-3
RAND13	5	250	22.2134	0.00	22.2134	22.2134	0.2	0	0
RAND14	30	250	93.9136	1.36	93.9136	93.9136	9.44	0	0
RAND15	60	250	1076.3503	12.33	1074.8568	1063.2535	890.00	14e-4	1.2e-2
RAND16	100	250	2707.3912	123.80	2706.2883	2667.7602	860.47	4e-4	1.5e-2

^{t}: truncated (Out of memory).
*: optimal.

5 Conclusions and further works

Assigning available resources to fire stations is a main task of fire department's administrator in a city. The importance of this problem increases when the number of available resources are inadequate. In this situation, the goal is to assign the limited available resources to fire stations such that the associated penalties of the shortages are minimized. Here, we first gave a mathematical approach to consider some penalties for the shortage. Next, we gave an integer program to minimize the sum of associated penalties. The proposed model can be used in many other problems arisen from health services, emergency management, and so on. We also proposed a heuristic to efficiently solve the problem in a reasonable time. Our proposed heuristic had two phases. In the first phase, using a greedy approach, our proposed heuristic constructed a proper feasible solution. Next, in the second phase, we proposed a local search to improve the quality of the solution constructed in the first phase. To show the efficiency of our proposed heuristic, we compared our proposed heuristic with CPLEX based on the running time and the quality of obtained solutions on two groups of problems (real-word prob-

lems and randomly generated problems). The numerical results showed that on the 80% of benchmark problems, the obtained solution was the same as CPLEX's solutions. Also, the running time of our proposed algorithm was almost 10 times better than CPLEX's running time, in average.
Finally, areas of further research within the framework of our proposed model and heuristic include

1. The second phase of our heuristic was a local search; so we expect that if we can propose a advanced search method such as tabu search or variable neighborhood search and so on, then the quality of the obtained solution will be increased.

2. In the proposed model, we assumed that some parameters were given in a numerical form (e.g., p_{kj} and pr_{kj}). We guess that the reliability of model is increased when the value of these parameters are given by intervals or even by linguistic variable. It is evident that in this situation, some advanced robust optimization method must be used to solve the model.

3. The proposed model is suitable when we have the shortage in our existing resources, and this shortage is very far from the minimum required resource to cover an event. When the level of shortage is not high or there is not a shortage in each resource, we can give a model to cover the events when they can be occurred in parallel.

Acknowledgements

Authors thank the Research Council of Fire Department of Mashhad for providing data. Also, authors are grateful to the two anonymous referees and the editor for their constructive comments. Consideration of all improved the presentation.

References

1. Chanta, S., Mayorga, M., and McLay, L. A. *Improving emergency service in rural areas: a bi-objective covering location model for EMS systems,* Ann. Oper. Res. 221 (2014), 133–159.

2. Chou, J .-S., Tsai, C. -F., Chen, Z. -Y., and Sun. M. -H. *Biological-based genetic algorithms for optimized disaster response resource allocation*, Computers & Industrial Engineering 74 (2014), 52–67.

3. Furness, A. and Muckett, M. *Introduction to Fire Safety Management*, Elsevier Ltd., 2007.

4. Ghanbari, R. and Mahdavi-Amiri, N. *Solving bus terminal location problems using evolutionary algorithms*, Applied Soft Computing, 11 (2011), 991–999.

5. Goldberg J. B. *Operations research models for the deployment of emergency services vehicles*, EMS Management Journal, 1 (2004), 20–39.

6. Guo, M., Li, B., Zhang, Z., Wu, S., and Song, J. *Efficiency evaluation for allocating community-based health services*, Computers & Industrial Engineering, 65 (2013), no. 3, 395–401.

7. Huang, Y., Fan, Y., and Cheu, R. *Optimal allocation of multiple emergency service resources for protection of critical transportation infrastructure*, Transportation Research Record: Journal of the Transportation Research Board 2022 (2007),1–8.

8. Maleki, M., Majlesinasab, N., and Sepehri, M. M. *Two new models for redeployment of ambulances*, Computers & Industrial Engineering 78 (2014), 271–274.

9. Marianov, V. and ReVelle, C. *The capacitated standard response fire protection siting problem: Deterministic and probabilistic models*, Annals of Operations Research 40 (1992), no. 1, 303–322.

10. Papadimitriou, C. H. and Steiglitz, K. *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1982.

11. Simpson, N. C. and Hancock, P. G. *Fifty years of operational research and emergency response*, Journal of Opertional Research Society 60 (2009), 126–139.

