Numerical Algorithms (2024) 96:799–826
https://doi.org/10.1007/s11075-023-01666-y

ORIGINAL PAPER

Optimal parameter of the SOR-like iteration method for solving absolute value equations

Cairong Chen[1] · Bo Huang[3] · Dongmei Yu[2] · Deren Han[3]

Received: 12 July 2023 / Accepted: 14 September 2023 / Published online: 16 October 2023

Abstract

The SOR-like iteration method for solving the system of absolute value equations of finding a vector x such that $Ax - |x| - b = 0$ with $\nu = \|A^{-1}\|_2 < 1$ is investigated. The convergence conditions of the SOR-like iteration method proposed by Ke and Ma (*Appl. Math. Comput.*, 311:195–202, 2017) are revisited and a new proof is given, which exhibits some insights in determining the convergent region and the optimal iteration parameter. Along this line, the optimal parameter which minimizes $\|T_\nu(\omega)\|_2$ with

$$T_\nu(\omega) = \begin{pmatrix} |1-\omega| & \omega^2\nu \\ |1-\omega| & |1-\omega| + \omega^2\nu \end{pmatrix}$$

and the approximate optimal parameter which minimizes an upper bound of $\|T_\nu(\omega)\|_2$ are explored. The optimal and approximate optimal parameters are iteration-independent, and the bigger value of ν is, the smaller convergent region of the iteration parameter ω is. Numerical results are presented to demonstrate that the SOR-like iteration method with the optimal parameter is superior to that with the approximate optimal parameter proposed by Guo et al. (*Appl. Math. Lett.*, 97:107–113, 2019).

✉ Dongmei Yu
yudongmei1113@163.com

Cairong Chen
cairongchen@fjnu.edu.cn

Bo Huang
bohuang0407@buaa.edu.cn

Deren Han
handr@buaa.edu.cn

1 School of Mathematics and Statistics & Key Laboratory of Analytical Mathematics and Applications (Ministry of Education) & Fujian Provincial Key Laboratory of Statistics and Artificial Intelligence & Fujian Key Laboratory of Analytical Mathematics and Applications (FJKLAMA) & Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University, Fuzhou 350117, People's Republic of China

2 Institute for Optimization and Decision Analytics, Liaoning Technical University, Fuxin 123000, People's Republic of China

3 LMIB-School of Mathematical Sciences, Beihang University, Beijing 100191, People's Republic of China

Keywords Absolute value equations · SOR-like iteration method · Optimal iteration parameter · Convergence condition

Mathematics Subject Classification (2010) 65F10 · 65H10 · 90C30

1 Introduction

In this paper, we consider the solution of the system of absolute value equations (AVE)

$$Ax - |x| - b = 0, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, and $|x| \in \mathbb{R}^n$ denotes the component-wise absolute value of the unknown vector $x \in \mathbb{R}^n$. The AVE (1.1) can be regarded as a special case of the more general AVE $Ax + B|x| - b = 0$ with $B \in \mathbb{R}^{n\times n}$, which was introduced in [26] and further investigated in [10, 17, 25, 30, 32]; just list a few. The AVE is equivalent to linear complementarity problem and mixed integer programming [12, 17, 20, 25], and it recently receives much attention in the optimization community; see [11, 36] and references therein.

It was known that determining the existence of a solution to the AVE is NP-hard [17], and if solvable, the problem of checking whether the AVE has unique or multiple solutions is NP-complete [25]. In addition, one sufficient condition for the AVE (1.1) being uniquely solvable for any b is described in the following proposition.

Proposition 1.1 ([20]) *Assume that $A \in \mathbb{R}^{n\times n}$ is nonsingular. If $\|A^{-1}\|_2 < 1$, then the AVE (1.1) has a unique solution x^* for any $b \in \mathbb{R}^n$.*

Hereafter, we assume that $A \in \mathbb{R}^{n\times n}$ and $\|A^{-1}\|_2 < 1$. Recently, there has been a surge of interest in solving the AVE (1.1) with $\|A^{-1}\|_2 < 1$, and a large number of numerical iteration methods have been proposed, including the generalized Newton method [16, 19], the smoothing Newton method [4], the exact and inexact Douglas-Rachford splitting methods [6], the Levenberg-Marquardt method [13], the SOR-like iteration method [9, 15], the generalization of the Gauss-Seidel iteration method [8], the relaxed-based matrix splitting methods [28], and others; see [1, 3, 5, 14, 18, 21–24, 27, 31, 35] and references therein.

Our work here is inspired by recent studies on the SOR-like iteration method for solving the AVE (1.1) [9, 15]. The SOR-like iteration method is one-parameter-dependent, and thus it is an important problem to determine the optimal iteration parameter. By an optimal iteration parameter, we mean it is the iteration parameter such that the SOR-like iteration method gets the fastest convergence rate. Usually, it seems not to be an easy task to find the optimal value of the involved iteration parameter, while it remains significance to find a somewhat optimal one. Recently, Guo, Wu, and Li [9] obtained an optimal iteration parameter that minimizes the spectral radius of the iteration matrix. However, it is associated with the spectral radius $\rho(D(x^{(k+1)})A^{-1})$, and thus it is iteration-dependent (that is, it may vary with the iteration sequence $\{x^{(k)}\}$). To compute the optimal parameter in every iterative step is expensive, especially when the dimension of A is large, and an approximate one is used in the numerical experiments of [9].

The goal of this paper is twofold: to revisit the convergence conditions of the SOR-like iteration method on solving the AVE (1.1) and to explore respectively an optimal iteration parameter and an approximate optimal iteration parameter for the SOR-like iteration method. It is important that our optimal and approximate optimal parameters are iteration-independent. Numerical results demonstrate that the SOR-like iteration method with our optimal iteration parameter is superior to that with the approximate one proposed in [9] for solving the AVE (1.1). The SOR-like iteration method with our approximate optimal iteration parameter is better than that with the approximate one proposed in [9] for solving the AVE (1.1) in some cases.

The rest of this paper is organized as follows: In Section 2, we revisit the convergence conditions of the SOR-like iteration method for solving the AVE (1.1). Section 3 characterizes the optimal and approximate optimal iteration parameters for the SOR-like iteration method. In Section 4, some numerical examples are given to demonstrate our claims made in the previous sections. Finally, some concluding remarks are given in Section 5.

Notations $\mathbb{R}^{n\times n}$ is the set of all $n \times n$ real matrices and $\mathbb{R}^n = \mathbb{R}^{n\times 1}$. I is the identity matrix with suitable dimension. For $x \in \mathbb{R}^n$, x_i refers to its ith entry, $|x|$ is in $\mathbb{R}^n$ with its ith entry $|x_i|$, and $|\cdot|$ denotes absolute value. $\operatorname{sgn}(x)$ denotes a vector with components equal to $-1, 0$, or 1 depending on whether the corresponding component of the vector x is negative, zero, or positive, respectively. For $x \in \mathbb{R}^n$, $\operatorname{diag}(x)$ represents a diagonal matrix with x_i as its diagonal entries for every $i = 1, 2, \cdots, n$. $\operatorname{tridiag}(u, v, w)$ denotes a tridiagonal matrix that has u, v, and w as the subdiagonal, main diagonal, and superdiagonal entries in the matrix, respectively; $\operatorname{Tridiag}(U, V, W)$ denotes a block tridiagonal matrix that has U, V, and W as the subdiagonal, main diagonal, and superdiagonal block entries in the matrix, respectively. For $X \in \mathbb{R}^{n\times n}$, $\|X\|_2$ denotes the spectral norm of X and is defined by the formula $\|X\|_2 \doteq \max\{\|Xx\|_2 : x \in \mathbb{R}^n, \|x\|_2 = 1\}$, where $\|x\|_2$ is the 2-norm of vector $x \in \mathbb{R}^n$. When X is a square matrix, we denote by $\rho(X)$ its spectral radius. For a symmetric matrix $X \in \mathbb{R}^{n\times n}$, $\lambda_{\max}(X)$ denotes its largest eigenvalue. For a function f, f' denotes its first derivative. $\lim_{x\to x_0^+} f(x)$ and $\lim_{x\to x_0^-} f(x)$ denote the right and left limits of a function $f(x)$ at a point x_0, respectively.

2 Revisit the convergence conditions of the SOR-like iteration method

The SOR-like iteration method for solving the AVE (1.1) is firstly proposed by Ke and Ma [15]. After that, in a different perspective, Guo, Wu, and Li in [9] present some new convergence conditions of the SOR-like iteration method. In this section, along the lines in [15], we will further study the convergence conditions of the SOR-like iteration method for solving the AVE (1.1). For this purpose, we first briefly review the SOR-like iteration method for solving the AVE (1.1).

The AVE (1.1) is equivalent to

$$\begin{cases} Ax - y = b, \\ y - |x| = 0, \end{cases} \tag{2.1}$$

that is,

$$\boldsymbol{A z} \doteq \begin{pmatrix} A & -I \\ -D(x) & I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} \doteq \boldsymbol{b},$$

where $D(x) \doteq \operatorname{diag}(\operatorname{sgn}(x))$. By splitting the coefficient matrix $\boldsymbol{A} = \mathcal{D} - \mathcal{L} - \mathcal{U}$ with

$$\mathcal{D} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}, \quad \mathcal{L} = \begin{pmatrix} 0 & 0 \\ D(x) & 0 \end{pmatrix}, \quad \mathcal{U} = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix},$$

one can obtain the following equation

$$(\mathcal{D} - \omega\mathcal{L})\boldsymbol{z} = [(1-\omega)\mathcal{D} + \omega\mathcal{U}]\,\boldsymbol{z} + \omega\boldsymbol{b},$$

where the parameter $\omega > 0$. That is,

$$\begin{pmatrix} A & 0 \\ -\omega D(x) & I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (1-\omega)A & \omega I \\ 0 & (1-\omega)I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \omega b \\ 0 \end{pmatrix}. \tag{2.2}$$

Based on (2.2), the following SOR-like iteration scheme is established in [15]:

$$\begin{pmatrix} A & 0 \\ -\omega D(x^{(k+1)}) & I \end{pmatrix} \begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = \begin{pmatrix} (1-\omega)A & \omega I \\ 0 & (1-\omega)I \end{pmatrix} \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} + \begin{pmatrix} \omega b \\ 0 \end{pmatrix}. \tag{2.3}$$

Specifically, the SOR-like iteration method is described in the following Algorithm 1, in which (2.4) is obtained from (2.3).

Algorithm 1 SOR-like iteration method for solving the AVE (1.1) [15]

Let $A \in \mathbb{R}^{n\times n}$ be a nonsingular matrix and $b \in \mathbb{R}^n$. Given initial vectors $x^{(0)} \in \mathbb{R}^n$ and $y^{(0)} \in \mathbb{R}^n$, for $k = 0, 1, 2, \cdots$ until the iteration sequence $\left\{(x^{(k)}, y^{(k)})\right\}_{k=0}^{\infty}$ is convergent, compute

$$\begin{cases} x^{(k+1)} = (1-\omega)x^{(k)} + \omega A^{-1}(y^{(k)} + b), \\ y^{(k+1)} = (1-\omega)y^{(k)} + \omega|x^{(k+1)}|, \end{cases} \tag{2.4}$$

where the iteration parameter ω is a positive constant.

Let (x^*, y^*) be the solution pair of the nonlinear equation (2.1). Throughout this paper, denote $e_k^x = x^* - x^{(k)}$ and $e_k^y = y^* - y^{(k)}$, where $(x^{(k)}, y^{(k)})$ is generated by Algorithm 1. Then, for the SOR-like iteration method, the following theorem is known.

