

A Modified Discrete-Time Jacobi Waveform Relaxation Iteration

Yong Liu^{1,2}, Shulin Wu³

¹Department of Instrument Science and Engineering, Southeast University, Nanjing, China

²Department of the Aeronautic Instruments and Electronic Engineering, The First Aeronautic Institute of Air Force, Xinyang, China

³School of Science, Sichuan University of Science and Engineering, Zigong, China

E-mail: sun_myth@sohu.com, wushulin_ylp@163.com

Received November 7, 2010; revised March 11, 2011; accepted March 14, 2011

Abstract

In this paper, we investigate an accelerated version of the discrete-time Jacobi waveform relaxation iteration method. Based on the well known Chebyshev polynomial theory, we show that significant speed up can be achieved by taking linear combinations of earlier iterates. The convergence and convergence speed of the new iterative method are presented and it is shown that the convergence speed of the new iterative method is sharper than that of the Jacobi method but blunter than that of the optimal SOR method. Moreover, at every iteration the new iterative method needs almost equal computation work and memory storage with the Jacobi method, and more importantly it can completely exploit the particular advantages of the Jacobi method in the sense of parallelism. We validate our theoretical conclusions with numerical experiments.

Keywords: Discrete-Time Waveform Relaxation, Convergence, Parallel Computation, Chebyshev Polynomial, Jacobi Iteration, Optimal SOR

1. Introduction

For very large scale initial value problems (IVPs), linear or nonlinear, the waveform relaxation (WR) iteration, also called dynamic iteration, is a very powerful method and has received much interest from many researchers in the past years, see [1-9] for more details about the history of this method. The difference of the classical iteration methods with the WR iteration method is that the WR method iterates with functions in a functional space (continuous-time WR), and solve each iteration by some numerical method (discrete-time WR), e.g. by the Runge-Kutta methods. In the past years, both the continuous-time WR iteration method and the discrete-time WR iteration method have been investigated widely. For example, one may refer to [1,3,9-13] for the the WR method dis-cussed in continuous time level, and to [9,12, 14-17] for the discrete-time WR method. There are so many excellent results in this field that we can not recount them detaildly.

For the linear system of IVPs on arbitrarily long time interval $[0, T]$ defined as

$$\begin{cases} y'(t) + Hy(t) = f(t), t \in [0, T], \\ y(t_0) = y_0, \end{cases} \quad (1)$$

where $f, y: \mathbb{R}^+ \rightarrow \mathbb{R}^n$, $y_0 \in \mathbb{R}^n$, $H \in \mathbb{R}^{n \times n}$, the Jacobi WR iteration can be written as

$$\frac{dy^{(k+1)}}{dt} + My^{(k+1)} = Ny^{(k)} + f, y^{(k+1)}(0) = y_0, k = 0, 1, \dots, \quad (2)$$

where $H = M - N$ and M is a pointwise or block-wise diagonal matrix. The discrete-time Jacobi WR iteration can be obtained by applying some numerical method, such as BDF method, Runge-Kutta method, etc., to discretize (2) for every iterative index k .

In Jacobi waveform relaxation, continuous-time or discrete-time, system (2) is decoupled into, say d , loosely coupled subsystems. If, on a parallel computer, these subsystems are assigned to d different processors, and they can be solved on $[0, T]$ simultaneously. This obvious type of parallelism is present in all waveform relaxation methods, see, e.g., [1,2-4,9-12,15-19] and references therein. However, in many cases, such as the Gauss-Seidel, optimal SOR waveform relaxation methods, exploiting this parallelism is only possible when approximations are exchanged between the different processors as soon as they have been computed (see, e.g., [20]). This leads to a large amount of communication during

each waveform relaxation sweep, which forms a severe drawback. In Jacobi waveform relaxation, communication is only necessary once per sweep, and this method is attractive for parallel implementation.

Therefore, we think that any improvement of the convergence speed of the Jacobi WR iteration is important. Based on this consideration, in this paper, we attempt to get speed up by taking linear combinations of earlier Jacobi iterates, and this leads to the following special iterative scheme:

$$\begin{aligned}
 &\text{for } k = 0, 1, 2, \dots \\
 &\quad x^{(0)} = y^{(k)}; \\
 &\text{for } m = 1, 2, \dots, \alpha \\
 &\quad \frac{dx^{(m)}(t)}{dt} + \mathbf{M}x^{(m)}(t) = \mathbf{N}x^{(m-1)}(t) + f(t); \quad (3) \\
 &\text{end} \\
 &\quad y^{(k+1)} = \sum_{m=0}^{\alpha} v_m x^{(m)}; \\
 &\text{end}
 \end{aligned}$$

where the parameters $\{v_m\}_{m=0}^{\alpha}$ are constant and satisfy

$$\sum_{m=0}^{\alpha} v_m = 1. \quad (4)$$

Since the discrete-time WR iteration is more favourable in practical applications, we will devote ourself to getting the speed up of the discrete-time version of iterative scheme (3). To solve the ODEs in (3), we first decompose the time interval $[0, T]$ into $N + s$ sub-intervals, say $[t_n, t_{n+1}]$, $n = 0, 1, \dots, N + s - 1$, with equal length h . And then we solve (3) numerically by a linear s -step formula with coefficients $(\alpha_j, \beta_j)_{j=0,1,\dots,s}$, which leads to the following iterative scheme written in compact form as

$$\begin{aligned}
 &\text{for } k = 0, 1, 2, \dots \\
 &\quad y_j^{(k)} = y_j^{(0)}, j = 0, 1, \dots, s - 1; \\
 &\quad x_n^{(0)} = y_n^{(k)}, n = 0, 1, \dots, N + s \\
 &\text{for } m = 1, 2, \dots, \alpha \\
 &\quad x_j^{(m)} = y_j^{(0)}, j = 0, 1, \dots, s - 1 \\
 &\text{for } n = 0, 1, 2, \dots, N \quad (5) \\
 &\quad \sum_{j=0}^s \alpha_j x_{n+j}^{(m)} = h \sum_{j=0}^s \beta_j (-\mathbf{M}x_{n+j}^{(m)} + \mathbf{N}x_{n+j}^{(m-1)} + f_{n+j}); \\
 &\text{end} \\
 &\text{end} \\
 &\quad y_{n+s}^{(k+1)} = \sum_{m=0}^{\alpha} v_m x_{n+s}^{(m)}, n = 0, \dots, N; \\
 &\text{end}
 \end{aligned}$$

where the values $y_j^{(0)}, j = 0, 1, \dots, s - 1$ are the s initial values which are obtained by, for example, the backward Euler method or some other one-step methods. We denote the new iterative scheme (5) by *Acc-Jacobi* in the remainder of this paper. We will show that the optimal parameters $\{v_m\}_{m=0}^{\alpha}$ which are used to accelerate the convergence of the original discrete-time Jacobi WR iteration relate closely to the coefficients of the α -th Chebyshev polynomial. Moreover, the effects of the step size h and the matrices \mathbf{M}, \mathbf{N} on the convergence speed of the *Acc-Jacobi* method are also presented. We note that, for the sake of economizing memory storage, formula (5) can be performed in special wise which is independent of the parameter α and nearly equals to that of the classical discrete-time Jacobi WR iteration.