12. Wright, P. D., Liberatore, M. J., and Nydick, R. L. *A survey of operations research models and applications in homeland security*, Interfaces 36 (2006), no. 6, 514–529.

13. Yang, L., Jones, B. F., and Yang, S. -H. *A fuzzy multi-objective programming for optimization of fire station locations through genetic algorithms*, European Journal of Operational Research 181 (2007), no. 2, 903–915.

Persian Translation of Abstracts

روش محاسباتی برای حل معادله انتگرال فردهلم با هسته منفرد ضعیف در فضاهای هسته بازتولید

دانیال حامدزاده و اسمعیل بابلیان

دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، گروه ریاضی

دریافت مقاله ۱۵ آذر ۱۳۹۵، دریافت مقاله اصلاح شده ۳ فروردین ۱۳۹۶، پذیرش مقاله ۲۴ خرداد ۱۳۹۶

چکیده : در این مقاله، روشی جدید برای حل دسته ای از معادلات انتگرال فردهلم نوع دوم با هسته منفرد ضعیف را در فضاهای هسته بازتولید معرفی میکنیم. با استفاده از بسط تیلور تابع مجهول تکینگی معادله را رفع کرده و در ادامه با استفاده از پایه فضاهای هسته بازتولید معادله را حل میکنیم. در نهایت با چند مثال عددی کارایی روش مفروض را نشان میدهیم.

کلمات کلیدی : هسته منفرد ضعیف؛ معادلات انتگرال فردهلم؛ سری تیلور؛ فضای هسته بازتولید.

یک روش فیلتر نایکنوای کاهشی جدید بدون اطلاعات گرادیان برای حل دستگاههای معادلات مقیاس بزرگ

فرشید ارزانی و محمدرضا پیغامی

مرکز پژوهشی محاسبات علمی در بهینه سازی و مهندسی سامانه ها، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران
دانشکده ریاضی، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران – صندوق پستی ۱۶۱۸-۱۶۳۱۵

دریافت مقاله ۱۸ شهریور ۱۳۹۴، دریافت مقاله اصلاح شده ۲۴ فروردین ۱۳۹۵، پذیرش مقاله ۲۱ تیر ۱۳۹۶

چکیده : در این مقاله، یک روش باقیمانده طیفی مشتق آزاد جدید برای حل دستگاه های معادلات مقیاس بزرگ ارایه می دهیم. الگوریتم پیشنهادی مجهز به یک فیلتر نایکنوای چند بعدی کاهشی است که از پوشش های آن زمانی که طول گام جستجوی خطی کاهش می یابد، کاسته می شود. همچنین، روش پیشنهادی با یک تکنیک جستجوی خطی نایکنوای رهاسازی شده آمیخته می شود که به الگوریتم اجازه می دهد تا از خاصیت نایکنوایی از ابتدا بهره بگیرد. تحت برخی شرایط استاندارد، خاصیت همگرایی سراسری الگوریتم پیشنهادی اثبات می شود. نتایج عددی روی مسایل آزمونی کارایی الگوریتم جدید را در عمل نیز نشان می دهد.

کلمات کلیدی : تکنیک فیلتر کاهشی؛ دستگاه معادلات؛ جستجوی خطی نایکنوا؛ همگرایی سراسری.

روش‌های وزن‌دار غیرنوسانی برای معادلات با مشتقات جزئی چندبعدی غیرخطی سهموی تباهیده

روح‌اله عابدیان

دانشگاه تهران، پردیس دانشکده‌های فنی، دانشکده علوم مهندسی

دریافت مقاله ۲۹بهمن ۱۳۹۵، دریافت مقاله اصلاح شده ۱۷ اردیبهشت ۱۳۹۶، پذیرش مقاله ۲۱ تیر ۱۳۹۶

چکیده : در این مقاله، یک روش برای تقریب جواب‌های معادلات غیرخطی سهموی تباهیده که ممکن است دارای ناپیوستگی در جواب‌ها باشند را ارائه می شود. در حالت یک بعدی، همانند ایده روش گلرکین ناپیوسته موضعی، ابتدا معادله سهموی تباهیده به صورت یک سیستم غیرخطی مرتبه یک در نظرگرفته می‌شود و سپس این سیستم با استفاده از یک روش تفاضلات متناهی وزن‌دار غیرنوسانی برای معادلات قوانین پایستگی حل می‌گردد. این برای اولین بار است که محدود کننده مینمد همراه با ایده روش وزن‌دار غیرنوسانی برای معادلات سهموی تباهیده به‌کار گرفته می‌شود. همچنین ذکر این نکته ضروری است که دقت روش جدید در نواحی هموار از مرتبه پنج و در نواحی تکینگی از مرتبه دو می‌باشد. با استفاده از مسائل گوناگون چند بعدی، خواص دقت، استواری و وضوح بالا روش جدید نمایش داده می‌شود.