Theorem 2.1 ([15]) *Assume that $A \in \mathbb{R}^{n\times n}$ is a nonsingular matrix and $b \in \mathbb{R}^n$. Denote*

$$\nu = \|A^{-1}\|_2, \quad a = |1-\omega| \quad \text{and} \quad c = \omega^2 \nu. \tag{2.5}$$

If

$$0 < \omega < 2 \qquad \text{and} \qquad a^4 - 3a^2 - 2ac - 2c^2 + 1 > 0, \tag{2.6}$$

then the following inequality

$$|||(e^x_{k+1}, e^y_{k+1})|||_\omega < |||(e^x_k, e^y_k)|||_\omega \tag{2.7}$$

holds for $k = 0, 1, 2, \cdots$. Here, the norm $|||\cdot|||_\omega$ is defined by

$$|||(e^x_k, e^y_k)|||_\omega := \sqrt{\|e^x_k\|_2^2 + \omega^{-2}\|e^y_k\|_2^2}. \tag{2.8}$$

Proof While a proof of Theorem 2.1 is given in [15], we will present the following proof because the expression of the eigenvalue (2.11) is needed later for discussing the optimal iteration parameter.

It follows from the proof of Theorem 3.1 in [15] that

$$\begin{pmatrix} \|e^x_{k+1}\|_2 \\ \omega^{-1}\|e^y_{k+1}\|_2 \end{pmatrix} \le \begin{pmatrix} |1-\omega| & \omega^2\nu \\ |1-\omega| & |1-\omega| + \omega^2\nu \end{pmatrix} \begin{pmatrix} \|e^x_k\|_2 \\ \omega^{-1}\|e^y_k\|_2 \end{pmatrix};$$

in other words,

$$|||(e^x_{k+1}, e^y_{k+1})|||_\omega \le \|T_\nu(\omega)\|_2 \cdot |||(e^x_k, e^y_k)|||_\omega \tag{2.9}$$

with

$$T_\nu(\omega) = \begin{pmatrix} a & c \\ a & a+c \end{pmatrix}, \tag{2.10}$$

where a and c are defined as in (2.5). In order to prove the inequality (2.7), we turn to consider the choice of the parameter ω such that $\|T_\nu(\omega)\|_2 < 1$.

Since[1]

$$H_\nu(\omega) = T_\nu(\omega)^T T_\nu(\omega) = \begin{pmatrix} 2a^2 & a^2 + 2ac \\ a^2 + 2ac & a^2 + 2c^2 + 2ac \end{pmatrix}$$

is a symmetric positive semi-definite matrix, we have $\|T_\nu(\omega)\|_2^2 = \rho\left(T_\nu(\omega)^T T_\nu(\omega)\right) = \lambda_{\max}\left(H_\nu(\omega)\right)$. Assume that λ is an eigenvalue of $H_\nu(\omega)$. Then,

$$(\lambda - 2a^2)[\lambda - (a^2 + 2c^2 + 2ac)] - (a^2 + 2ac)^2 = 0,$$

namely,

$$\lambda^2 - (3a^2 + 2c^2 + 2ac)\lambda + a^4 = 0,$$

from which we obtain

$$\lambda = \frac{3a^2 + 2c^2 + 2ac \pm \sqrt{(3a^2 + 2c^2 + 2ac)^2 - 4a^4}}{2}.$$

[1] The rest of the proof seems new.

Consequently,

$$\lambda_{\max}(H_\nu(\omega)) = \frac{3a^2 + 2c^2 + 2ac + \sqrt{(3a^2 + 2c^2 + 2ac)^2 - 4a^4}}{2}. \tag{2.11}$$

In particular,

$$\begin{aligned}\lambda_{\max}(H_\nu(\omega)) < 1 &\Longleftrightarrow 3a^2 + 2c^2 + 2ac + \sqrt{(3a^2 + 2c^2 + 2ac)^2 - 4a^4} < 2 \\ &\Longleftrightarrow \sqrt{(3a^2 + 2c^2 + 2ac)^2 - 4a^4} < 2 - (3a^2 + 2c^2 + 2ac).\end{aligned}$$

Hence, a sufficient condition for the convergence is

$$\begin{cases} 3a^2 + 2c^2 + 2ac < 2, \\ 3a^2 + 2c^2 + 2ac < 1 + a^4. \end{cases} \tag{2.12}$$

From (2.12), we have $\lambda_{\max}(H_\nu(\omega)) < 1$ provided that $1 + a^4 < 2$ and $3a^2 + 2c^2 + 2ac < 1 + a^4$. It is easy to check that $1 + a^4 < 2$ is equivalent to $0 < \omega < 2$, which completes the proof. □

The inequality (2.9) together with $\|T_\nu(\omega)\|_2 < 1$ implies that the sequence $\left\{(x^{(k)}, y^{(k)})\right\}_{k=0}^{\infty}$ generated by the SOR-like iteration scheme (2.4) is linearly convergent. However, the second inequality in (2.6) seems fussy at the first glance, which may make (2.6) harder to be checked than (2.13) in the following Corollary 2.1. This motivates us to take a closer look at (2.6).

Corollary 2.1 ([15]) *Let $A \in \mathbb{R}^{n\times n}$ be a nonsingular matrix, $b \in \mathbb{R}^n$ and $\nu = \|A^{-1}\|_2$. If*

$$\nu < 1 \quad and \quad 1 - \tau < \omega < \min\left\{1 + \tau, \sqrt{\frac{\tau}{\nu}}\right\}, \tag{2.13}$$

where $\tau = \frac{2}{3+\sqrt{5}}$. Then the following inequality

$$|||(e_{k+1}^x, e_{k+1}^y)|||_\omega < |||(e_k^x, e_k^y)|||_\omega$$

holds for $k = 0, 1, 2, \cdots$. Here, $|||(e_k^x, e_k^y)|||_\omega$ is defined as in (2.8).

Let

$$\begin{aligned} f_\nu(\omega) &= 3a^2 + 2c^2 + 2ac - a^4 - 1 \\ &= 3(\omega - 1)^2 + 2\nu^2\omega^4 + 2\nu\omega^2|\omega - 1| - (\omega - 1)^4 - 1 \\ &= \begin{cases} (2\nu^2 - 1)\omega^4 + (4 - 2\nu)\omega^3 + (2\nu - 3)\omega^2 - 2\omega + 1 \doteq f_\nu^1(\omega), & \text{if } \ 0 < \omega < 1; \\ (2\nu^2 - 1)\omega^4 + (4 + 2\nu)\omega^3 - (2\nu + 3)\omega^2 - 2\omega + 1 \doteq f_\nu^2(\omega), & \text{if } \ \omega \geq 1. \end{cases} \end{aligned} \tag{2.14}$$

On the number of real zeros of the polynomials $f_\nu^1(\omega)$ and $f_\nu^2(\omega)$, we obtain the following results by using symbolic computation method of solving semi-algebraic systems.[2]

Proposition 2.1 *Consider the two semi-algebraic systems*

$$\mathcal{S}_1 = \{f_\nu^1(\omega) = 0, \quad \omega > 0, \quad 1-\omega > 0, \quad \nu > 0, \quad 1-\nu > 0\}$$
$$\mathcal{S}_2 = \{f_\nu^2(\omega) = 0, \quad \omega - 1 \geq 0, \quad 2-\omega > 0, \quad \nu > 0, \quad 1-\nu > 0\}$$

formed by (2.14). *We have that the system $\mathcal{S}_1$ (or $\mathcal{S}_2$) has exactly one real solution with respect to ω if and only if $R_1 = 2\nu^2 - 1 < 0$; the system $\mathcal{S}_1$ has exactly two real solutions with respect to ω if and only if $R_1 = 2\nu^2 - 1 > 0$. Moreover, there is no other given number of real solution(s) for the system $\mathcal{S}_1$ (or $\mathcal{S}_2$).*

Proof Here, we only show the proof for $\mathcal{S}_1$ to have exactly one real solution since the other statements can be proved similarly. In fact, using the RegularChains package bulit-in Maple to solve the semi-algebraic system $\mathcal{S}_1$, one can obtain the desired result $R_1 = 2\nu^2 - 1 < 0$.

The detailed procedures in Maple are the following:

```
> with(RegularChains):
> with(ParametricSystemTools):
> with(SemiAlgebraicSetTools):
> R:= PolynomialRing([ω,ν]):
> rrc:= RealRootClassification([f_ν^1(ω)], [ ], [ν, 1 − ν, ω, 1 − ω], [ ], 1, 1, R)[3];
```

□

Remark 2.1 When $R_1 = 2\nu^2 - 1 = 0$, it is easy to verify that the system $\mathcal{S}_1$ (or $\mathcal{S}_2$) has exactly one real solution with respect to ω.

Particularly, if $\nu \neq \frac{\sqrt{2}}{2}$, the zeros of $f_\nu^1(\omega)$ are

$$\begin{aligned}
&-\frac{\alpha}{4\beta} + \frac{\gamma}{2\beta} + \frac{\sqrt{-(8\nu^3 - 16\nu^2 + 4\nu - 1) - (8\nu^2 - 2\nu)\gamma}}{2\beta},\\
&-\frac{\alpha}{4\beta} + \frac{\gamma}{2\beta} - \frac{\sqrt{-(8\nu^3 - 16\nu^2 + 4\nu - 1) - (8\nu^2 - 2\nu)\gamma}}{2\beta},\\
&-\frac{\alpha}{4\beta} - \frac{\gamma}{2\beta} + \frac{\sqrt{-(8\nu^3 - 16\nu^2 + 4\nu - 1) + (8\nu^2 - 2\nu)\gamma}}{2\beta},\\
&-\frac{\alpha}{4\beta} - \frac{\gamma}{2\beta} - \frac{\sqrt{-(8\nu^3 - 16\nu^2 + 4\nu - 1) + (8\nu^2 - 2\nu)\gamma}}{2\beta},
\end{aligned} \tag{2.15}$$

[2] The definition of the semi-algebraic system can be found in [33].

[3] RealRootClassification computes conditions on the parameters for the system to have a given number of real solutions; the detailed algorithm can be found in [34].

and the zeros of $f_\nu^2(\omega)$ are

$$\begin{aligned}
&-\frac{\xi}{4\beta}+\frac{\zeta}{2\beta}+\frac{\sqrt{(8\nu^3+16\nu^2+4\nu+1)-(8\nu^2+2\nu)\zeta}}{2\beta},\\
&-\frac{\xi}{4\beta}+\frac{\zeta}{2\beta}-\frac{\sqrt{(8\nu^3+16\nu^2+4\nu+1)-(8\nu^2+2\nu)\zeta}}{2\beta},\\
&-\frac{\xi}{4\beta}-\frac{\zeta}{2\beta}+\frac{\sqrt{(8\nu^3+16\nu^2+4\nu+1)+(8\nu^2+2\nu)\zeta}}{2\beta},\\
&-\frac{\xi}{4\beta}-\frac{\zeta}{2\beta}-\frac{\sqrt{(8\nu^3+16\nu^2+4\nu+1)+(8\nu^2+2\nu)\zeta}}{2\beta},
\end{aligned} \tag{2.16}$$

where

$$\begin{aligned}
&\alpha=4-2\nu>0, \quad \beta=2\nu^2-1, \quad \xi=4+2\nu>0,\\
&\gamma=\sqrt{-(\nu-1)(\nu+5)}>0, \quad \zeta=\sqrt{-(\nu+1)(\nu-5)}>0.
\end{aligned} \tag{2.17}$$

If $\nu=\frac{\sqrt{2}}{2}$, then the zeros of $f_\nu^1(\omega)$ are

$$-\frac{1}{7}-\frac{1}{28}\sqrt{2}+\frac{1}{28}\sqrt{242+64\sqrt{2}}\approx 0.4579, \quad -\frac{1}{7}-\frac{1}{28}\sqrt{2}-\frac{1}{28}\sqrt{242+64\sqrt{2}}\approx -0.8446, \quad 1,$$

and the zeros of $f_\nu^2(\omega)$ are

$$-\frac{1}{7}+\frac{1}{28}\sqrt{2}+\frac{1}{28}\sqrt{242-64\sqrt{2}}\approx 0.3472, \quad -\frac{1}{7}+\frac{1}{28}\sqrt{2}-\frac{1}{28}\sqrt{242-64\sqrt{2}}\approx -0.5319, \quad 1.$$