The remainder of this paper is organized as follows. In Section 2, we recall the definitions of the so-called $H_B^{(1)}$ -block matrix and \mathbf{M} -matrix, and some related properties. In Section 3, the convergence speed of the discrete-time WR iteration (5) is derived and some comparisons about convergence speed between the *Acc-Jacobi*, the classical Jacobi and the optimal SOR methods are given. In Section 4, we present some numerical results which validate our theoretical conclusions very well.

2. Some Basic Knowledge of Matrix

Throughout this paper, the partial orderings ' \leq ', ' $<$ ' and the absolute value $|\cdot|$ in \mathbb{R}^n and $\mathbb{R}^{n \times n}$ are interpreted in componentwise. For a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, let $q (\leq n)$ and $n_i (\leq n), i = 1, \dots, q$, be some positive integers satisfying $\sum_{i=1}^q n_i = n$. Then define the block-wise vector and matrix spaces as follows (see also [11]):

$$\begin{aligned}
 V_q &= \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = (\mathbf{x}_1^T, \dots, \mathbf{x}_q^T)^T, \mathbf{x}_i \in \mathbb{R}^{n_i} \right\}, \\
 L_q &= \left\{ \mathbf{A} \in \mathbb{R}^{n \times n} \mid \mathbf{A} = (\mathbf{A}_{ij})_{q \times q}, \right. \\
 &\quad \left. \mathbf{A}_{ij} \in \mathbb{R}^{n_i \times n_j}, \mathbf{A}_{ii}^{-1} \text{ is exist for } i = 1, \dots, q \right\}. \quad (6)
 \end{aligned}$$

Moreover, for any matrix $\mathbf{G} \in \mathbb{R}^{n \times n}$, let

$$\begin{aligned}
 \mathbf{D}_G &= \text{diag}(g_{11}, \dots, g_{nn}), \mathbf{B}_G = \mathbf{D}_G - \mathbf{G}, \\
 \text{and } \mathbf{J}_G &= \mathbf{D}_G^{-1} \mathbf{B}_G, \quad (7)
 \end{aligned}$$

provided that the quantities $q_{ii} \neq 0, i = 1, 2, \dots, n$.

Definition 1 ([21]). A real matrix $\mathbf{A} = (a_{ij})_{n \times n}$ with $a_{ij} \leq 0$ for all $i \neq j$ is an *M-matrix* if \mathbf{A} is nonsingular and $\mathbf{A}^{-1} \geq 0$.

Definition 2 ([11, 22]). $\mathbf{H} \in L_q$ is an $H_B^{(1)}$ -block matrix, if its block comparison matrix $\langle \mathbf{H} \rangle$ defined by

$$\langle \mathbf{H} \rangle_{ij} = \begin{cases} \|\mathbf{H}_{ii}^{-1}\|^{-1}, & i = 1, \dots, q, \\ -\|\mathbf{H}_{ij}\|, & i \neq j, i, j = 1, \dots, q. \end{cases} \quad (8)$$

is an M -matrix, here and hereafter $\|\cdot\|$ denotes the standard Euclidean norm.

Clearly, any M -matrix $\mathbf{A} = (a_{ij})$ with positive diagonal elements is an $H_B^{(1)}$ -block matrix.

For $\mathbf{H} \in L_q$, we use $[\mathbf{H}] = (\|\mathbf{H}_{ij}\|) \in \mathbb{R}^{q \times q}$ to represent the block absolute value. The block absolute value of a vector $\mathbf{x} \in V_q$ can be defined in a similar way, i.e., $[\mathbf{x}] = (\|\mathbf{x}_1^T\|, \dots, \|\mathbf{x}_q^T\|)^T$.

Definition 3 ([21]). For $n \times n$ real matrices \mathbf{H} , \mathbf{M} and \mathbf{N} , $\mathbf{H} = \mathbf{M} - \mathbf{N}$ is a regular splitting of matrix \mathbf{H} , if \mathbf{M} is nonsingular with $\mathbf{M}^{-1} \geq 0$ and $\mathbf{N} \geq 0$.

Lemma 1 ([11, 22]). Let $\mathbf{L}, \mathbf{M} \in L_q$, $\mathbf{x}, \mathbf{y} \in V_q$ and $\chi \in \mathbb{C}$. Then

- 1) $\|[\mathbf{L}] - [\mathbf{M}]\| \leq [\mathbf{L} + \mathbf{M}]$
 $\leq [\mathbf{L}] + [\mathbf{M}]$ ($\|[\mathbf{x}] - [\mathbf{y}]\| \leq [\mathbf{x} + \mathbf{y}] \leq [\mathbf{x}] + [\mathbf{y}]$);
- 2) $[\mathbf{LM}] \leq [\mathbf{L}][\mathbf{M}]$ ($[\mathbf{Mx}] \leq [\mathbf{M}][\mathbf{x}]$);
- 3) $[\chi\mathbf{M}] \leq |\chi|[\mathbf{M}]$ ($[\chi\mathbf{x}] \leq |\chi|[\mathbf{x}]$);
- 4) $\rho(\mathbf{M}) \leq \rho([\mathbf{M}])$.

Lemma 2 ([11]). Let $\mathbf{H} \in L_q$ be an $H_B^{(1)}$ -block matrix, then

- 1) \mathbf{H} is nonsingular;
- 2) $[\mathbf{H}^{-1}] \leq [\mathbf{H}]^{-1}$;
- 3) $\rho(\mathbf{J}_{\langle \mathbf{H} \rangle}) < 1$,

where $\mathbf{J}_{\langle \mathbf{H} \rangle} = \mathbf{D}_{\langle \mathbf{H} \rangle}^{-1} \mathbf{B}_{\langle \mathbf{H} \rangle}$ with $\mathbf{J}_{\langle \mathbf{H} \rangle} = \mathbf{D}_{\langle \mathbf{H} \rangle}^{-1} \mathbf{B}_{\langle \mathbf{H} \rangle} \mathbf{D}_{\langle \mathbf{H} \rangle} = \text{diag}(\|\mathbf{H}_{11}^{-1}\|^{-1}, \dots, \|\mathbf{H}_{qq}^{-1}\|^{-1})$ and $\mathbf{B}_{\langle \mathbf{H} \rangle} = \langle \mathbf{H} \rangle - \mathbf{D}_{\langle \mathbf{H} \rangle}$.