کلمات کلیدی : روش‌های وزن‌دار غیرنوسانی؛ روش تفاضلات متناهی؛ معادله چندبعدی غیرخطی سهموی تباهیده؛ معادله محیط متخلخل.

یک روش ترکیبی جدید بر اساس عملگرهای شبه دیفرانسیل و موجک هار برای حل معادلات دیفرانسیل معمولی

محمد باقر قائمی، مجید جمال‌پور بیرگانی و المیرا نبی زاده مرصل فرد

دانشکده ریاضی، دانشگاه علم و صنعت ایران

دریافت مقاله ۱۴ تیر ۱۳۹۴، دریافت مقاله اصلاح شده ۲۷ خرداد ۱۳۹۶، پذیرش مقاله ۲۰ آبان ۱۳۹۶

چکیده : در این مقاله ما با ترکیب عملگرهای شبه دیفرانسیل و موجک هار یک روش جدیدی برای حل معادلات دیفرانسیل خطی و غیر خطی ارائه می‌دهیم. روش ارائه شده از لحاظ کارایی و سادگی بسیار عالی عمل می کند.

کلمات کلیدی : موجک هار؛ تبدیل فوریه؛ معادلات دیفرانسیل؛ جواب عددی؛ عملگرهای شبه دیفرانسیل.

بازی‌های دوماتریسی مقید با آرمان‌های فازی و کاربرد آن در مذاکرات هسته‌ای

حمید بیگدلی[1] ، حسن حسن‌پور [2] و جواد طیبی [3]

[1] پژوهشگر، پژوهشکده عالی جنگ، دانشگاه فرماندهی و ستاد آجا
[2] استادیار، گروه ریاضی، دانشکده علوم ریاضی و آمار، دانشگاه بیرجند
[3] استادیار، گروه صنایع، دانشکده مهندسی صنایع و کامپیوتر، دانشگاه صنعتی بیرجند

دریافت مقاله ۴ اردیبهشت ۱۳۹۵، دریافت مقاله اصلاح شده ۳۰ بهمن ۱۳۹۵، پذیرش مقاله ۱ آذر ۱۳۹۶

چکیده : هدف این تحقیق، حل بازی‌های دوماتریسی مقید در محیط فازی است. این دسته از بازی‌های مجموع ناصفر دونفره، با راهبردهای متناهی و آرمان‌های فازی در نظر گرفته شده است که در آن مجموعه محدودیت‌های خطی روی راهبردها اعمال شده اند. بازی‌های مجموع ناصفر مقید به صورت تک‌هدفی و چندهدفی مورد بررسی قرار گرفته است. نشان داده شده که جواب تعادل در حالت تک‌هدفی از حل یک مسئله‌ی برنامه‌ریزی درجه دوم با محدودیت‌های خطی به دست می‌آید. همچنین برای یافتن نقطه‌ی تعادل در مسائل چندهدفی با محدودیت‌های قطعی و فازی، تعدادی مسئله‌ی برنامه‌ریزی ریاضی معرفی شده است. در نهایت، یک کاربرد سیاسی از چنین بازی‌هایی ارائه شده که در مورد مذاکرات هسته‌ای بین دو کشور می‌باشد.

کلمات کلیدی : بازی چندهدفی؛ بازی مقید؛ بازی مقید فازی؛ مذاکرات هسته‌ای.

برخی روشهای انتگرالگیری نردسیک کارا برای مسایل مقدار اولیه

نسرین برقی اسکویی، علی عبدی و غلامرضا حجتی

دانشکده علوم ریاضی، دانشگاه تبریز

دریافت مقاله ۹ مرداد ۱۳۹۵، دریافت مقاله اصلاح شده ۲۰ آبان ۱۳۹۵، پذیرش مقاله ۱ آذر ۱۳۹۶

چکیده : در این مقاله، در ادامه ساخت روش های عددی کارا برای حل مسایل مقدار اولیه سخت، روش های خطی عمومی با مشتق دوم نردسیک نوع دو با مرتبه p=s را می سازیم، که در آن s تعداد مراحل میانی است و مرتبه مرحله ای q=p می باشد. پیاده سازی روش های ساخته شده با طول گام ثابت و متغیر بحث قرار گرفته است که کارایی روشها را تأیید می کند.