Using Proposition 2.1 and Remark 2.1, we can determine that the two real roots of $f_\nu(\omega)=0$ located in $(0,2)$ for $\nu\in(0,1)$ are as follows:

If $0<\nu<\frac{\sqrt{2}}{2}$:

$$\omega_1(\nu)=-\tfrac{\alpha}{4\beta}+\tfrac{\gamma}{2\beta}-\tfrac{\sqrt{-(8\nu^3-16\nu^2+4\nu-1)-(8\nu^2-2\nu)\gamma}}{2\beta} \tag{2.18}$$

$$\omega_2(\nu)=-\tfrac{\xi}{4\beta}-\tfrac{\zeta}{2\beta}+\tfrac{\sqrt{(8\nu^3+16\nu^2+4\nu+1)+(8\nu^2+2\nu)\zeta}}{2\beta} \tag{2.19}$$

If $\frac{\sqrt{2}}{2}<\nu<1$:

$$\omega_3(\nu)=-\tfrac{\alpha}{4\beta}+\tfrac{\gamma}{2\beta}+\tfrac{\sqrt{-(8\nu^3-16\nu^2+4\nu-1)-(8\nu^2-2\nu)\gamma}}{2\beta} \tag{2.20}$$

$$\omega_4(\nu)=-\tfrac{\alpha}{4\beta}-\tfrac{\gamma}{2\beta}+\tfrac{\sqrt{-(8\nu^3-16\nu^2+4\nu-1)+(8\nu^2-2\nu)\gamma}}{2\beta} \tag{2.21}$$

If $\nu=\frac{\sqrt{2}}{2}$:

$$\omega_5=-\tfrac{1}{7}-\tfrac{\sqrt{2}}{28}+\tfrac{\sqrt{242+64\sqrt{2}}}{28}\approx 0.4579,$$

$$\omega_6=1,$$

where α, β, γ, ξ, and ζ are defined as in (2.17). Indeed, we have $0 < \omega_5 < 1 = \omega_6 < 2$. In Appendix A, we prove that $\omega_i(\nu)$ ($i = 1, 2, 3, 4$) are real-valued functions (with respect to ν) in their domains, respectively. In addition, the curves of the derivatives of $\omega_1(\nu)$ and $\omega_2(\nu)$ in $\nu \in (0, \frac{\sqrt{2}}{2})$ and the curves of the derivatives of $\omega_3(\nu)$ and $\omega_4(\nu)$ in $\nu \in (\frac{\sqrt{2}}{2}, 1)$ are shown in Fig. 1, from which we can find that the derivative of $\omega_1(\nu)$ in $\nu \in (0, \frac{\sqrt{2}}{2})$ and the derivative of $\omega_3(\nu)$ in $\nu \in (\frac{\sqrt{2}}{2}, 1)$ are always positive, and the derivative of $\omega_2(\nu)$ in $\nu \in (0, \frac{\sqrt{2}}{2})$ and the derivative of $\omega_4(\nu)$ in $\nu \in (\frac{\sqrt{2}}{2}, 1)$ are always negative. Hence, $\omega_1(\nu)$ is strictly monotonously increasing in $\nu \in (0, \frac{\sqrt{2}}{2})$, $\omega_3(\nu)$ is strictly monotonously increasing in $\nu \in (\frac{\sqrt{2}}{2}, 1)$, $\omega_2(\nu)$ is strictly monotonously decreasing in $\nu \in (0, \frac{\sqrt{2}}{2})$, and $\omega_4(\nu)$ is strictly monotonously decreasing in $\nu \in (\frac{\sqrt{2}}{2}, 1)$, respectively. Furthermore, by using the L'Hospital's rule once (or twice if needed), we have

$$\lim_{\nu\to 0^+} \omega_1(\nu) = \frac{3-\sqrt{5}}{2}, \quad \lim_{\nu\to \frac{\sqrt{2}}{2}^-} \omega_1(\nu) = \omega_5, \quad \lim_{\nu\to 0^+} \omega_2(\nu) = \frac{\sqrt{5}+1}{2}, \quad \lim_{\nu\to \frac{\sqrt{2}}{2}^-} \omega_2(\nu) = 1, \tag{2.22}$$

$$\lim_{\nu\to \frac{\sqrt{2}}{2}^+} \omega_3(\nu) = \omega_5, \quad \lim_{\nu\to 1^-} \omega_3(\nu) = \frac{\sqrt{5}-1}{2}, \quad \lim_{\nu\to \frac{\sqrt{2}}{2}^+} \omega_4(\nu) = 1, \quad \lim_{\nu\to 1^-} \omega_4(\nu) = \frac{\sqrt{5}-1}{2}. \tag{2.23}$$

It follows from (2.22), (2.23), and the strictly monotonous properties of $\omega_i(\nu)$ ($i = 1, 2, 3, 4$) that $\omega_1(\nu) \in (\frac{3-\sqrt{5}}{2}, \omega_5) \subset (0, 2)$ and $\omega_2(\nu) \in (1, \frac{\sqrt{5}+1}{2}) \subset (0, 2)$ and thus $0 < \omega_1(\nu) < 1 < \omega_2(\nu) < 2$ for $\nu \in (0, \frac{\sqrt{2}}{2})$. Similarly, $\omega_3(\nu) \in (\omega_5, \frac{\sqrt{5}-1}{2}) \subset (0, 1)$ and $\omega_4(\nu) \in (\frac{\sqrt{5}-1}{2}, 1) \subset (0, 1)$ and thus $0 < \omega_3(\nu) < \omega_4(\nu) < 1$ for $\nu \in (\frac{\sqrt{2}}{2}, 1)$. Numerically, for a given $\nu \in (0, 1)$, we can obtain the desired real zeros of $f_\nu(\omega)$ from (2.15) or (2.16). For example, by taking $\nu = 3/4$ in (2.15), the approximate values of the four functions are

$$[0.4667, -0.8750, 0.9536, -20.5452],$$

and the corresponding approximate values of $f_\nu^1(\omega)$ are

$$[-5.7732 \times 10^{-15}, -2.5757 \times 10^{-14}, 0, -3.0340 \times 10^{-12}].$$

We see that the first and third functions in (2.15) are the desired roots of $f_\nu^1(\omega) = 0$.

According to Proposition 1.1 and Theorem 2.1, for $0 < \nu < 1$, the problem to find $\omega \in (0, 2)$ such that $f_\nu(\omega) < 0$ is of interest. The fact that $f_\nu(\omega)$ has only two zeros in $(0, 2)$ together with $f_\nu(0) = 1 > 0$ and $f_\nu(2) = 32\nu^2 + 8\nu + 1 > 0$ implies that $f_\nu(\omega) < 0$ if one of the following conditions holds:

$$\omega \in (\omega_1(\nu), \omega_2(\nu)) \doteq \Omega_1, \quad \text{when} \quad \nu \in (0, \frac{\sqrt{2}}{2}), \tag{2.24}$$

$$\omega \in (\omega_3(\nu), \omega_4(\nu)) \doteq \Omega_2, \quad \text{when} \quad \nu \in (\frac{\sqrt{2}}{2}, 1), \tag{2.25}$$

$$\omega \in (\omega_5, \omega_6) \doteq \Omega_3, \quad \text{when} \quad \nu = \frac{\sqrt{2}}{2}. \tag{2.26}$$

In addition, it can be concluded from (2.22), (2.23), and the the strictly monotonous properties of $\omega_i(\nu)$ $(i = 1, 2, 3, 4)$ that the bigger value of ν is, the smaller the range of ω will be. In addition, since $\lim_{\nu\to 1^-} \omega_3(\nu) = \lim_{\nu\to 1^-} \omega_4(\nu) = \frac{\sqrt{5}-1}{2}$, $(\omega_3(\nu), \omega_4(\nu))$ must close to the singleton set $\left\{\frac{\sqrt{5}-1}{2} \approx 0.6180\right\}$ as $\nu \to 1^-$. Figure 2 plots the curves of the function $f_\nu(\omega)$ for some values of ν and the range of the iteration parameter ω, from which our claims are intuitively shown.

Finally, according to the discussion above, we will rewrite Theorem 2.1 as follows, i.e., Theorem 2.2, which exhibits more information than the original one.

Theorem 2.2 *Assume that $A \in \mathbb{R}^{n\times n}$ is a nonsingular matrix and $b \in \mathbb{R}^n$. If one of the conditions* (2.24)–(2.26) *holds, then the inequality* (2.9) *with $\|T_\nu(\omega)\|_2 < 1$ holds for $k = 0, 1, 2, \cdots$, which implies that the sequence $\{x^{(k)}\}_{k=0}^{\infty}$ generated by the SOR-like iteration scheme (2.4) linearly converges to the unique solution of the AVE* (1.1)*. Here, $|||(e_k^x, e_k^y)|||_\omega$ is defined as in* (2.8) *and $T_\nu(\omega)$ is defined as in* (2.10).

3 Optimal and approximate optimal iteration parameters for the SOR-like iteration method

3.1 Optimal iteration parameter

As is known, the SOR-like iteration method involves one parameter and to determine the somewhat optimal iteration parameter is an important problem. From (2.9), one can obtain that

$$0 \le |||(e_{k+1}^x, e_{k+1}^y)|||_\omega \le \|T_\nu(\omega)\|_2 \cdot |||(e_k^x, e_k^y)|||_\omega \le \cdots \le \|T_\nu(\omega)\|_2^{k+1} \cdot |||(e_0^x, e_0^y)|||_\omega. \tag{3.1}$$

If the conditions of Theorem 2.2 hold, then we have $\|T_\nu(\omega)\|_2 < 1$ and $\lim_{k\to\infty} |||(e_k^x, e_k^y)|||_\omega = 0$, that is, the sequence $\{x^{(k)}\}_{k=0}^{\infty}$ generated by the SOR-like iteration scheme (2.4) converges to the unique solution of the AVE (1.1). In addition, from (3.1), the smaller value of $\|T_\nu(\omega)\|_2$ is, the faster the SOR-like iteration method will converge later on. The question is for what $\omega \in (0, 2)$, $\|T_\nu(\omega)\|_2$ is minimized with some given $\nu \in (0, 1)$. According to the proof of Theorem 2.1, it is equiva-

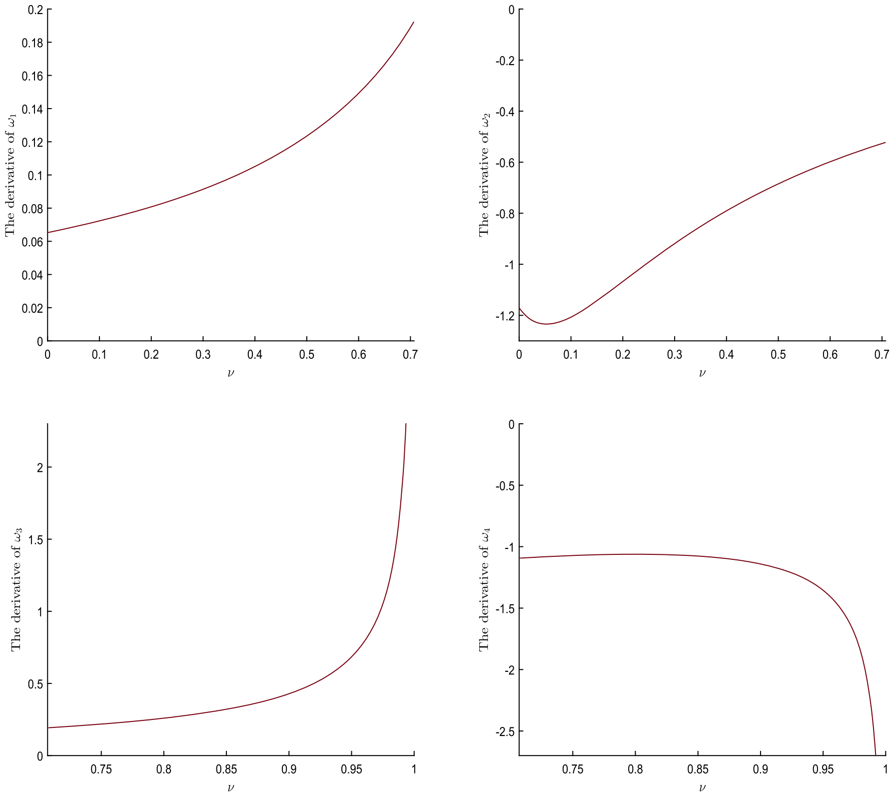

Fig. 1 Curves for derivatives of $\omega_i(\nu)$ with $i = 1, 2, 3, 4$

lent to determining the optimal parameter $\omega \in (0, 2)$ that minimizes the eigenvalue $\lambda_{\max}(H_\nu(\omega))$ defined as in (2.11) for a given $\nu \in (0, 1)$.

