

## ON CONTRACTION AND SEMI-CONTRACTION FACTORS OF GSOR METHOD FOR AUGMENTED LINEAR SYSTEMS\*

Fang Chen

*Department of Mathematical Sciences, Xi'an Jiaotong University, Xi'an 710049, China*  
*School of Science, Xi'an University of Post and Telecommunications, Xi'an 710121, China*  
*Email: chenfreesky@gmail.com*

Yao-Lin Jiang

*Department of Mathematical Sciences, Xi'an Jiaotong University, Xi'an 710049, China*  
*Email: yljjiang@mail.xjtu.edu.cn*

Bing Zheng

*School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China*  
*Email: bzheng@lzu.edu.cn*

### Abstract

The *generalized successive overrelaxation (GSOR)* method was presented and studied by Bai, Parlett and Wang [Numer. Math. 102(2005), pp.1-38] for solving the augmented system of linear equations, and the optimal iteration parameters and the corresponding optimal convergence factor were exactly obtained. In this paper, we further estimate the contraction and the semi-contraction factors of the GSOR method. The motivation of the study is that the convergence speed of an iteration method is actually decided by the contraction factor but not by the spectral radius in finite-step iteration computations. For the nonsingular augmented linear system, under some restrictions we obtain the contraction domain of the parameters involved, which guarantees that the contraction factor of the GSOR method is less than one. For the singular but consistent augmented linear system, we also obtain the semi-contraction domain of the parameters in a similar fashion. Finally, we use two numerical examples to verify the theoretical results and the effectiveness of the GSOR method.

*Mathematics subject classification:* 65F10, 65F50; CR: G1.3.

*Key words:* Contraction and semi-contraction factors, Augmented linear system, GSOR method, Convergence.

### 1. Introduction

We study an iterative solution of the augmented linear system

$$\begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ q \end{pmatrix}, \quad \text{or} \quad \mathcal{A}z = f, \quad (1.1)$$

where  $A \in \mathbb{R}^{m \times m}$  is symmetric positive definite, and  $B \in \mathbb{R}^{m \times n}$  is a rectangular matrix. Here  $m \geq n$ , and  $b \in \mathbb{R}^m$  and  $q \in \mathbb{R}^n$  are given vectors, and  $x \in \mathbb{R}^m$  and  $y \in \mathbb{R}^n$  are unknown vectors, respectively. We use  $B^T$  to denote the transpose of the matrix  $B$ . When  $B$  is of full column-rank, we know that the augmented linear system (1.1) has a unique solution. When

---

\* Received July 31, 2009 / Revised version received November 13, 2009 / Accepted December 7, 2009 /  
Published online August 9, 2010 /

$B$  is rank-deficient and  $q \in R(B^T)$  (the range of  $B^T$ ), the augmented linear system (1.1) has infinitely many solutions; which is called the singular but consistent augmented linear system.

The augmented linear system (1.1) results from a wide variety of scientific and engineering applications such as mixed and hybrid finite element approximations of the elliptic problems, Stokes equations, weighted least-squares problems, computer graphics, electronic networks and others; see [1, 2]. The augmented linear system is also called as a saddle point problem, or a *Karush-Kuhn-Tucker* (**KKT**) system. Recently, the augmented linear system has attracted more and more researchers and various kinds of iteration methods have been established and discussed. For example, the Uzawa-type methods [13, 15], the preconditioned Krylov subspace methods [7, 9, 18], the relaxation methods [8, 10, 14, 17], and the Hermitian and skew-Hermitian splitting methods [3–6, 12], etc. Moreover, the singular augmented linear system has been specially studied in [11, 19].

The oldest and famous iteration method for solving the augmented linear system is the Uzawa method [1]. Gloub et al. proposed an SOR-like method for solving the linear system (1.1) in [16]. Based on this idea, Bai et al. established and discussed the GSOR method in [8] and obtained the optimal parameters and the corresponding optimal convergence factor; see also [10].

The GSOR method has the following form.

**Method 1.1.** ([8]) (*The GSOR Method*).

Let  $Q \in \mathbb{R}^{n \times n}$  be a symmetric and nonsingular matrix. Given initial vectors  $x^{(0)} \in \mathbb{R}^m$  and  $y^{(0)} \in \mathbb{R}^n$ , and two relaxation factors  $\omega, \tau$  with  $\omega, \tau \neq 0$ . For  $k = 0, 1, 2, \dots$  until the iteration sequence  $\{(x^{(k)T}, y^{(k)T})^T\}$  is convergent, compute

$$\begin{cases} x^{(k+1)} = (1 - \omega)x^{(k)} + \omega A^{-1}(b - By^{(k)}), \\ y^{(k+1)} = y^{(k)} + \tau Q^{-1}(B^T x^{(k+1)} + q). \end{cases}$$

Here,  $Q$  is an approximate (preconditioning) matrix of the Schur complement matrix  $B^T A^{-1} B$ .

We know that an iteration method is convergent when the spectral radius of the corresponding iteration matrix is less than one. As a matter of fact, the convergence speed of an iteration method is, however, decided by the contraction factor, but not by the spectral radius in practical computations. Therefore, to estimate the contraction factor of an iteration method is a practically important task.

In this paper, firstly, we give the iteration matrix of the GSOR method and introduce a new norm. According to this norm, we propose the concept about the contraction factor of the GSOR method. Usually, it is difficult to obtain the optimal parameters which minimize the contraction factor. Hence, we turn to estimate an upper bound of the contraction factor proposed. The domain makes the upper bound be less than one. Moreover, we extend these results to the singular but consistent augmented linear system.