Lemma 3 ([21]). Let $\mathbf{H} = \mathbf{M} - \mathbf{N}$ be a regular splitting of the matrix \mathbf{H} . Then \mathbf{H} is nonsingular with $\mathbf{H}^{-1} \geq 0$, if and only if $\rho(\mathbf{M}^{-1}\mathbf{N}) < 1$.

Lemma 4 ([21]). Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, satisfy $\mathbf{A} \geq \mathbf{B} \geq 0$. Then $\rho(\mathbf{A}) \geq \rho(\mathbf{B})$.

Lemma 5 ([23]). Let $\mathbf{H} \in \mathbb{R}^{n \times n}$, and can be splitted as $\mathbf{H} = \beta \mathbf{I} - \mathbf{B}$, where $\beta > 0$, $\mathbf{B} \geq 0$, then the following results are equivalent:

- 1) $\mathbf{H}^{-1} \geq 0$;
- 2) $\rho(\mathbf{B}) < \beta$;
- 3) for $i = 1, 2, \dots, n$, $\text{Re}(\lambda_i) \geq 0$, where λ_i is the eigenvalue of the matrix.

Lemma 6 ([21]). Let $\mathbf{H}^{-1} \geq 0$ and $\mathbf{H} = \mathbf{M}_1 - \mathbf{N}_1 = \mathbf{M}_2 - \mathbf{N}_2$ be two regular splittings of \mathbf{H} . Then $1 > \rho(\mathbf{M}_2^{-1}\mathbf{N}_2) \geq \rho(\mathbf{M}_1^{-1}\mathbf{N}_1) \geq 0$ if $\mathbf{N}_2 \geq \mathbf{N}_1 \geq 0$; $1 > \rho(\mathbf{M}_2^{-1}\mathbf{N}_2) > \rho(\mathbf{M}_1^{-1}\mathbf{N}_1) > 0$ if $\mathbf{N}_2 \geq \mathbf{N}_1 \geq 0$.

3. Convergence Analysis

As supposed in [11], we assume in this paper that the matrix $\mathbf{H} (\in \mathbb{R}^{n \times n})$ and its splitting $\mathbf{H} = \mathbf{M} - \mathbf{N}$ satisfy the following assumption.

Assumption (A). Assume that all the off diagonal elements of \mathbf{H} are non-positive and \mathbf{H} is a symmetric and $H_B^{(1)}$ -block matrix; its splitting $\mathbf{H} = \mathbf{M} - \mathbf{N}$ satisfies for some integer $q (1 \leq q \leq n)$ that $\mathbf{M} = \text{diag}(\mathbf{H}_{11}, \dots, \mathbf{H}_{qq})$ is a symmetric positive definite matrix and \mathbf{M}, \mathbf{N} are commutative matrices, i.e., $\mathbf{MN} = \mathbf{NM}$.

Under this assumption, we know that \mathbf{N} is also a symmetric matrix. Therefore, the eigenvalues of the product \mathbf{MN} are all real numbers, since for symmetric matrices \mathbf{M} and \mathbf{N} , \mathbf{M}, \mathbf{N} are commutative matrices if and only if $(\mathbf{MN})^T = \mathbf{MN}$, see, e.g., [21,23]. We next make an assumption for the linear s -step formula as follows.

Assumption (B). Assume that $\frac{\alpha_s}{\beta_s} > 0$, where α_s and β_s are the coefficients of the linear s -step formula.

Lemma 7. Let matrix \mathbf{H} and its splitting $\mathbf{H} = \mathbf{M} - \mathbf{N}$ satisfy assumption (A). Then for any real number $\bar{h} \geq 0$, $\bar{h}\mathbf{I} + \mathbf{H}$ is an M -matrix and $\rho((\bar{h}\mathbf{I} + \mathbf{M})^{-1}\mathbf{N}) < 1$.

Proof. Since $\mathbf{M} = \text{diag}(\mathbf{H}_{11}, \dots, \mathbf{H}_{qq})$ is a symmetric positive definite matrix with non-positive off diagonal elements, by the conclusions 1, 3 of lemma 5 and the results given in [11], we know that for any $\bar{h} \geq 0$ it holds

$$(\bar{h}\mathbf{I} + \mathbf{M})^{-1} \geq 0 \text{ and } \|(\bar{h}\mathbf{I} + \mathbf{M})^{-1}\| \leq \|\mathbf{M}^{-1}\|.$$

Clearly, the inequality $\|(\bar{h}\mathbf{I} + \mathbf{M})^{-1}\| \leq \|\mathbf{M}^{-1}\|$ implies

$0 \leq \mathbf{D}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle}^{-1} \leq \mathbf{D}_{\langle \mathbf{H} \rangle}^{-1}$. Thus, by definition 3,

$\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle = \mathbf{D}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle} - \mathbf{B}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle}$ is a regular splitting of matrix $\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle$ and $\mathbf{D}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle}^{-1} \mathbf{B}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle} \leq \mathbf{D}_{\langle \mathbf{H} \rangle}^{-1} \mathbf{B}_{\langle \mathbf{H} \rangle}$, since $\mathbf{B}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle} = \mathbf{B}_{\langle \mathbf{H} \rangle} \geq 0$, where the definition of the matrices $\mathbf{D}_{\langle \mathbf{H} \rangle}$ and $\mathbf{B}_{\langle \mathbf{H} \rangle}$ has been given in lemma 2.

Then it follows from lemma 4 and the conclusion 3 of lemma 2 that

$$\rho(\mathbf{D}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle}^{-1} \mathbf{B}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle}) \leq \rho(\mathbf{D}_{\langle \mathbf{H} \rangle}^{-1} \mathbf{B}_{\langle \mathbf{H} \rangle}) < 1. \quad (9)$$

Moreover, by definition 2 and the notations given in lemma 2, it is clear that $[\mathbf{N}] = \mathbf{B}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle} \geq 0$,

$\langle \bar{h}\mathbf{I} + \mathbf{M} \rangle = \mathbf{D}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle}$ and the matrix $\langle \bar{h}\mathbf{I} + \mathbf{M} \rangle$ is an $H_B^{(1)}$ -block matrix, since $\mathbf{M} = \text{diag}(\mathbf{H}_{11}, \dots, \mathbf{H}_{qq})$ is a sym-

metric positive definite matrix.