کلمات کلیدی : معادلات دیفراتسیل سخت؛ روشهای خطی عمومی با مشتق دوم نردسیک؛ $A-$ و $L-$ پایداری؛ پیاده سازی با طول گام متغیر.

مساله تخصیص منابع در حالت کمبود به پایگاههای آتش نشانی

عفت السادات علوی و رضا قنبری

دانشگاه فردوسی مشهد، دانشکده علوم ریاضی، گروه ریاضی کاربردی

دریافت مقاله ۸ آبان ۱۳۹۳، دریافت مقاله اصلاح شده ۲۷ خرداد ۱۳۹۶، پذیرش مقاله ۱۳ دی ۱۳۹۶

چکیده : تخصیص منابع موجود به پایگاه های آتش نشانی یکی از مهمترین وظایف سازمان اتش نشانی در یک شهر است. اهمیت این مساله هنگامی که در تامین منابع دچار کمبود هستیم، افزایش می یابد. در این وضعیت، هدف تخصیص منابع دارای کمبود به پایگاه هایی است که جریمه کمبود در کل کمینه شود. در این جا، با استفاده از یک روش تحلیلی مقادیر جریمه را برای کمبود محاسبه می کنیم. سپس، یک مدل برنامه ریزی عدد صحیح برای کمینه کردن جریمه ها ارایه می کنیم. این مدل می تواند برای بسیاری از مسایل دیگر مانند سرویس های سلامت و یا مدیریت اورژانس و غیره به کار آید. افزون براین، در این جا یک روش ابتکاری برای حل مساله در یک زمان معقول ارایه خواهیم داد. الگوریتم ابتکاری پیشنهادی دو مرحله دارد. در مرحله اول، با استفاده از یک ایده حریصانه یک جواب مناسب شدنی می سازیم. سپس در مرحله دوم، با استفاده از یک الگوریتم جستجو محلی کیفیت جواب مرحله یک را بهبود می بخشیم. برای نشان دادن کارایی الگوریتم پیشنهادی، رفتار عددی آن را با نتایج CPLEX از منظر سرعت و کیفیت جواب به رو ی دو دسته از مسایل (نمونه های واقعی و نمونه های تصادفی) مقایسه می کنیم. نتایج عددی نشان می دهد که روی ۸۰٪ مساله های آزمون جواب به دست آمده مشابه جواب CPLEX است درحالیکه به طور متوسط سرعت الگوریتم پیشنهادی ۱۰ برابر CPLEX است.

کلمات کلیدی : مساله تخصیص منابع؛ پایگاه‌های آتش نشانی؛ کمبود؛ برنامه‌ریزی عدد صحیح؛ الگوریتم ابتکاری.

Aims and scope

Iranian Journal of Numerical Analysis and Optimization (IJNAO) is published twice a year by the Department of Applied Mathematics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad. Papers dealing with different aspects of numerical analysis and optimization, theories and their applications in engineering and industry are considered for publication.

Journal Policy

After receiving an article, the editorial committee will assign referees. Refereeing process can be followed via the web site of the Journal.
The manuscripts are accepted for review with the understanding that the work has not been published and also it is not under consideration for publication by any other journal. All submissions should be accompanied by a written declaration signed by the author(s) that the paper has not been published before and has not been submitted for consideration elsewhere.

Instruction for Authors

The Journal publishes all papers in the fields of numerical analysis and optimization. Articles must be written in English.
All submitted papers will be refereed and the authors may be asked to revise their manuscripts according to the referee's reports. The Editorial Board of the Journal keeps the right to accept or reject the papers for publication.
The papers with more than one authors, should determine the corresponding author. The e-mail address of the corresponding author must appear at the end of the manuscript or as a footnote of the first page.
It is strongly recommended to set up the manuscript by Latex or Tex, using the template provided in the web site of the Journal. Manuscripts should be typed double-spaced with wide margins to provide enough room for editorial remarks.
References should be arranged in alphabetical order by the surname of the first author as examples below:
[1] Stoer, J. and Bulirsch, R. *Introduction to Numerical Analysis*, Springer-verlag, New York, 2002.
[2] Brunner, H. *A survey of recent advances in the numerical treatment of Volterra integral and integro-differential equations*, J. Comput. Appl. Math. 8 (1982), 213-229.