From (2.11), let

$$
\begin{aligned}
g_\nu(\omega) &= 3a^2 + 2c^2 + 2ac + \sqrt{(3a^2 + 2c^2 + 2ac)^2 - 4a^4} \\
&= 3(\omega-1)^2 + 2\nu^2\omega^4 + 2\nu\omega^2|\omega-1| + \sqrt{\left[3(\omega-1)^2 + 2\nu^2\omega^4 + 2\nu\omega^2|\omega-1|\right]^2 - 4(\omega-1)^4},
\end{aligned} \tag{3.2}
$$

then to minimize $\lambda_{\max}(H_\nu(\omega))$ is equivalent to minimizing $g_\nu(\omega)$. Notice that the function $g_\nu(\omega)$ in (3.2) is continuous but non-smooth with $0 < \nu < 1$ due to the non-differentiability of the absolute value function. Indeed, it is just not differentiable at $\omega = 1$. In addition, by simple calculation, we have

$$
g'_\nu(\omega) = \begin{cases} s_\nu(\omega) + \frac{r_\nu(\omega)s_\nu(\omega) - 8(\omega-1)^3}{\sqrt{[r_\nu(\omega)]^2 - 4(\omega-1)^4}} \doteq g_\nu^1(\omega), & \text{if } \ 0 < \omega < 1, \\ t_\nu(\omega) + \frac{\left[3(\omega-1)^2 + 2\nu^2\omega^4 + 2\nu\omega^2(\omega-1)\right]t_\nu(\omega) - 8(\omega-1)^3}{\sqrt{\left[3(\omega-1)^2 + 2\nu^2\omega^4 + 2\nu\omega^2(\omega-1)\right]^2 - 4(\omega-1)^4}} \doteq g_\nu^2(\omega), & \text{if } \ 1 < \omega < 2, \end{cases} \tag{3.3}
$$

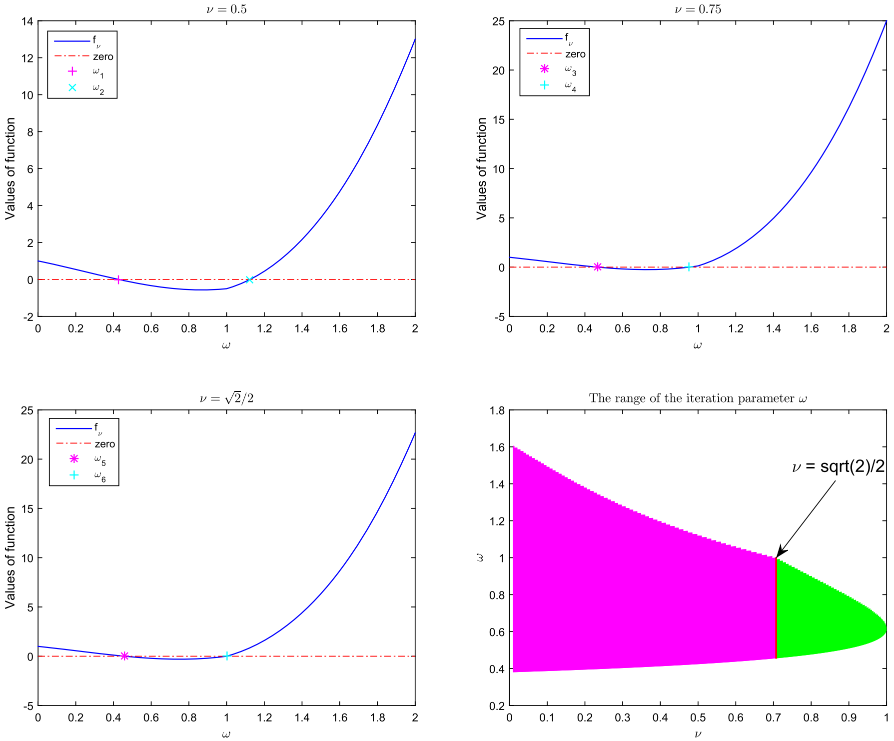

Fig. 2 Curves for function f_ν with $\nu = 0.5$ (top-left), $\nu = 0.75$ (top-right), $\nu = \frac{\sqrt{2}}{2}$ (bottom-left), and the range of the iteration parameter ω (bottom-right)

where

$$r_\nu(\omega) = 3\,(\omega - 1)^2 + 2\,\nu^2\omega^4 + 2\,\nu\omega^2\,(1 - \omega)\,,$$
$$s_\nu(\omega) = 6(\omega - 1) + 8\nu^2\omega^3 + 2\nu(2\omega - 3\omega^2),$$
$$t_\nu(\omega) = 6(\omega - 1) + 8\nu^2\omega^3 + 2\nu(3\omega^2 - 2\omega).$$

It is easy to check that $g_\nu^2(\omega) > 0$ for $1 < \omega < 2$ and $0 < \nu < 1$. Thus, for any $0 < \nu < 1$, $g_\nu(\omega)$ is strictly monotonously increasing when $1 < \omega < 2$. For any $0 < \nu < 1$, now we turn to consider $g_\nu^1(\omega)$ in (3.3) with $0 < \omega < 1$. By direct computation, we have

$$g_\nu^1(0) = -6 - 2\sqrt{5} < 0, \tag{3.4}$$
$$g_\nu^1(1) = 4\nu(4\nu - 1). \tag{3.5}$$

It follows from (3.5) that

$$g_\nu^1(1) \le 0, \quad \text{if} \quad 0 < \nu \le \frac{1}{4}, \tag{3.6}$$
$$g_\nu^1(1) > 0, \quad \text{if} \quad \frac{1}{4} < \nu < 1. \tag{3.7}$$

In addition, we have

$$(g_\nu^1)'(\omega) = s_\nu'(\omega) + \frac{g_\nu^{1u}(\omega)}{g_\nu^{1l}(\omega)} - \frac{g_\nu^{2u}(\omega)}{[g_\nu^{1l}(\omega)]^3}, \tag{3.8}$$

where

$$\begin{aligned}
g_\nu^{1u}(\omega) &= [s_\nu(\omega)]^2 + s_\nu'(\omega) r_\nu(\omega) - 24(\omega-1)^2,\\
g_\nu^{1l}(\omega) &= \sqrt{[r_\nu(\omega)]^2 - 4(\omega-1)^4},\\
g_\nu^{2u}(\omega) &= [r_\nu(\omega)s_\nu(\omega) - 8(\omega-1)^3]^2.
\end{aligned}$$

Next, we will rigorously prove that $(g_\nu^1)'(\omega) > 0$ for $\omega \in (0, 1)$ and $\nu \in (0, 1)$ by using methods of symbolic computation. For this, we write (3.8) in the following form:

$$[g_\nu^{1l}(\omega)]^3 (g_\nu^1)'(\omega) = [g_\nu^{1l}(\omega)]^3 s_\nu'(\omega) + G(\nu, \omega),$$

where

$$\begin{aligned}
G(\nu,\omega) &= [g_\nu^{1l}(\omega)]^2 g_\nu^{1u}(\omega) - g_\nu^{2u}(\omega)\\
&= 192\nu^8\omega^{14} - 672\nu^7\omega^{13} + 608\nu^7\omega^{12} + 1776\nu^6\omega^{12} - 3264\nu^6\omega^{11} + 1536\nu^6\omega^{10}\\
&\quad - 2784\nu^5\omega^{11} + 7584\nu^5\omega^{10} - 6960\nu^5\omega^9 + 3192\nu^4\omega^{10} + 2160\nu^5\omega^8 - 11344\nu^4\omega^9\\
&\quad + 15144\nu^4\omega^8 - 2440\nu^3\omega^9 - 9024\nu^4\omega^7 + 10584\nu^3\omega^8 + 2032\nu^4\omega^6 - 18288\nu^3\omega^7\\
&\quad + 1260\nu^2\omega^8 + 15760\nu^3\omega^6 - 6432\nu^2\omega^7 - 6792\nu^3\omega^5 + 13608\nu^2\omega^6 - 360\nu\omega^7\\
&\quad + 1176\nu^3\omega^4 - 15312\nu^2\omega^5 + 2040\nu\omega^6 + 9708\nu^2\omega^4 - 4860\nu\omega^5 + 50\omega^6\\
&\quad - 3312\nu^2\omega^3 + 6300\nu\omega^4 - 300\omega^5 + 480\nu^2\omega^2 - 4800\nu\omega^3 + 750\omega^4\\
&\quad + 2160\nu\omega^2 - 1000\omega^3 - 540\nu\omega + 750\omega^2 + 60\nu - 300\omega + 50.
\end{aligned}$$

It is obvious that $s_\nu'(\omega) = (3\nu\omega - 2)^2 + 15\nu^2\omega^2 + 4\nu + 2 > 0$ and $g_\nu^{1l}(\omega) > 0$ for $\omega \in (0, 1)$ and $\nu \in (0, 1)$. So, it is sufficient to show that $G(\nu, \omega) > 0$ for $\omega \in (0, 1)$ and $\nu \in (0, 1)$. Let $\bar{D} = [0, 1] \times [0, 1]$ be a compact set of (ν, ω). We want to find the minimum value of $G(\nu, \omega)$ on $\bar{D}$. First, we compute its gradient $\nabla G = [G_\nu, G_\omega]$, where

$$\begin{aligned}
G_\nu &= 1536\,\nu^7\omega^{14} - 4704\,\nu^6\omega^{13} + 4256\,\nu^6\omega^{12} + 10656\,\nu^5\omega^{12} - 19584\,\nu^5\omega^{11} + 9216\,\nu^5\omega^{10}\\
&\quad - 13920\,\nu^4\omega^{11} + 37920\,\nu^4\omega^{10} - 34800\,\nu^4\omega^9 + 12768\,\nu^3\omega^{10} + 10800\,\nu^4\omega^8 - 45376\,\nu^3\omega^9\\
&\quad + 60576\,\nu^3\omega^8 - 7320\,\nu^2\omega^9 - 36096\,\nu^3\omega^7 + 31752\,\nu^2\omega^8 + 8128\,\nu^3\omega^6 - 54864\,\nu^2\omega^7\\
&\quad + 2520\,\nu\,\omega^8 + 47280\,\nu^2\omega^6 - 12864\,\nu\,\omega^7 - 20376\,\nu^2\omega^5 + 27216\,\nu\,\omega^6 - 360\,\omega^7 + 3528\,\nu^2\omega^4\\
&\quad - 30624\,\nu\,\omega^5 + 2040\,\omega^6 + 19416\,\nu\,\omega^4 - 4860\,\omega^5 - 6624\,\nu\,\omega^3 + 6300\,\omega^4 + 960\,\nu\,\omega^2\\
&\quad - 4800\,\omega^3 + 2160\,\omega^2 - 540\,\omega + 60,
\end{aligned}$$