Therefore, we have

$$\begin{aligned} \rho\left((\bar{h}\mathbf{I} + \mathbf{M})^{-1} \mathbf{N}\right) &\leq \rho\left[\left[(\bar{h}\mathbf{I} + \mathbf{M})^{-1}\right][\mathbf{N}]\right] \\ &\leq \rho\left(\langle \bar{h}\mathbf{I} + \mathbf{M} \rangle^{-1} [\mathbf{N}]\right) \quad (10) \\ &= \rho\left(\mathbf{D}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle}^{-1} \mathbf{B}_{\langle \bar{h}\mathbf{I} + \mathbf{H} \rangle}\right) < 1, \end{aligned}$$

where in the first, second and the last inequalities we have used the conclusions 2, 4 of lemma 1, the conclusion 2 of lemma 2 and (9), respectively.

Finally, from (10) we know that $\bar{h}\mathbf{I} + \mathbf{H} = (\bar{h}\mathbf{I} + \mathbf{M}) - \mathbf{N}$ is a regular splitting of matrix $\bar{h}\mathbf{I} + \mathbf{H}$ which satisfies $\rho\left((\bar{h}\mathbf{I} + \mathbf{H})^{-1} \mathbf{N}\right) < 1$. Hence, according to lemma 3, it is easy to know that $\bar{h}\mathbf{I} + \mathbf{H}$ is an M -matrix. \square

Lemma 8 Let matrix H and its splitting $\mathbf{H} = \mathbf{M} - \mathbf{N}$ satisfy assumption (A). Then for any real number r , all the eigenvalues of matrix $(r\mathbf{I} + \mathbf{M})^{-1} \mathbf{N}$ are real numbers.

Proof. In assumption (A), we have mentioned that the matrices \mathbf{M} and \mathbf{N} are commutative. Therefore, by the results in [21,23], we know that there exists a non-singular matrix, say \mathbf{Q} , such that

$$\begin{aligned} \mathbf{QM}\mathbf{Q}^{-1} &= \text{diag}\left(\lambda_1^M, \dots, \lambda_n^M\right) \text{ and} \\ \mathbf{QN}\mathbf{Q}^{-1} &= \text{diag}\left(\lambda_1^N, \dots, \lambda_n^N\right), \end{aligned}$$

where $\{\lambda_i^M\}_{i=1}^n$ and $\{\lambda_i^N\}_{i=1}^n$ are the real eigenvalues of matrices \mathbf{M} and \mathbf{N} , respectively.

Thus, for any real number r , it is obvious that the n eigenvalues of the matrix $(r\mathbf{I} + \mathbf{M})^{-1} \mathbf{N}$ are $\left\{\frac{\lambda_i^N}{r + \lambda_i^M}\right\}_{i=1}^n$,

which are real numbers. \square

Now, we turn our attention to the convergence analysis of the Acc-Jacobi method defined by (5). We denote the convergent solution of the Acc-Jacobi method by y_j , $j = 0, 1, \dots, N + s$ with $y_j = y_j^{(0)}$ for $j = 0, 1, \dots, s - 1$. We thus define the notations as Equation (11) shows.

Therefore, by careful routine calculations we can equivalently rewritten the Acc-Jacobi iteration (5) as

$$\mathbf{X}^m = \tilde{\mathbf{M}}^{-1} \tilde{\mathbf{N}} \mathbf{X}^{(m-1)} + \tilde{\mathbf{M}}^{-1} (\tilde{\mathbf{f}} + \tilde{\mathbf{r}}) = (\tilde{\mathbf{M}}^{-1} \tilde{\mathbf{N}})^m \mathbf{Y}^{(k)} + \tilde{\mathbf{c}}_m, \quad (12)$$

where $\tilde{\mathbf{c}}_m = \left(\sum_{j=0}^{m-1} (\tilde{\mathbf{M}}^{-1} \tilde{\mathbf{N}})^j\right) \tilde{\mathbf{M}}^{-1} (\tilde{\mathbf{f}} + \tilde{\mathbf{r}})$ with $(\tilde{\mathbf{M}}^{-1} \tilde{\mathbf{N}})^0 = \mathbf{I}$. This implies the following relation between $\mathbf{Y}^{(k+1)}$ and $\mathbf{Y}^{(k)}$ as

$$\mathbf{Y}^{(k+1)} = \left(\sum_{m=0}^{\alpha} v_m (\tilde{\mathbf{M}}^{-1} \tilde{\mathbf{N}})^m\right) \mathbf{Y}^{(k)} + \sum_{m=0}^{\alpha} v_m \tilde{\mathbf{c}}_m. \quad (13)$$

Hence, it follows from the convergence theory of Jacobi iteration [21,23] that the vector $\mathbf{Y}^{(k)}$ generated by (13) converges to some value, say \mathbf{Y} for example, if and only if it holds that

$$\rho\left(\sum_{m=0}^{\alpha} v_m (\tilde{\mathbf{M}}^{-1} \tilde{\mathbf{N}})^m\right) = \rho\left(\sum_{m=0}^{\alpha} v_m (\Delta_s^{-1} \tilde{\mathbf{N}}_s)^m\right) < 1. \quad (14)$$

Therefore, we arrive at the following question: how to select parameters $\{v_m\}_{m=0}^{\alpha}$ which minimize the argument $\rho\left(\sum_{m=0}^{\alpha} v_m (\Delta_s^{-1} \tilde{\mathbf{N}}_s)^m\right)$, and this leads to following minimum problem

$$\min_{v_0 + v_1 + \dots + v_{\alpha} = 1} \rho\left(\sum_{m=0}^{\alpha} v_m C^m\right), \quad (15a)$$