Submission of Manuscripts

Authors may submit their manuscripts by either of the following ways:

a) Online submission (pdf or dvi files) via the website of the Journal at:

http://ijnao.um.ac.ir

b) Via journal's email mjms@um.ac.ir

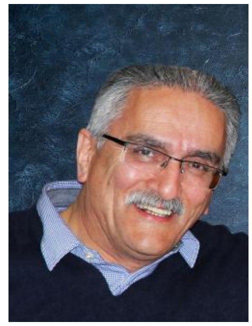
Copyright Agreement

Upon the acceptance of an article by the Journal, the corresponding author will be asked to sign a ”Copyright Transfer Agreement” (see the web site) and send it to the Journal address. This will permit the publisher to publish and distribute the work.

Reviewers Acknowledgment

Peer reviewers are key to advancing scholarship and contributing to the quality of scholarly journals. We would like to sincerely thank the following reviewers who have taken part in the peer-review process for Iranian Journal of Numerical Analysis and Optimization during the last year.

Abbasbandi, S.	Afsharnezhad, Z.
Alizadeh Afrouzi, G.	Amani Rad, J.
Amini, K.	Aminikhah, H.
Ashrafi, A.	Aziz, I.
Babaie-Kafaki, S.	Babolian, E.
Bargerd, R.	Baghani, O.
Beik, F.	Bras, M.
Cen, Z.	Davoodi, N.
Dolati, A.	Effati, S.
Emrouznejad, A.	Erfanian, M.
Eslami, M.	Fakharzadeh Jahromi, A.
Farahi, M. H.	Faraz, N.
Fathali, J.	Gachpazan, M.
Garshasbi, M.	Ghanbari, R.
Ghorbani, A.	Hadizadeh Yazdi, M.
Hajarian, M.	Hassanpur, H.
Hesameddini, E.	Jafari, H.
Jiwari, R.	Kadkhodayan, M.
Kaeppeli, R.	Karimi, K.
Keyanpour, M.	Kheiri, H.
Khojasteh, D.	Kmet, S.
Lakestani, M.	Liu, Y.
Lv, X.	Mahdavi-Amiri, N.
Markolefas, S.	Mehdizadeh Khalsaraei, M.
Mirzazadeh, M.	Mokhtari, R.
Mungkas, S.	Naghipoor, J.
Namjoo, M.	Navidi, H.
Noori Skandari, M.	Pourgholi, R.
Rahbarnia, F.	Riaz, S.
Sadeghpour Gildeh, B.	Safi, M.
Salahi, M.	Salehi, R.
Salehi Fathabadi, H.	Sarafrazi, S.
Sharifi, S.	Shirazian, M.
Soheili, Ali R.	Soleymani Fard, O.
Sun, W.	Sun, Y.
Toutounian, F.	Vahidian Kamyad, A.
Zhang, S.	Zhao, G.



دکتر حامد رضا طارقیان

ما نه آنیم که در بازی تکراری این چرخ فلک　　　هر که از دیده مان رفت ز خاطر ببریم
یا که چون فصل خزان آمد و گل رفت به خواب　　　دل به عشق دگری داده ز آنجا برویم

حامد رضا طارقیان، در چهاردهم مهرماه ۱۳۳۵ در تربت حیدریه و در خانواده ای از اهل علم و ادب چشم به جهان گشود. خود او می گوید به دلیل ارادت ویژه ای که پدر ایشان به حضرت رضا (ع) داشته اند، نام حامد رضا را برای ایشان برگزیدند. سه سال اول دبستان را در زادگاه خود و بقیه دوران تحصیل را تا اخذ دیپلم در مشهد گذرانید. در دوران تحصیل همواره جزء شاگردانی بود که در مسابقات علمی منطقه ای و ملی به عنوان دانش آموز برگزیده شرکت می نمود و همواره جزء دانش آموزان برتر و نمونه بود. در سال ۱۳۵۴ برای ادامه تحصیل عازم انگلستان گردید و پس از طی مراحل آموزش زبان انگلیسی و دروس مورد نیاز برای ورود به دانشگاه های انگلیس، در سپتامبر ۱۹۷۶ دوره کارشناسی را شروع نموده که بلافاصله بعد از اتمام دوره مذکور، دوره های کارشناسی ارشد و دکتری را ادامه می دهد. وی در ۲۸ می ۱۹۸۶ از رساله دکتری خود تحت عنوان **طراحی یک سیستم زمان بندی برای آمبولانس سرویس انگلستان** دفاع نمود. بعد از اتمام تحصیل با وجود تمایل و اصرار دانشگاه محل تحصیل برای استخدام ایشان؛ بعد از یک سال، در خردادماه ۱۳۶۶ به ایران مراجعه و با وجود اینکه از چندین دانشگاه در تهران دعوت به کار شده بود، از بهمن ماه ۱۳۶۶ همکاری خود را با گروه ریاضی دانشکده علوم ریاضی (در آن زمان دانشکده علوم) دانشگاه فردوسی مشهد شروع نمود.