$$\begin{aligned} G_\omega = {} & 2688\,\nu^8\omega^{13} - 8736\,\nu^7\omega^{12} + 7296\,\nu^7\omega^{11} + 21312\,\nu^6\omega^{11} - 35904\,\nu^6\omega^{10} + 15360\,\nu^6\omega^9 \\ & - 30624\,\nu^5\omega^{10} + 75840\,\nu^5\omega^9 - 62640\,\nu^5\omega^8 + 31920\,\nu^4\omega^9 + 17280\,\nu^5\omega^7 - 102096\,\nu^4\omega^8 \\ & + 121152\,\nu^4\omega^7 - 21960\,\nu^3\omega^8 - 63168\,\nu^4\omega^6 + 84672\,\nu^3\omega^7 + 12192\,\nu^4\omega^5 - 128016\,\nu^3\omega^6 \\ & + 10080\,\nu^2\omega^7 + 94560\,\nu^3\omega^5 - 45024\,\nu^2\omega^6 - 33960\,\nu^3\omega^4 + 81648\,\nu^2\omega^5 - 2520\,\nu\,\omega^6 \\ & + 4704\,\nu^3\omega^3 - 76560\,\nu^2\omega^4 + 12240\,\nu\,\omega^5 + 38832\,\nu^2\omega^3 - 24300\,\nu\,\omega^4 + 300\,\omega^5 \\ & - 9936\,\nu^2\omega^2 + 25200\,\nu\,\omega^3 - 1500\,\omega^4 + 960\,\nu^2\omega - 14400\,\nu\,\omega^2 + 3000\,\omega^3 + 4320\,\nu\,\omega \\ & - 3000\,\omega^2 - 540\,\nu + 1500\,\omega - 300. \end{aligned}$$

To determine where $\nabla G = 0$, we apply the Maple bulit-in program directly to search all the common roots of G_ν and G_ω on $\bar{D}$.

> with(RegularChains):
> with(ChainTools):
> with(SemiAlgebraicSetTools):
> R:= PolynomialRing([ν, ω]):
> L:= RealRootIsolate([G_ν, G_ω], [ν, $1-\nu$, ω, $1-\omega$], [ ], [ ], R, 'abserr' = $1/2^5$):
> map(BoxValues, L, R);

$$[[\nu = 0, \omega = 1]].$$

We see that there is only one critical point of $G(\nu, \omega)$, at $(\nu, \omega) = (0, 1)$. Hence, $G(\nu, \omega)$ reaches its minimum value at the boundaries of $\bar{D}$ (see, e.g., Section 14.7 in [29]). Direct computation gives

$$\min_{\bar{D}} G(\nu, \omega) = \min\{G(0,0), G(0,1), G(1,0), G(1,1)\} = 0.$$

Then, we prove that $G(\nu, \omega) > 0$ for $\omega \in (0, 1)$ and $\nu \in (0, 1)$. As a result, $(g_\nu^1)'(\omega) > 0$ holds for any $\omega \in (0, 1)$ and $\nu \in (0, 1)$.

Combining the above fact (which means that $g_\nu^1(\omega)$ is strictly monotonously increasing in $\omega \in (0, 1)$) with (3.4), (3.6), and (3.7), we can conclude that $g_\nu^1(\omega)$ has at most one root whenever $0 < \omega < 1$. If the root exists (it happens when $\frac{1}{4} < \nu < 1$), which is denoted by ω_{opt}, then $g_\nu(\omega)$ is strictly monotonously decreasing in $(0, \omega_{opt})$ and strictly monotonously increasing in $(\omega_{opt}, 2)$, and the optimal iteration parameter which minimizes $\|T_\nu(\omega)\|_2$ is $\omega^*_{opt} = \omega_{opt}$. Otherwise, if $0 < \nu \le \frac{1}{4}$, then $g_\nu(\omega)$ is strictly monotonously decreasing in $(0, 1)$ and strictly monotonously increasing in $(1, 2)$, and the optimal iteration parameter which minimizes $\|T_\nu(\omega)\|_2$ is $\omega^*_{opt} = 1$. In conclusion, the optimal iteration parameter is

$$\omega^*_{opt} = \begin{cases} \omega_{opt}, & \text{if } \frac{1}{4} < \nu < 1, \\ 1, & \text{if } 0 < \nu \le \frac{1}{4}. \end{cases} \tag{3.9}$$

When $\frac{1}{4} < \nu < 1$, the root of $g_\nu^1(\omega)$ located in $(0, 1)$, ω_{opt}, can be efficiently numerically calculated by the classical bisection method; see Algorithm 2 for more

Algorithm 2 The bisection method [2, Pages 72-73]

Require: a continuous function f on the closed interval $[s, t]$ satisfied $f(s)f(t) < 0$. Let $\varepsilon > 0$ be the tolerance.
Ensure: a root of f on (s, t).
repeat
  define $r = (s + t)/2$.
  if $t - r \le \varepsilon$ and $|f(r)| \le \varepsilon$, then accept r as the root and stop.
  if $\mathrm{sgn}(f(t)) \cdot \mathrm{sgn}(f(r)) \le 0$, then set $s = r$. Otherwise, set $t = r$.
until convergence;
return the last r as the approximation to the root of f on (s, t).

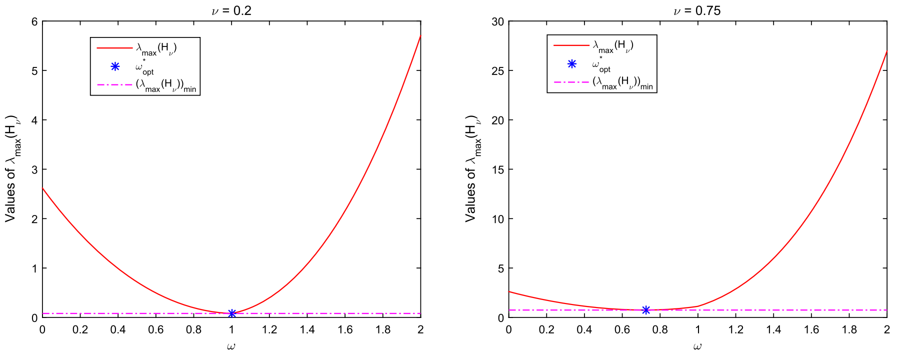

Fig. 3 Curves of $\lambda_{\max}(H_\nu(\omega))$

details.[4] Figure 3 shows some curves of $\lambda_{\max}(H_\nu(\omega))$ which visually demonstrates our claims mentioned above.

3.2 Approximate optimal iteration parameter

In this subsection, we devote to discussing the approximate optimal iteration parameter ω^*_{aopt} which will minimize $\eta_\nu(\omega) = \max\{|1 - \omega| \doteq h(\omega), \omega^2\nu \doteq h_\nu(\omega)\}$ with $\omega \in (0, 2)$.

This $\eta_\nu(\omega)$ appears in the proof of Corollary 3.1 in [15] and satisfies $\|T_\nu(\omega)\|_2 \le \frac{\eta_\nu(\omega)}{\tau}$. That is, $\frac{\eta_\nu(\omega)}{\tau}$ is an upper bound of $\|T_\nu(\omega)\|_2$ with $\omega \in (0, 2)$, which is the reason that we call ω^*_{aopt} the approximate optimal iteration parameter. Since $h(\omega)$ is strictly monotonously decreasing in $(0, 1)$ and strictly monotonously increasing in $(1, 2)$ and $h_\nu(\omega)$ is strictly monotonously increasing in $(0, 2)$ with $\nu \in (0, 1)$, ω^*_{aopt} must satisfy $1 - \omega = \nu\omega^2 (0 < \omega < 2)$, that is,

$$\omega^*_{aopt}(\nu) = \frac{\sqrt{4\nu + 1} - 1}{2\nu}. \tag{3.10}$$

Figure 4 intuitively confirms this result.

[4] In our numerical experiments, we set $s = 0$, $t = 1$, and $\varepsilon = 10^{-10}$.

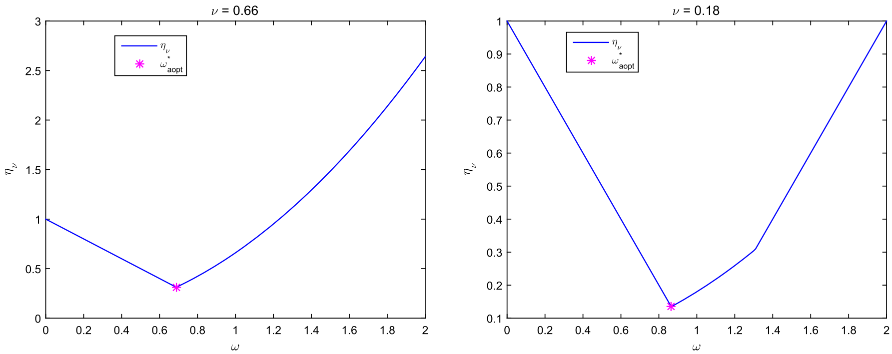

Fig. 4 Curves of $\eta_\nu(\omega)$

3.3 Some discussions on the optimal and approximate optimal iteration parameters

Thanks to (3.10), the approximate optimal iteration parameter $\omega^*_{aopt}(\nu)$ is strictly monotonously decreasing with respect to $\nu \in (0, 1)$ and it approaches to 1 (the optimal one when $0 < \nu \le \frac{1}{4}$) as $\nu \to 0^+$ and closes to $\frac{\sqrt{5}-1}{2}$ as $\nu \to 1^-$. Thus, $\omega^*_{aopt}(\nu)$ must locate in the range of ω determined by conditions (2.24)–(2.26) for $\nu \in (0, 1)$. On the other hand, since the optimal iteration parameter ω^*_{opt} must satisfy the conditions (2.24)–(2.26), it also locates in the range of ω determined by conditions (2.24)–(2.26) and it closes to $\frac{\sqrt{5}-1}{2}$ as $\nu \to 1^-$. Furthermore, from the definition of the optimal iteration parameter, we have $\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))} \le \sqrt{\lambda_{\max}(H_\nu(\omega^*_{aopt}))}$,[5] which implies that the SOR-like iteration method with optimal iteration parameter (3.9) will converge no slower than that with the approximate optimal iteration parameter (3.10). In addition, we have $\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))} \le \frac{\eta_\nu(\omega^*_{aopt})}{\tau}$. Indeed, when $\nu \in (0, \frac{1}{4})$, we have $\omega^*_{aopt} < 1 = \omega^*_{opt}$, and we can conclude that $\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))} < \sqrt{\lambda_{\max}(H_\nu(\omega^*_{aopt}))}$. When $\nu \in (\frac{1}{4}, 1)$, the analytical formulation of ω^*_{opt} is unknown, and, numerically, we also find that $\omega^*_{aopt} < \omega^*_{opt}$ whenever ω^*_{opt} is obtained. Thus, $\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))} < \sqrt{\lambda_{\max}(H_\nu(\omega^*_{aopt}))}$ for $\nu \in (0, 1)$, which means that the SOR-like iteration method with optimal iteration parameter (3.9) converges faster than that with the approximate optimal iteration parameter (3.10). As a by-product, we have $\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))} < \frac{\eta_\nu(\omega^*_{aopt})}{\tau}$ for $\nu \in (0, 1)$. Figure 5 intuitively clarifies our arguments here. Moreover, our numerical results in the next section will further demonstrate these arguments.