$$\begin{aligned} \mathbf{X}^{(m)} &= \left(\left(\mathbf{x}_s^{(m)}\right)^T, \left(\mathbf{x}_{s+1}^{(m)}\right)^T, \dots, \left(\mathbf{x}_{N+s}^{(m)}\right)^T\right)^T, \mathbf{Y}^{(k)} = \left(\left(\mathbf{y}_s^{(k)}\right)^T, \left(\mathbf{y}_{s+1}^{(k)}\right)^T, \dots, \left(\mathbf{y}_{N+s}^{(k)}\right)^T\right)^T, \\ m &= 0, 1, \dots, \alpha, \text{ and } k = 0, 1, \dots, \\ \tilde{\mathbf{r}} &= \left(h \sum_{j=0}^{s-1} (\beta_j \mathbf{I} + h\mathbf{H} \beta_j) \left(\mathbf{y}_j^{(0)}\right)^T, h \sum_{j=0}^{s-2} (\beta_j \mathbf{I} + h\mathbf{H} \beta_j) \left(\mathbf{y}_j^{(0)}\right)^T, \dots, h(\beta_0 \mathbf{I} + h\mathbf{H} \beta_0) \left(\mathbf{y}_0^{(0)}\right)^T, 0, \dots, 0\right)^T \\ \tilde{\mathbf{f}} &= \left(h \sum_{j=0}^s \beta_j f_j^T, h \sum_{j=0}^s \beta_j f_{1+j}^T, \dots, h \sum_{j=0}^s \beta_j f_{N+j}^T\right)^T, \Delta_i = \alpha_i \mathbf{I} + h\beta_i \mathbf{M}, \tilde{\mathbf{N}}_i = h\beta_i \mathbf{N}, i = 0, 1, \dots, s, \\ \tilde{\mathbf{M}} &= \begin{pmatrix} \Delta_s & & & & & \\ \Delta_{s-1} & \Delta_s & & & & \\ \dots & \dots & \dots & & & \\ \Delta_0 & \Delta_1 & \dots & \Delta_s & & \\ 0 & \ddots & \ddots & \ddots & \Delta_s & \\ 0 & 0 & \Delta_0 & \Delta_1 & \dots & \Delta_s \end{pmatrix}, \tilde{\mathbf{N}} = \begin{pmatrix} \tilde{\mathbf{N}}_s & & & & & \\ \tilde{\mathbf{N}}_{s-1} & \tilde{\mathbf{N}}_s & & & & \\ \dots & \dots & \dots & & & \\ \tilde{\mathbf{N}}_0 & \tilde{\mathbf{N}}_1 & \dots & \tilde{\mathbf{N}}_s & & \\ 0 & \ddots & \ddots & \ddots & \tilde{\mathbf{N}}_s & \\ 0 & 0 & \tilde{\mathbf{N}}_0 & \tilde{\mathbf{N}}_1 & \dots & \tilde{\mathbf{N}}_s \end{pmatrix} \end{aligned} \quad (11)$$

here and hereafter we set

$$C = \Delta_s^{-1} \tilde{N}_s \tag{15b}$$

for convenience.

Definition 4 Let S_α be the set of all polynomials $p_\alpha(x)$ of degree α with $p_\alpha(1) = 1$.

By this definition, we can rewrite problem (15) as

$$\min_{p_\alpha \in S_\alpha} \rho(p_\alpha(C)). \tag{16}$$

It is very difficult to solve problem (17) exactly. However, if we have $\rho = \rho(C)$ at hand, we may consider the following problem

$$\min_{p_\alpha \in S_\alpha} \left\{ \max_{-\rho \leq x \leq \rho} |p_\alpha(x)| \right\}. \tag{17}$$

Clearly,

$$\min_{p_\alpha \in S_\alpha} \rho(p_\alpha(C)) \leq \min_{p_\alpha \in S_\alpha} \left\{ \max_{-\rho \leq x \leq \rho} |p_\alpha(x)| \right\}. \tag{18}$$

The *max-min* problem (18) is classical and has been investigated deeply by the famous mathematician Pafnuty Chebyshev (1821-1894), see, e.g., [24]. The solution of this problem can be obtained by the so called Chebyshev polynomial $U_\alpha(x)$ which is defined as follows.

Definition 5 ([21,24])

$$U_\alpha(x) = \begin{cases} \cos(\alpha \cos^{-1}(x)), & \text{if } -1 \leq x \leq 1, \\ \cosh(\alpha \cosh^{-1}(x)), & \text{if } x > 1, \end{cases} \tag{19}$$

where $\alpha \geq 1$ is an integer.

The following two lemmas about the function $U_\alpha(x)$ are useful to our subsequent analysis.

Lemma 9 ([21,24]) The polynomial $U_\alpha(x)$ defined in (20) satisfies the three-term recursion relation as

$$U_{\alpha+1}(x) = 2xU_\alpha(x) - U_{\alpha-1}(x), \tag{20}$$

with $U_0(x) = 1$ and $U_1(x) = x$.

Lemma 10 The function $\frac{1}{U_\alpha(1/x)}$ ($\alpha \geq 1$) for $x \in (0,1)$

is strictly monotone increasing.

Proof. Define $g(x) = \frac{x}{1 + \sqrt{1-x^2}}$ and $f(x) = \frac{2x^\alpha}{1+x^{2\alpha}}$,

$0 < x < 1$. Then routine computation shows that

$$g'(x) = \frac{1}{1 + \sqrt{1-x^2}} + \frac{x^2}{\sqrt{1-x^2} (1 + \sqrt{1-x^2})^2} > 0$$

and

$$f'(x) = \frac{2\alpha x^{\alpha-1} (1-x^{2\alpha})}{(1+y^{2\alpha})^2} > 0$$

for $0 < x < 1$ and $\alpha \geq 1$. This implies that $f(g(x))$ is

a strictly monotone increasing function for $0 < x < 1$ and $\alpha \geq 1$. Therefore, we complete the proof by noting that $\frac{1}{U_\alpha(1/x)} = f(g(x))$ for $0 < x < 1$. \square

With the Chebyshev polynomial $U_\alpha(x)$, we have the following conclusion for the *max-min* problem (18).

Theorem 1 ([21]) The polynomial

$$p_\alpha(x) = \frac{U_\alpha(x/\rho)}{U_\alpha(1/\rho)}, \tag{21}$$

is the unique polynomial which solves problem (18).

We note that the coefficients of the polynomial p_α can be computed conveniently by the three-term recursion relation given by (21).