دکتر طارقیان در خردادماه ۱۳۷۸ به مرتبه دانشیاری و در فروردین ماه ۱۳۸۷ به مرتبه استادی ارتقا یافت. ایشان در دوران خدمت در دانشگاه فردوسی مشهد، مسئولیت های مهمی را بر عهده داشته است. در جایی می گوید "به مسئولیت هایی که در دوران خدمت، اعم از آموزشی، پژوهشی و اجرایی بر عهده می گرفته ام، همواره همچون فریضه ای مقدس می نگریسته و همواره سعی می داشته ام تا شرایط لازم برای انجام مسئولیت هایی با این ویژگی را در خود ایجاد کنم. این نگاه باعث می شده تا دست حمایت پروردگار را در لحظه لحظه دوران کاری خود به عینه مشاهده کنم". بعضی از مسئولیت های ایشان به شرح زیر است:

مسئولیت کتابخانه پروفسور فاطمی در دانشکده علوم ریاضی، مدیریت دفتر نظارت و ارزیابی دانشگاه، ریاست دانشکده علوم ریاضی، مدیر مسئول نشریه پژوهش های علوم ریاضی، عضو هیات ممیزه دانشگاه فردوسی مشهد، مدیر مسئول مجله ایرانی آنالیز عددی و بهینه سازی، مسئول کمیته تخصصی علوم پایه در هیات ممیزه دانشگاه فردوسی مشهد، عضو بنیاد دانشگاه فردوسی مشهد.

ایشان در دوران خدمت در دانشگاه فردوسی مشهد، دانشجویان بسیاری را در مقاطع کارشناسی، کارشناسی ارشد و دکتری فارغ التحصیل نموده است. در این فاصله، سه کتاب تالیفی، پنج کتاب ترجمه، ده ها طرح پژوهشی مصوب شورای دانشگاه در بخش های درون دانشگاهی و برون دانشگاهی، چاپ بیش از پنجاه و پنج مقاله در نشریات معتبر خارجی و داخلی و ارائه بیش از چهل مقاله در کنفرانس ها و همایش های ملی و بین المللی ارائه نموده است.

دکتر حامد رضا طارقیان، معلمی زحمتکش، صبور و گشاده روی، متعهد، با تقوی و با ایمان و همکاری رفیق، پاکدست، خوشرو، همراه و پژوهشگری دقیق، منظم، با حوصله و با پشتکار بود.

در طول بیش از چهار سالی که ایشان مدیر مسئول مجله ایرانی آنالیز عددی و بهینه سازی بود، این مجله مسیر رشد خیره کننده ای را از پایه گذاری اولیه تا تبدیل به یک مجله معتبر علمی را طی نمود. حوصله در پیگیری امور، دقت و وسواس ایشان در انجام صحیح و منظم امور مجله، تشویق دائمی دست اندرکاران مجله برای پیشرفت اهداف مجله، حضور و اختصاص بی ریای وقت برای دنبال کردن مسایل مربوط به مجله، واقعاً توصیف ناپذیر است. غیبت و نبود او، غیبت مشی نایاب یک معلم پژوهشگر است. روحش شاد باد.

گرچه از هر ماتمی خیزد غمی　　　فرق دارد ماتمی با ماتمی
لیک در مرگ مردان بزرگ　　　گفت باید ای دریغا عالمی

سردبیر

Iranian Journal of Numerical Analysis and Optimization

Former : MRJMS

CONTENTS Vol. 8, No.1, pp 1-141, 2018

web site : http://ijnao.um.ac.ir

Email : mjms@um.ac.ir

ISSN : 2423-6977