Remark 3.1 In [9], the authors give some convergence conditions from the involved iteration matrix of the SOR-like iteration method for solving the AVE (1.1), which are

[5] According to the property of $\sqrt{\lambda_{\max}(H_\nu(\omega))}$, the equality holds if and only if $\omega^*_{opt} = \omega^*_{aopt}$.

different from the results of [15]. In addition, if all the eigenvalues of $D(x^{(k+1)})A^{-1}$ are positive, then the optimal iteration parameter which minimizes the spectral radius of the iteration matrix is $\omega_o^* = \frac{2}{1+\sqrt{1-\varrho}}$ with $\varrho = \rho(D(x^{(k+1)})A^{-1})$ [9]. Apparently, the optimal iteration parameter is iteration-dependent and it will be expensive to compute it in every iterative step, and the authors use the approximate optimal one $\omega_o = \frac{2}{1+\sqrt{1-\tilde{\varrho}}}$ with $\tilde{\varrho} = \rho(A^{-1})$ in their numerical experiments. Though the approximate optimal parameter is iteration-independent, however, as it will be shown in the next section, ω_o will be out of the range of ω determined by (2.24)–(2.26), which means that the SOR-like method with this ω_o may be divergent (see Example 4.2 for more details).

Remark 3.2 From our analysis above, ω_{opt}^* is the optimal iteration parameter which minimizes $\|T_\nu(\omega)\|_2$, an upper bound of the linear convergence factor for the SOR-like iteration method in the metric $|||\cdot|||_\omega$ of (e_k^x, e_k^y) (see (2.9)). However, it may not be the really optimal one since the upper bound may not be tight; see the numerical results of Example 4.2 for an instance.

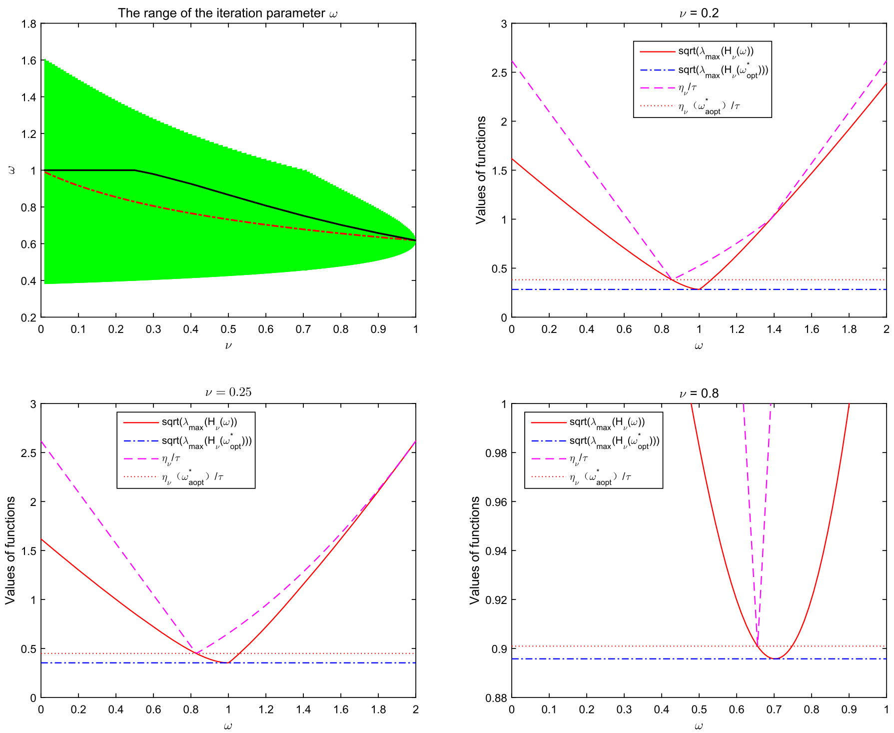

Fig. 5 Curves of optimal (top left: black solid line) and approximate optimal (top left: red dots line) iteration parameters and the comparisons of minimal points and minimums (others)

4 Numerical experiments

In this section, we will present two numerical examples to illustrate the superior performance of the SOR-like iteration method with our optimal iteration parameter for solving AVE (1.1). Four algorithms will be tested.

1. SORLo: the SOR-like iteration method with the approximate optimal iteration parameter $\omega_o = \frac{2}{1+\sqrt{1-\tilde{\varrho}}}$ proposed in [9]. Here, $\tilde{\varrho} = \rho(A^{-1})$.
2. SORLaopt: the SOR-like iteration method with the approximate optimal iteration parameter ω^*_{aopt} defined as in (3.10).
3. SORLopt: the SOR-like iteration method with the optimal iteration parameter ω^*_{opt} defined as in (3.9).
4. SORLno: the SOR-like iteration method with the numerically optimal iteration parameter ω_{no}, which is selected from $\omega = [0.001 : 0.001 : 1.999]$ and is the first one to reach the minimal number of iteration of the method. Practically, determining optimal iteration parameter in this way is unwise because of that it is always very time consuming, especially for large scale problems (see the following numerical results for detail). However, it is still necessary to be tested here in order to demonstrate the claim in Remark 3.2.

Note that in all these algorithms, the main task per iteration is solving a system of linear equations. In this paper, the tested methods are implemented in conjunction with the Cholesky factorization since the coefficient matrix is symmetric positive definite. Specifically, we use $dA = \textbf{decomposition}(A, \text{'chol'})$ to generate the Cholesky decomposition of A, where **decomposition** is the routine in MATLAB, which returns the corresponding decomposition of a matrix A that can be used to solve the linear system $Ax = b$ efficiently. The call $x = dA\backslash b$ returns the same vector as $A\backslash b$, but is typically faster.

In the numerical results, we will report “IT" (the number of iteration), “CPU" (the elapsed CPU time in seconds), and “RES" (the residual error). RES is defined by

$$\text{RES} = \|Ax^{(k)} - |x^{(k)}| - b\|.$$

All tests are started from the initial zero vector and terminated if the current iteration satisfies RES $\leq 10^{-8}$ (except the last problem in Example 4.2) or the number of prescribed maximal iteration steps $k_{\max} = 100$ is exceeded (denoted by “–”). In order to obtain more accurate CPU time, we run all test problems in each method five times and take the average. All computations are done in MATLAB R2017b with a machine precision 2.22×10^{-16} on a personal computer with 2.60GHz central processing unit (Intel Core i7), 16GB memory and MacOS operating system.

Table 1 Parameters for Example 4.1

m	ν	ω_o	ω_{no}	ω^*_{aopt}	ω^*_{opt}	$\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))}$	$\frac{\eta(\omega^*_{aopt})}{\tau}$	Ω_1
8	0.2358	1.0671	0.991	0.8354	1	0.3335	0.4309	(0.3994, 1.3447)
16	0.2458	1.0704	0.984	0.8305	1	0.3476	0.4438	(0.4003, 1.3347)
32	0.2489	1.0714	0.992	0.8290	1	0.3520	0.4478	(0.4005, 1.3316)
64	0.2497	1.0717	0.983	0.8286	1	0.3531	0.4488	(0.4006, 1.3308)

Example 4.1 ([9]) Consider the AVE (1.1) with

$$A = \mathrm{Tridiag}(-I_m, S_m, -I_m) = \begin{bmatrix} S_m & -I_m & 0 & \cdots & 0 & 0 \\ -I_m & S_m & -I_m & \cdots & 0 & 0 \\ 0 & -I_m & S_m & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & S_m & -I_m \\ 0 & 0 & 0 & \cdots & -I_m & S_m \end{bmatrix} \in \mathbb{R}^{n\times n},$$

$$S_m = \mathrm{tridiag}(-1, 8, -1) = \begin{bmatrix} 8 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 8 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 8 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 8 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 8 \end{bmatrix} \in \mathbb{R}^{m\times m}$$

and $b = Ax^* - |x^*|$, where $x^* = [-1, 1, -1, 1, \cdots, -1, 1]^T \in \mathbb{R}^n$. Here, we have $n = m^2$.

The parameters for this example are displayed in Table 1, from which we can find that $\omega^*_{aopt} < \omega^*_{opt}, \omega_o, \omega_{no}, \omega^*_{aopt}, \omega^*_{opt} \in \Omega_1$, and the larger the value of ν is, the smaller the range of ω is. Meanwhile, $\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))} < \frac{\eta(\omega^*_{aopt})}{\tau} < 1$, which implies that SORLopt will converge faster than SORLaopt.

Numerical results for this example are reported in Table 2. From Table 2, we find that SORLopt always performs better than SORLo and SORLaopt in terms of IT and CPU. SORLopt and SORLno take the same number of iteration, but SORLopt is better

Table 2 Numerical results for Example 4.1

Method		m			
		8	16	32	64
SORLo	IT	20	21	22	22
	CPU	0.0027	0.0049	0.0136	0.0518
	RES	3.1565×10^{-9}	4.0382×10^{-9}	3.3876×10^{-9}	7.4133×10^{-9}
SORLaopt	IT	23	24	25	26
	CPU	0.0034	0.0054	0.0135	0.0534
	RES	4.2462×10^{-9}	5.3287×10^{-9}	4.9212×10^{-9}	4.0809×10^{-9}
SORLopt	IT	13	14	14	15
	CPU	0.0022	0.0044	0.0118	0.0507
	RES	4.9340×10^{-9}	3.0140×10^{-9}	6.7716×10^{-9}	1.8270×10^{-9}
SORLno	IT	13	14	14	15
	CPU	1.4552	2.1608	5.2270	20.0681
	RES	9.0770×10^{-9}	9.4222×10^{-9}	9.2093×10^{-9}	9.7421×10^{-9}

than SORLno in terms of CPU time and the reason is that determining the numerically optimal iteration parameter needs more CPU than computing the optimal iteration parameter. In addition, SORLo is better than SORLaopt in terms of IT and CPU.

Example 4.2 ([7, 15]) Consider the AVE (1.1) with the matrix $A \in \mathbb{R}^{n\times n}$ arises from six different test problems listed in Table 3. These matrices are sparse and symmetric positive definite and $\|A^{-1}\| < 1$. In addition, let $b = Ax^* - |x^*|$ with $x^* = [-1, 1, -1, 1, \cdots, -1, 1]^T \in \mathbb{R}^n$.

Table 4 displays the parameters for Example 4.2. It follows from Table 4 that $\omega^*_{aopt} < \omega^*_{opt}$, $\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))} < \frac{\eta(\omega^*_{aopt})}{\tau} < 1$, and the range of ω becomes smaller as the value of ν becomes larger. Furthermore, all values of ω_{no}, ω^*_{aopt}, and ω^*_{opt} belong to Ω_1 or Ω_2, while ω_o does not belong to Ω_1 or Ω_2 for the former three test problems and it belongs to Ω_1 for the later three test problems.

Table 5 reports the numerical results for Example 4.2. From Table 5, for the former three test problems, SORLo is divergent and the reason is that ω_o does not belong to Ω_1 or Ω_2 in these cases. For the test problems $Trefethen\_20b$ and $Trefethen\_200b$, all methods converge within 100 numbers of iteration. For the $Trefethen\_20000b$ problem, RES of all methods do not reach about $O(10^{-9})$ within 100 numbers of iteration (Indeed, it is still the case if we set the maximal iteration number to 2000 or more). However, for the test problem $Trefethen\_20000b$, if we set RES $\leq 10^{-6}$ as the new stopping criterion, then SORLo, SORLaopt, SORLopt, and SORLno converge at iterative steps 53, 22, 14, and 12, respectively, and the elapsed CPU time are 12.4365, 9.4597, *8.6735*, and 13575.4682, respectively. The RES for SORLo, SORLaopt, SORLopt, and SORLno are 7.8509×10^{-7}, 6.7046×10^{-7}, 4.7848×10^{-7}, and 9.6870×10^{-7}, respectively.

In conclusion, for the tested problems in this example, we find that SORLopt is always better than SORLo and SORLaopt in terms of IT and CPU; SORLno is better than SORLopt with respect to IT (see the possible reason in Remark 3.2) but behaves worst in terms of CPU; SORLaopt is better than SORLo in terms of IT and CPU.