Next, we focus on deriving the spectral radius of the matrix $\sum_{m=0}^{\alpha} v_m C^m$, when $\{v_m\}_{m=0}^{\alpha}$ are the coefficients of the polynomial p_α defined by (22). Let $\{\lambda_i\}_{i=1}^n$ be the n eigenvalues of the matrix C (the matrix C is defined by (16)), and by lemma 8 we know that they are all real numbers. Therefore, it holds that

$$\begin{aligned} \rho\left(\sum_{m=0}^{\alpha} v_m C^m\right) &= \rho(p_\alpha(C)) = \max_{1 \leq i \leq n} |p_\alpha(\lambda_i)| \\ &= \max_{1 \leq i \leq n} \frac{|U_\alpha(\lambda_i/\rho)|}{U_\alpha(1/\rho)}, \end{aligned} \tag{22}$$

where $\rho = \rho(C)$. Then it follows, by noting that $|U_\alpha(\pm 1)| = 1$ and there must exist some $\tilde{\lambda} \in \{\lambda_i\}_{i=1}^n$ such that $|\tilde{\lambda}/\rho(C)| = 1$, that

$$\begin{aligned} \rho\left(\sum_{m=0}^{\alpha} v_m C^m\right) &= \rho(p_\alpha(C)) = \max_{1 \leq i \leq n} \frac{|U_\alpha(\lambda_i/\rho)|}{U_\alpha(1/\rho)} \\ &= \frac{1}{U_\alpha(1/\rho)}. \end{aligned} \tag{23}$$

Now, by assumptions (A) and (B), and lemma 7, we know that

$$\rho = \rho(C) = \rho(\Delta_s^{-1} \tilde{N}_s) = \rho\left(\left(\frac{\alpha_s}{h\beta_s} I + M\right)^{-1} N\right) < 1. \tag{24}$$

Then by the definition of $U_\alpha(x)$ given in (24), routine computation yields

$$\rho\left(\sum_{m=0}^{\alpha} v_m C^m\right) = \frac{1}{U_\alpha(1/\rho)} = \frac{2(\omega-1)^{\alpha/2}}{1+(\omega-1)^\alpha}, \tag{25}$$

where

$$\omega = \frac{2}{1 + \sqrt{1-\rho^2}} = 1 + \left(\frac{\rho}{1 + \sqrt{1-\rho^2}}\right)^2. \tag{26}$$

We next compare the convergence speed of the Acc-Jacobi method with the classical Jacobi iteration and the optimal SOR iteration. To this end, we first review the discrete-time Jacobi WR iteration and the discrete-time optimal SOR WR iteration investigated in [17,12] and [25,3,4], respectively. For simplicity, at the moment we just introduce the continuous-time Jacobi and SOR WR iterations, and the discrete-time version can be obtained straightforwardly by applying the linear s -step formula.

The Jacobi WR iteration can be written as

$$\frac{dy^{(k+1)}}{dt} + My^{(k+1)} = Ny^{(k)} + f; \quad (27)$$

where $H = M - N$ and M is a point or block diagonal matrix. Clearly, the Jacobi WR iteration (28) can be equivalently written as

$$\begin{aligned} x^{(0)} &= y^{(k)}, \\ \frac{dx^{(m)}}{dt} + Mx^{(m)} &= Nx^{(m-1)} + f, \quad m = 0, 1, \dots, \alpha, \\ y^{(k+1)} &= x^{(\alpha)}. \end{aligned} \quad (28)$$

Analogously, the optimal SOR WR iteration can be written as follows:

$$\begin{aligned} x^{(0)} &= y^{(k)}, \\ \frac{dx^{(m)}}{dt} + M_\omega x^{(m)} &= N_\omega x^{(m-1)} + f, \quad m = 0, 1, \dots, \alpha, \\ y^{(k+1)} &= x^{(\alpha)}, \end{aligned} \quad (29)$$

where $M_\omega = \frac{1}{\omega}(M - \omega L)$, $N_\omega = \frac{1-\omega}{\omega}M + U$ with the argument ω been defined by (27), and $H = M - U - L$. The matrix M is the point or block diagonal matrix of H just as mentioned in the Jacobi WR iteration (29) and L, U are the strictly lower and upper triangle matrices of the matrix $M - H$, respectively.

Similar to our above analysis, it is easy to know that the convergence speed of the discrete-time Jacobi iteration (after applying the linear s -step formula to (29)) and the discrete-time optimal SOR WR iteration (after applying the linear s -step formula to (30)) can be specified as

$$\rho_{Jac} = \rho^\alpha \text{ and } \rho_{SOR} = \left(\frac{\rho}{1 + \sqrt{1 - \rho^2}} \right)^{2\alpha}, \quad (30)$$

respectively, where $\rho = \rho(C)$ and $C = \left(\frac{\alpha_s}{h\beta_s} I + M \right)^{-1} N$.

For more details, see [25,12]. Let

$$\rho_{Che} = \rho \left(\sum_{m=0}^{\alpha} y_m C^m \right) = \frac{1}{U_\alpha(1/\rho)} = \frac{2(\omega-1)^{\alpha/2}}{1 + (\omega-1)^\alpha}, \quad (31)$$

where $\{y_m\}_{m=0}^\alpha$ are the coefficients of the α -th Chebyshev polynomial given by (20) and ω is define by (27). In **Figure 1**, we plot the curves of the ρ_{Che} , ρ_{Jac} and ρ_{SOR} as functions of $\rho = \rho(C)$ for $\rho \in (0,1)$ and $\alpha = 4, 8, 11$.

One can see clearly from these four panels that the convergence speed of the Acc-Jacobi WR iteration method is sharper than that of the classical Jacobi WR iteration, while blunter than that of the optimal SOR method. Moreover, as the argument α becomes larger, the spectral radiuses of the *Acc-Jacobi* and the optimal SOR methods become nearer.

4. Numerical Results

To validate our theoretical results given in Section 3, we consider the following problem:

$$\begin{cases} y'(t) + Hy(t) = f(t), t \in [0, t_f], \\ y(0) = y_0, \end{cases} \quad (32)$$

where $f(t) = \left(\sin\left(\frac{j}{j+1}t\right) \right)_{j=1,2,\dots,100}^T$, $y_0 = (1, 1, \dots, 1)^T$ (33)

and $H \in \mathbb{R}^{100 \times 100}$ is defined by

$$\begin{aligned} H &= \begin{pmatrix} A & -B & & \\ -B & A & \ddots & \\ & \ddots & \ddots & -B \\ & & -B & A \end{pmatrix}, \\ \text{with } A &= \begin{pmatrix} 100 & -25 & 0 & 0 \\ -25 & 100 & -25 & 0 \\ 0 & -25 & 100 & -25 \\ 0 & 0 & -25 & 100 \end{pmatrix} \\ \text{and } B &= \begin{pmatrix} 28.375 & 1 & 0 & 0 \\ 1 & 28.375 & 1 & 0 \\ 0 & 1 & 28.375 & 1 \\ 0 & 0 & 1 & 28.375 \end{pmatrix}. \end{aligned} \quad (34)$$

One can verify that the eigenvalue of the matrix H defined by (34) changes from 4×10^{-4} to 194, and thus system (33) is really stiff. To perform the discrete-time WR iteration, we choose the backward Euler method with step size $h = 0.02$. Throughout all of our experiments,

we choose $\alpha = 5$ and $N = 250$, where $N = \frac{t_f}{h}$ is the total steps when the time interval $[0, t_f]$ is discretized by step size h . We note that the discrete-time optimal SOR WR iteration is a typical sequential algorithm, and therefore we only focus on comparing the convergence speed of the *Acc-Jacobi* WR iteration defined by (5) and

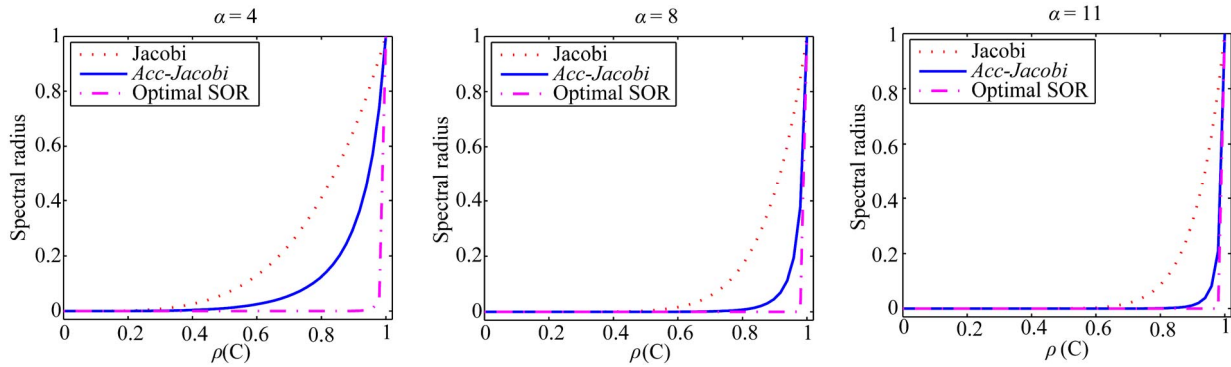


Figure 1. The spectral radius of the three discrete-time WR iterative methods; from left to right: $\alpha = 4, 8, 11$.

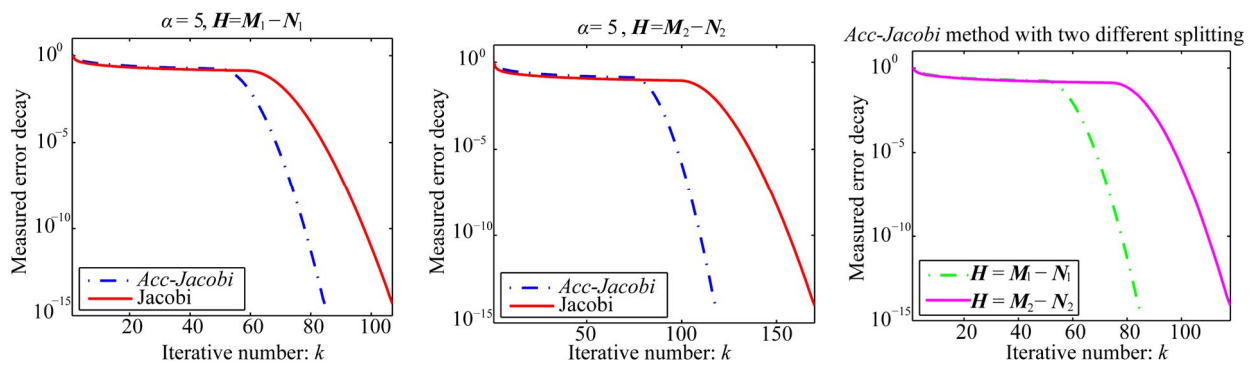


Figure 2. Measured convergence speed of the *Acc-Jacobi* and discrete-time *Jacobi* WR methods.

Table 1. Coefficients $\{v_m\}_{m=0}^5$ for splitting $H = M_1 - N_1$ and $H = M_2 - N_2$.

splitting	coefficients $\{v_m\}_{m=0}^5$					
	v_0	v_1	v_2	v_3	v_4	v_5
$H = M_1 - N_1$	0	0.041 468 677 620 31	0	-0.561 375 694 613 35	0	1.519 907 016 993 04
$H = M_2 - N_2$	0	0.121 948 099 474 61	0	-1.097 542 735 328 93	0	1.975 594 635 854 32

the Jacobi WR iteration. To make a fair comparison, the discrete-time Jacobi WR iteration is performed as (29). The measured error at iteration k for these two methods is defined as:

- **Jacobi:** $\text{Err}_{Jac}(k) = \max_{1 \leq n \leq N} \|\tilde{y}_n^{(k)} - \tilde{y}_n\|_\infty$, where $\{\tilde{y}_n^{(k)}\}_{n=1}^N$ and $\{\tilde{y}_n\}_{n=1}^N$ are the k -th iterative solution and the convergent solution of the discrete-time Jacobi WR method, respectively.
- **Acc-Jacobi:** $\text{Err}_{Acc-Jac}(k) = \max_{1 \leq n \leq N} \|y_n^{(k)} - y_n\|_\infty$, where $\{y_n^{(k)}\}_{n=1}^N$ and $\{y_n\}_{n=1}^N$ are the k -th iterative solution and the convergent solution of the Acc-Jacobi method, respectively.

We choose two different splitting $H = M_1 - N_1 = M_2 - N_2$ with $M_1 = \text{diag}(A, A, \dots, A)$,

$M_2 = \text{diag}(100, 100, \dots, 100)$ and $N_1 = M_1 - H$, $N_2 = M_2 - H$. And by computer it is easy to get

$$\rho\left(\left(\frac{1}{h}I + M_1\right)^{-1} N_1\right) \approx 0.543 6,$$

$$\text{and } \rho\left(\left(\frac{1}{h}I + M_2\right)^{-1} N_2\right) \approx 0.666 7.$$

For the splitting $H = M_1 - N_1$ and $H = M_2 - N_2$, the coefficients $\{v_m\}_{m=0}^5$ are given in Table 1.

In Figure 2, we plot in the left and middle panels the measured error decay of the two methods with splitting $H = M_1 - N_1$ and $H = M_2 - N_2$, respectively, where one can see clearly that the convergence speed of the *Acc-Jacobi* method is much sharper than that of the *Jacobi* method. For these two splitting $H = M_1 - N_1$

and $H = M_2 - N_2$ with $N_2 \geq N_1$, theorem 1 predicts that the *Acc-Jacobi* WR method with the former choice shall converge faster than the case with the latter one. This theoretical conclusion is clearly depicted in the right panel of **Figure 2**.