The numerical results of this section allow us to conclude that SORLopt is better than SORLo and SORLaopt in terms of IT and CPU. SORLopt is superior to SORLno in terms of CPU. However, SORLno is no worse than SORLopt in terms of IT, which demonstrates that our optimal iteration parameter is not exactly the optimal one; it is optimal in the sense that it minimizes $\|T_\nu(\omega)\|_2$, an upper bound of the linear convergence factor of the SOR-like iteration method in the metric $\||\cdot|\|_\omega$ for (e^x_k, e^y_k). Nevertheless, SORLopt has performed pretty good. In addition, SORLaopt is better than SORLo in some situations.

Table 3 Test problems for Example 4.2

Problem	n	Problem	n
$mesh1e1$	48	$Trefethen\_20b$	19
$mesh1em1$	48	$Trefethen\_200b$	199
$mesh2e1$	306	$Trefethen\_20000b$	19999

Table 4 Parameters for Example 4.2

Problem	ν	ω_o	ω_{no}	ω^*_{aopt}	ω^*_{opt}	$\sqrt{\lambda_{\max}(H_\nu(\omega^*_{opt}))}$	$\frac{\eta(\omega^*_{aopt})}{\tau}$	Range
$mesh1e1$	0.5747	1.2105	0.946	0.7102	0.8218	0.7301	0.7588	$\Omega_1 = (0.4361, 1.0753)$
$mesh1em1$	0.6397	1.2498	0.921	0.6929	0.7848	0.7845	0.8040	$\Omega_1 = (0.4460, 1.0367)$
$mesh2e1$	0.7615	1.3438	0.944	0.6641	0.7210	0.8717	0.8793	$\Omega_2 = (0.4692, 0.9413)$
$Trefethen\_20b$	0.4244	1.1372	0.939	0.7569	0.9114	0.5783	0.6365	$\Omega_1 = (0.4175, 1.1785)$
$Trefethen\_200b$	0.4265	1.1381	0.940	0.7561	0.9102	0.5807	0.6384	$\Omega_1 = (0.4177, 1.1769)$
$Trefethen\_20000b$	0.4268	1.1382	0.943	0.7561	0.9101	0.5810	0.6387	$\Omega_1 = (0.4177, 1.1767)$

Table 5 Numerical results for Example 4.2 (RES $\leq 10^{-8}$)

Problem		Method $SORLo$	$SORLaopt$	$SORLopt$	$SORLno$
mesh1e1	IT	–	42	32	23
	CPU	–	0.0047	0.0039	1.6561
	RES	–	9.5347×10^{-9}	9.2561×10^{-9}	9.5398×10^{-9}
mesh1em1	IT	–	43	35	24
	CPU	–	0.0064	0.0057	1.6590
	RES	–	8.5392×10^{-9}	5.9469×10^{-9}	9.7687×10^{-9}
mesh2e1	IT	–	53	46	26
	CPU	–	0.0252	0.0250	3.0330
	RES	–	7.1097×10^{-9}	7.6843×10^{-9}	9.9663×10^{-9}
Trefethen_20b	IT	68	27	18	16
	CPU	0.0036	0.0028	0.0013	1.3719
	RES	7.9973×10^{-9}	8.2040×10^{-9}	5.6398×10^{-9}	9.9390×10^{-9}
Trefethen_200b	IT	69	27	18	16
	CPU	0.0141	0.0114	0.0102	3.8434
	RES	9.2407×10^{-9}	8.6263×10^{-9}	6.1932×10^{-9}	9.4897×10^{-9}
Trefethen_20000b	IT	–	–	–	–
	CPU	–	–	–	–
	RES	–	–	–	–

5 Conclusions

In this paper, by revisiting the convergence conditions of the SOR-like iteration method proposed in [15] for solving the AVE (1.1), the convergent range of the iteration parameter is given and optimal and approximate optimal iteration parameters for the SOR-like iteration method are determined. Our analysis is from the view of the iteration error, which is different from that of iteration matrix [9]. Furthermore, our optimal and approximate optimal iteration parameters are iteration-independent. Numerical results are provided to illustrate that the SOR-like iteration method with our optimal iteration parameter converges faster than that with the approximate optimal parameter proposed in [9] on solving the AVE (1.1) with $\|A^{-1}\|_2 < 1$. In addition, our approximate optimal parameter behaves better than that of [9] in some situations.

Appendix A. The functions $\omega_i(\nu)(i = 1,\ 2,\ 3,\ 4)$ are real-valued

Recall that a real-valued function is a function that assigns a real number to each member of its domain. Our main result in this section is the next theorem.

Theorem 5.1 *The functions $\omega_i(\nu)$ $(i = 1,2,3,4)$ defined as in (2.18)-(2.21) are real-valued functions.*

Proof Firstly, we consider $\omega_1(\nu)$. According to (2.17) and (2.18), we have

$$\begin{aligned}&\omega_1(\nu)\\&=\frac{\nu-2+\sqrt{(1-\nu)(\nu+5)}-\sqrt{-(8\nu^3-16\nu^2+4\nu-1)-(8\nu^2-2\nu)\sqrt{(1-\nu)(\nu+5)}}}{2(2\nu^2-1)},\\&\nu\in(0,\frac{\sqrt{2}}{2}).\end{aligned}$$

Since $\nu \in (0, \frac{\sqrt{2}}{2})$, $(1-\nu)(\nu+5) > 0$ and thus $\sqrt{(1-\nu)(\nu+5)}$ is real-valued. Let

$$\Delta_1(\nu) = -(8\nu^3-16\nu^2+4\nu-1)-(8\nu^2-2\nu)\sqrt{(1-\nu)(\nu+5)} \tag{5.1}$$

$$= -(8\nu^3-16\nu^2+4\nu-1)+\frac{8\nu^4+30\nu^3-48\nu^2+10\nu}{\sqrt{(1-\nu)(\nu+5)}}. \tag{5.2}$$

Then $\omega_1(\nu)$ is real-valued in $(0, \frac{\sqrt{2}}{2})$ if $\Delta_1(\nu) > 0$ whenever $\nu \in (0, \frac{\sqrt{2}}{2})$. According to (5.2), $\Delta_1(\nu) > 0$ is equivalent to

$$\begin{aligned}&p_1(\nu) \doteq 8\nu^4+30\nu^3-48\nu^2+10\nu > (8\nu^3-16\nu^2+4\nu-1)\sqrt{(1-\nu)(\nu+5)}\\&\doteq p_2(\nu)\sqrt{(1-\nu)(\nu+5)}.\end{aligned} \tag{5.3}$$

Since $p_2'(\nu) = 4(6\nu^2 - 8\nu + 1) > 0$ in $(0, \frac{4-\sqrt{10}}{6})$ and $p_2'(\nu) < 0$ in $(\frac{4-\sqrt{10}}{6}, \frac{\sqrt{2}}{2})$, $p_2(\nu)$ is strictly monotonously increasing in $(0, \frac{4-\sqrt{10}}{6})$ and strictly monotonously decreasing in $(\frac{4-\sqrt{10}}{6}, \frac{\sqrt{2}}{2})$. Thus, for any $\nu \in (0, \frac{\sqrt{2}}{2})$, $p_2(\nu) \le p_2(\frac{4-\sqrt{10}}{6}) = \frac{-83+20\sqrt{10}}{27} < 0$. On the other hand, $p_1(\nu) = \nu(4\nu - 1)(2\nu + 10)(\nu - 1) \ge 0$ when $0 < \nu \le \frac{1}{4}$ and $p_1(\nu) < 0$ when $\frac{1}{4} < \nu < \frac{\sqrt{2}}{2}$. Thus, from (5.3), we have $\Delta_1(\nu) > 0$ when $0 < \nu \le \frac{1}{4}$. Now we turn to the case that $\frac{1}{4} < \nu < \frac{\sqrt{2}}{2}$. In this case, (5.3) is equivalent to

$$(8\nu^3 - 16\nu^2 + 4\nu - 1)^2(1 - \nu)(\nu + 5) > (8\nu^4 + 30\nu^3 - 48\nu^2 + 10\nu)^2,$$

which is simplified as

$$\begin{aligned} p_3(\nu) &\doteq (2\nu^2 - 1)^2(\nu + 5)(\nu - 1)(-32\nu^2 + 8\nu - 1) \\ &\doteq (2\nu^2 - 1)^2(\nu + 5)(\nu - 1)p_4(\nu) > 0. \end{aligned}$$

Since $p_4(\nu)$ is strictly monotonously decreasing in $(\frac{1}{4}, \frac{\sqrt{2}}{2})$, $p_4(\nu) < p_4(\frac{1}{4}) = -1 < 0$ with $\nu \in (\frac{1}{4}, \frac{\sqrt{2}}{2})$. Then we have $p_3(\nu) > 0$ since $(2\nu^2 - 1)^2(\nu + 5)(\nu - 1) < 0$ for $\nu \in (\frac{1}{4}, \frac{\sqrt{2}}{2})$. Hence, if $0 < \nu < \frac{\sqrt{2}}{2}$, it holds that $\Delta_1(\nu) > 0$ and thus $\omega_1(\nu)$ is a real-valued function in $(0, \frac{\sqrt{2}}{2})$.

Secondly, we turn to $\omega_2(\nu)$. According to (2.17) and (2.19), we have

$$\omega_2(\nu) = \frac{-\nu - 2 - \sqrt{(\nu+1)(5-\nu)} + \sqrt{(8\nu^3+16\nu^2+4\nu+1) + (8\nu^2+2\nu)\sqrt{(\nu+1)(5-\nu)}}}{2(2\nu^2-1)}, \quad \nu \in (0, \frac{\sqrt{2}}{2}).$$

Since $\nu \in (0, \frac{\sqrt{2}}{2})$, $(\nu + 1)(5 - \nu) > 0$ and thus $\sqrt{(\nu + 1)(5 - \nu)}$ is real-valued. Let

$$\begin{aligned} \Delta_2(\nu) &= (8\nu^3 + 16\nu^2 + 4\nu + 1) + (8\nu^2 + 2\nu)\sqrt{(\nu + 1)(5 - \nu)} \\ &= (8\nu^3 + 16\nu^2 + 4\nu + 1) - \frac{8\nu^4 - 30\nu^3 - 48\nu^2 - 10\nu}{\sqrt{(\nu + 1)(5 - \nu)}}. \end{aligned} \tag{5.4}$$

Then $\omega_2(\nu)$ is real-valued in $(0, \frac{\sqrt{2}}{2})$ if $\Delta_2(\nu) > 0$ whenever $\nu \in (0, \frac{\sqrt{2}}{2})$. According to (5.4), $\Delta_2(\nu) > 0$ is equivalent to

$$p_5(\nu) \doteq 8\nu^4 - 30\nu^3 - 48\nu^2 - 10\nu < (8\nu^3 + 16\nu^2 + 4\nu + 1)\sqrt{(\nu + 1)(5 - \nu)} \doteq p_6(\nu)\sqrt{(\nu + 1)(5 - \nu)}. \tag{5.5}$$

Since $p_5(\nu) = \nu(\nu - 5)(\nu + 1)(8\nu + 2)$, $p_5(\nu) < 0$ for $\nu \in (0, \frac{\sqrt{2}}{2})$. In addition, $p_6(\nu)\sqrt{(\nu + 1)(5 - \nu)} > 0$ for $\nu \in (0, \frac{\sqrt{2}}{2})$. Hence, (5.5) holds for $\nu \in (0, \frac{\sqrt{2}}{2})$ and thus $\omega_2(\nu)$ is a real-valued function in $(0, \frac{\sqrt{2}}{2})$.

Thirdly, we consider $\omega_3(\nu)$. According to (2.17), (2.20) and (5.1), we have

$$\omega_3(\nu) = \frac{\nu - 2 + \sqrt{(1-\nu)(\nu+5)} + \sqrt{\Delta_1(\nu)}}{2(2\nu^2-1)}, \quad \nu \in (\frac{\sqrt{2}}{2}, 1).$$

Since $\nu \in (\frac{\sqrt{2}}{2}, 1)$, $(1-\nu)(\nu+5) > 0$ and thus $\sqrt{(1-\nu)(\nu+5)}$ is real-valued. Then $\omega_3(\nu)$ is real-valued in $(\frac{\sqrt{2}}{2}, 1)$ if $\Delta_1(\nu) > 0$ whenever $\nu \in (\frac{\sqrt{2}}{2}, 1)$. Recall (5.3), $p_2(\nu)$ is strictly monotonously decreasing in $(\frac{\sqrt{2}}{2}, 1)$ since $p_2'(\nu) = 4(6\nu^2 - 8\nu + 1) < 0$ in $(\frac{\sqrt{2}}{2}, 1)$. Thus $p_2(\nu) < p_2(\frac{\sqrt{2}}{2}) = 4\sqrt{2} - 9 < 0$ for $\nu \in (\frac{\sqrt{2}}{2}, 1)$. On the other hand, $p_1(\nu) = \nu(4\nu-1)(2\nu+10)(\nu-1)$ implies that $p_1(\nu) < 0$ whenever $\nu \in (\frac{\sqrt{2}}{2}, 1)$. Hence, (5.3) is equivalent to

$$p_7(\nu) \doteq -128\nu^8 - 480\nu^7 + 892\nu^6 + 304\nu^5 - 776\nu^4 + 56\nu^3 + 171\nu^2 - 44\nu + 5 > 0.$$

Since

$$\begin{aligned} p_7(\nu) &= (2\nu^2-1)^2(\nu+5)(\nu-1)(-38\nu^2+8\nu-1) \\ &\doteq (2\nu^2-1)^2(\nu+5)(\nu-1)p_8(\nu) \end{aligned}$$

and $(2\nu^2-1)^2(\nu+5)(\nu-1) < 0$ for $\nu \in (\frac{\sqrt{2}}{2}, 1)$, it is sufficient to prove $p_8(\nu) < 0$ for $\nu \in (\frac{\sqrt{2}}{2}, 1)$. Indeed, $p_8'(\nu) = -76\nu + 8 < 0$ implies that $p_8(\nu) < p_8(\frac{\sqrt{2}}{2}) = -20 + 4\sqrt{2} < 0$. This completes the proof that $\omega_3(\nu)$ is a real-valued function in $(\frac{\sqrt{2}}{2}, 1)$.

Finally, we complete the proof by showing that $\omega_4(\nu)$ is a real-valued function in $(\frac{\sqrt{2}}{2}, 1)$. According to (2.17) and (2.21), we have

$$\omega_4(\nu) = \frac{\nu - 2 - \sqrt{(1-\nu)(5+\nu)} + \sqrt{-(8\nu^3 - 16\nu^2 + 4\nu - 1) + (8\nu^2 - 2\nu)\sqrt{(1-\nu)(5+\nu)}}}{2(2\nu^2-1)}, \quad \nu \in (\frac{\sqrt{2}}{2}, 1).$$

Since $\nu \in (\frac{\sqrt{2}}{2}, 1)$, $(1-\nu)(\nu+5) > 0$ and thus $\sqrt{(1-\nu)(\nu+5)}$ is real-valued. Let

$$\begin{aligned} \Delta_3(\nu) &= -(8\nu^3 - 16\nu^2 + 4\nu - 1) + (8\nu^2 - 2\nu)\sqrt{(1-\nu)(5+\nu)} \\ &= -(8\nu^3 - 16\nu^2 + 4\nu - 1) - \frac{8\nu^4 + 30\nu^3 - 48\nu^2 + 10\nu}{\sqrt{(1-\nu)(\nu+5)}}. \end{aligned}$$

Then $\omega_4(\nu)$ is real-valued in $(\frac{\sqrt{2}}{2}, 1)$ if $\Delta_3(\nu) > 0$ whenever $\nu \in (\frac{\sqrt{2}}{2}, 1)$. $\Delta_3(\nu) > 0$ is equivalent to

$$p_2(\nu)\sqrt{(1-\nu)(\nu+5)} < -(8\nu^4 + 30\nu^3 - 48\nu^2 + 10\nu) \doteq p_9(\nu). \tag{5.6}$$

Earlier, we have proved that $p_2(\nu) < 0$ when $\nu \in (\frac{\sqrt{2}}{2}, 1)$. On the other hand, $p_9(\nu) = -\nu(4\nu - 1)(2\nu + 10)(\nu - 1) > 0$ for $\nu \in (\frac{\sqrt{2}}{2}, 1)$. Thus, (5.6) holds for $\nu \in (\frac{\sqrt{2}}{2}, 1)$.
$\square$

Acknowledgements The authors are grateful to the two anonymous reviewers and the editor for their helpful comments and suggestions that have helped to improve the paper.

Author Contributions All authors read and approved the final manuscript. Cairong Chen and Dongmei Yu mainly contribute to the theoretical analysis and the numerical experiments. Bo Huang is mainly responsible for the symbolic computation. Deren Han mainly contributes to the theoretical analysis.

Funding The research of Cairong Chen is supported by the Natural Science Foundation of Fujian Province (No. 2021J01661) and the National Natural Science Foundation of China (No. 11901024). The research of Bo Huang is supported by the National Natural Science Foundation of China (No. 12101032). The research of Dongmei Yu is supported by the National Natural Science Foundation of China (No. 12201275), the Ministry of Education in China of Humanities and Social Science Project (No. 21YJCZH204) and the Natural Science Foundation of Liaoning Province (No. 2020-MS-301). The research of Deren Han is supported by the National Natural Science Foundation of China (No. 12131004) and the Ministry of Science and Technology of China (No. 2021YFA1003600).

Data Availability Data sharing not applicable to this paper as no datasets were generated or analyzed during the current study.

Declarations

Ethics approval and consent to participate Not Applicable.

Conflict of interest The authors declare no competing interests.

References

1. Abdallah, L., Haddou, M., Migot, T.: Solving absolute value equation using complementarity and smoothing functions. J. Comput. Appl. Math. **327**, 196–207 (2018)
2. Atkinson, K., Han, W.-M.: Elementary Numerical Analysis. John Wiley & Sons Inc, NJ (2004)
3. Bello Cruz, J.Y., Ferreira, O.P., Prudente, L.F.: On the global convergence of the inexact semi-smooth Newton method for absolute value equation. Comput. Optim. Appl. **65**, 93–108 (2016)
4. Caccetta, L., Qu, B., Zhou, G.-L.: A globally and quadratically convergent method for absolute value equations. Comput. Optim. Appl. **48**, 45–58 (2011)
5. Chen, C.-R., Yang, Y.-N., Yu, D.-M., Han, D.-R.: An inverse-free dynamical system for solving the absolute value equations. Appl. Numer. Math. **168**, 170–181 (2021)
6. Chen, C.-R., Yu, D.-M., Han, D.-R.: Exact and inexact Douglas-Rachford splitting methods for solving large-scale sparse absolute value equations. IMA J. Numer. Anal. **43**, 1036–1060 (2023)
7. Davis, T.: University of Florida Sparse Matrix Collection. University of Florida, Gainesville, FL. http://www.cise.ufl.edu.research/sparse/matrices/
8. Edalatpour, V., Hezari, D., Salkuyeh, D.K.: A generalization of the Gauss-Seidel iteration method for solving absolute value equations. Appl. Math. Comput. **293**, 156–167 (2017)
9. Guo, P., Wu, S.-L., Li, C.-X.: On the SOR-like iteration method for solving absolute value equations. Appl. Math. Lett. **97**, 107–113 (2019)
10. Hladík, M.: Bounds for the solutions of absolute value equations. Comput. Optim. Appl. **69**, 243–266 (2018)
11. Hladík, M.: Properties of the solution set of absolute value equations and the related matrix classes. SIAM J. Matrix Anal. Appl. **44**(1), 175–195 (2023)
12. Hu, S.-L., Huang, Z.-H.: A note on absolute value equations. Optim. Lett. **4**, 417–424 (2010)

13. Iqbal, J., Iqbal, A., Arif, M.: Levenberg-Marquardt method for solving systems of absolute value equations. J. Comput. Appl. Math. **282**, 134–138 (2015)
14. Ke, Y.-F.: The new iteration algorithm for absolute value equation. Appl. Math. Lett. **99**, 105990 (2020)
15. Ke, Y.-F., Ma, C.-F.: SOR-like iteration method for solving absolute value equations. Appl. Math. Comput. **311**, 195–202 (2017)
16. Lian, Y.-Y., Li, C.-X., Wu, S.-L.: Weaker convergent results of the generalized Newton method for the generalized absolute value equations. J. Comput. Appl. Math. **338**, 221–226 (2018)
17. Mangasarian, O.L.: Absolute value programming. Comput. Optim. Appl. **36**, 43–53 (2007)
18. Mangasarian, O.L.: Absolute value equation solution via concave minimization. Optim. Lett. **1**, 3–8 (2007)
19. Mangasarian, O.L.: A generalized Newton method for absolute value equations. Optim. Lett. **3**, 101–108 (2009)
20. Mangasarian, O.L., Meyer, R.R.: Absolute value equations. Linear Algebra Appl. **419**, 359–367 (2006)
21. Mansoori, A., Erfanian, M.: A dynamic model to solve the absolute value equations. J. Comput. Appl. Math. **333**, 28–35 (2018)
22. Mansoori, A., Eshaghnezhad, M., Effati, S.: An efficient neural network model for solving the absolute value equations. IEEE T. Circuits-II **65**, 391–395 (2017)
23. Miao, X.-H., Yang, J.-T., Saheya, B., Chen, J.-S.: A smoothing Newton method for absolute value equation associated with second-order cone. Appl. Numer. Math. **120**, 82–96 (2017)
24. Noor, M.A., Iqbal, J., Noor, K.I., Al-Said, E.: On an iterative method for solving absolute value equations. Optim. Lett. **6**, 1027–1033 (2012)
25. Prokopyev, O.: On equivalent reformulations for absolute value equations. Comput. Optim. Appl. **44**, 363–372 (2009)
26. Rohn, J.: A theorem of the alternatives for the equation $Ax + B|x| = b$. Linear Multilinear Algebra **52**, 421–426 (2004)
27. Saheya, B., Yu, C.-H., Chen, J.-S.: Numerical comparisons based on four smoothing functions for absolute value equation. J. Appl. Math. Comput. **56**, 131–149 (2018)
28. Song, J., Song, Y.-Z.: Relaxed-based matrix splitting methods for solving absolute value equations. Comp. Appl. Math. **42**, 19 (2023)
29. Stewart, J.: Calculus Early Transcendentals, 6th edn. Thomson Learning, USA (2008)
30. Wang, A., Cao, Y., Chen, J.-X.: Modified Newton-type iteration methods for generalized absolute value equations. J. Optim. Theory Appl. **181**, 216–230 (2019)
31. Wang, H.-J., Cao, D.-X., Liu, H., Qiu, L.: Numerical validation for systems of absolute value equations. Calcolo **54**, 669–683 (2017)
32. Wu, S.-L., Li, C.-X.: A note on unique solvability of the absolute value equation. Optim. Lett. **14**, 1957–1960 (2020)
33. Xia, B.-C.: DISCOVERER: a tool for solving semi-algebraic systems. ACM Commun. Comput. Algebra **41**, 102–103 (2007)
34. Yang, L., Hou, X.-R., Xia, B.-C.: A complete algorithm for automated discovering of a class of inequality-type theorems. Sci. China Ser. F **44**, 33–49 (2001)
35. Yu, D.-M., Chen, C.-R., Han, D.-R.: A modified fixed point iteration method for solving the system of absolute value equations. Optimization **71**, 449–461 (2022)
36. Zamani, M., Hladík, M.: Error bounds and a condition number for the absolute value equations. Math. Program. **198**, 85–113 (2023)