5. Acknowledgments

This paper is supported by the the NSF of Sichuan University of Science and Engineering (2010XJKRL005, 2009JKRL011) and the Sichuan Provincial Education Department Foundation of China (10ZB098). The authors are grateful to the anonymous referees for the careful reading of a preliminary version of the manuscript and their valuable suggestions and comments, which really improve the quality of this paper.

6. References

- [1] C. Lubich and A. Ostermann, "Multi-grid Dynamic Iteration for Parabolic Equations," *BIT Numerical Mathematics*, Vol. 27, No. 2, 1987, pp. 216-234. doi:10.1007/BF01934186
- [2] E. Lelarasmee, A. E. Ruehli and A. L. Sangiovanni-Vincentelli, "The Waveform Relaxation Methods for Time-domain Analysis of Large Scale Integrated Circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 1, No. 3, 1982, pp. 131-145. doi:10.1109/TCAD.1982.1270004
- [3] U. Miekka and O. Nevanlinna, "Convergence of Dynamic Iteration Methods for Initial Value Problems," *SIAM Journal on Scientific and Statistical Computing*, Vol. 8, No. 4, 1987, pp. 459-482. doi:10.1137/0908046
- [4] U. Miekka and O. Nevanlinna, "Sets of Convergence and Stability Regions," *BIT Numerical Mathematics*, Vol. 27, No.4, 1987, pp. 554-584. doi:10.1007/BF01937277
- [5] U. Miekka, "Dynamic Iteration Methods Applied to linear DAE Systems," *Journal of Computational and Applied Mathematics*, Vol. 25, No. 2, 1989, pp. 133-151. doi:10.1016/0377-0427(89)90044-7
- [6] O. Nevanlinna, "Remarks on Picard-Lindelöf Iteration, Part I," *BIT Numerical Mathematic*, Vol. 29, No. 2, 1989, pp. 328-346.
- [7] O. Nevanlinna, "Remarks on Picard-Lindelöf Iteration, Part II," *BIT Numerical Mathematic*, Vol. 29, No. 3, 1989, pp. 535-562. doi:10.1007/BF02219239
- [8] O. Nevanlinna, "Linear Acceleration of Picard-Lindelöf Iteration," *Numerische Mathematik*, Vol. 57, No. 1, 1990, pp. 147-156. doi:10.1007/BF01386404
- [9] S. Vandewalle, "Parallel Multigrid Waveform Relaxation for Parabolic Problems," B. G. Teubner, Stuttgart, 1993.
- [10] J. Janssen and S. Vandewalle, "Multigrid Waveform Relaxation of Spatial Finite Element Meshes: The Continuous-Time Case," *SIAM Journal on Numerical Analysis*, Vol. 33, No. 2, 1996, pp. 456-474. doi:10.1137/0733024
- [11] J. Y. Pan and Z. Z. Bai, "On the Convergence of Waveform Relaxation Methods for Linear Initial Value Problems," *Journal of Computational Mathematics*, Vol. 22, No. 5, 2004, pp. 681-698.
- [12] J. Sand and K. Burrage, "A Jacobi Waveform Relaxation Method for ODEs," *SIAM Journal on Scientific Computing*, Vol. 20, No. 2, 1998, pp. 534-552. doi:10.1137/S1064827596306562
- [13] J. Wang and Z. Z. Bai, "Convergence Analysis of Two-stage Waveform Relaxation Method for the Initial Value Problems," *Journal of Applied Mathematics and Computing*, Vol. 172, No. 2, 2006, pp. 797-808. doi:10.1016/j.amc.2004.11.031
- [14] A. Bellen, Z. Jackiewicz and M. Zennaro, "Contractivity of Waveform Relaxation Runge-Kutta Iterations and Related Limit Methods for Dissipative Systems in the Maximum Norm," *SIAM Journal on Numerical Analysis*, Vol. 31, No. 2, 1994, pp. 499-523. doi:10.1137/0731027
- [15] L. Galeone and R. Garrappa, "Convergence Analysis of Time-Point Relaxation Iterates for Linear Systems of Differential Equations," *Journal of Computational and Applied Mathematics*, Vol. 80, No. 2, 1997, pp. 183-195. doi:10.1016/S0377-0427(97)00004-6
- [16] R. Garrappa, "An Analysis of Convergence for Two-stage Waveform Relaxation Methods," *Journal of Computational and Applied Mathematics*, Vol. 169, No. 2, 2004, pp. 377-392. doi:10.1016/j.cam.2003.12.031
- [17] J. Janssen and S. Vandewalle, "Multigrid Waveform Relaxation on Spatial Finite Element Meshes: The Discrete-Time Case," *SIAM Journal on Scientific Computing*, Vol. 17, No. 1, 1996, pp. 133-155. doi:10.1137/0917011
- [18] K. Burrage, Z. Jackiewicz and B. Welfert, "Block-Toeplitz Preconditioning for Static and Dynamic Linear Systems," *Linear Algebra and its Applications*, Vol. 279, No. 1-3, 1998, pp. 51-74. doi:10.1016/S0024-3795(98)00007-X
- [19] J. M. Bahi, K. Rhofir and J. C. Miellou, "Parallel Solution of Linear DAEs by Multisplitting Waveform Relaxation Methods," *Linear Algebra and its Applications*, Vol. 332-334, 2001, pp. 181-196. doi:10.1016/S0024-3795(00)00199-3
- [20] C. W. Gear, "Massive Parallelism Across Space in ODEs," *Applied Numerical Mathematics*, Vol. 11, No. 1-3, 1993, pp. 27-43. doi:10.1016/0168-9274(93)90038-S
- [21] R. S. Varga, "Matrix Iterative Analysis," Springer Verlag, Berlin, New York, 2000. doi:10.1007/978-3-642-05156-2
- [22] Z. Z. Bai, "Parallel Matrix Multisplitting Block Relaxation Iteration Methods," *Mathematica Numerica Sinica*, Vol. 17, No. 3, 1995, pp. 238-252.
- [23] D. W. Young, "Iterative Solution of Large Linear Systems," Academic Press, New York, 1971.
- [24] J. C. Mason and D. C. Handscomb, "Chebyshev Polynomials," Chapman & Hall/CRC, Florida, 2003.
- [25] J. Janssen and S. Vandewalle, "On SOR Waveform Relaxation Methods," *SIAM Journal on Numerical Analysis*, Vol. 34, No. 6, 1997, pp. 2456-2481. doi:10.1137/S0036142995294292