The paper is organized as follows. In Section 2, the contraction and semi-contraction factors of the GSOR method are established, and the convergence and the semi-convergence of the GSOR method are analyzed. In Section 3, the domains of parameters which guarantee the contraction and the semi-contraction factors to be less than one are obtained. Numerical examples are given in Section 4.

### 2. Formulas of Contraction and Semi-Contraction Factors

For the GSOR method, we compute  $z^{(k+1)}$  from  $z^{(k)}$  by

$$z^{(k+1)} = \mathcal{H}(\omega, \tau)z^{(k)} + \mathcal{M}(\omega, \tau)^{-1}f,$$

where the iteration matrix  $\mathcal{H}(\omega, \tau)$  can be expressed as

$$\mathcal{H}(\omega, \tau) = \begin{pmatrix} (1 - \omega)I & -\omega A^{-1}B \\ (1 - \omega)\tau Q^{-1}B^T & I - \omega\tau Q^{-1}B^T A^{-1}B \end{pmatrix},$$

$$\mathcal{M}(\omega, \tau) = \begin{pmatrix} \frac{1}{\omega}A & 0 \\ -B^T & \frac{1}{\tau}Q \end{pmatrix},$$

with  $I$  being the identity matrix. Let

$$\mathcal{N}(\omega, \tau) = \mathcal{M}(\omega, \tau) - \mathcal{A} = \begin{pmatrix} (\frac{1}{\omega} - 1)A & -B \\ 0 & \frac{1}{\tau}Q \end{pmatrix}.$$

Then  $\mathcal{A} = \mathcal{M}(\omega, \tau) - \mathcal{N}(\omega, \tau)$  is a splitting of the matrix  $\mathcal{A}$ . When the spectral radius of the iteration matrix  $\mathcal{H}(\omega, \tau)$  is less than one, the GSOR iteration method is convergent; see [8].

Assume that the matrix  $Q$  is symmetric positive definite. We define

$$G = \begin{pmatrix} A^{\frac{1}{2}} & 0 \\ 0 & Q^{\frac{1}{2}} \end{pmatrix}$$

and let  $\tilde{\mathcal{H}}(\omega, \tau) = G\mathcal{H}(\omega, \tau)G^{-1}$ . Then

$$\tilde{\mathcal{H}}(\omega, \tau) = \begin{pmatrix} (1 - \omega)I & -\omega\tilde{B} \\ (1 - \omega)\tau\tilde{B}^T & I - \omega\tau\tilde{B}^T\tilde{B} \end{pmatrix},$$

where  $\tilde{B} = A^{-\frac{1}{2}}BQ^{-\frac{1}{2}}$ .

Now, we introduce a vector norm  $|||x||| = \|Gx\|_2$  (for all  $x \in \mathbb{R}^n$ ). The corresponding matrix norm is  $|||X||| = \|GXG^{-1}\|_2$  (for all  $X \in \mathbb{R}^{n \times n}$ ); see [5]. At this situation,

$$|||\mathcal{H}(\omega, \tau)||| = |||\tilde{\mathcal{H}}(\omega, \tau)|||_2.$$

It is easy to know that the rank of  $\tilde{B}$  is the same as that of  $B$ . In the following, we define the contraction and the semi-contraction factors  $|||\mathcal{H}(\omega, \tau)|||$  according to two cases.

Case (a)  $B$  is of full column-rank. We assume that the matrix  $\tilde{B}$  has the following singular value decomposition:

$$U\tilde{B}V^* = \Sigma = \begin{pmatrix} \Lambda & \\ & 0 \end{pmatrix},$$

where  $U$  and  $V$  are unitary matrices, and  $V^*$  is the conjugate transpose of  $V$ . Denote by

$$\mathcal{P} = \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}.$$

Then from the structure of the matrix  $\mathcal{P}$  we know that  $\mathcal{P}$  is also unitary. Hence,

$$\begin{aligned} \hat{\mathcal{H}}(\omega, \tau) &= \mathcal{P}\tilde{\mathcal{H}}(\omega, \tau)\mathcal{P}^* \\ &= \begin{pmatrix} (1-\omega)I & -\omega U\tilde{B}V^* \\ (1-\omega)\tau V\tilde{B}^T U^* & I - \omega\tau V\tilde{B}^T\tilde{B}V^* \end{pmatrix} \\ &= \begin{pmatrix} (1-\omega)I & -\omega\Sigma \\ (1-\omega)\tau\Sigma^T & I - \omega\tau\Sigma^T\Sigma \end{pmatrix} \\ &= \begin{pmatrix} (1-\omega)I & 0 & -\omega\Lambda \\ 0 & (1-\omega)I & 0 \\ (1-\omega)\tau\Lambda & 0 & I - \omega\tau\Lambda^2 \end{pmatrix}. \end{aligned}$$

It follows that the matrix  $\hat{\mathcal{H}}(\omega, \tau)$  is unitarily similar to the matrix  $\tilde{\mathcal{H}}(\omega, \tau)$ , and

$$\|\tilde{\mathcal{H}}(\omega, \tau)\|_2 = \|\hat{\mathcal{H}}(\omega, \tau)\|_2.$$

Now, we define the contraction factor  $\|\|\mathcal{H}(\omega, \tau)\|\|$  of the GSOR method as

$$\|\|\mathcal{H}(\omega, \tau)\|\| = \|\|\mathcal{H}(\omega, \tau)\|\| = \|\hat{\mathcal{H}}(\omega, \tau)\|_2.$$

Therefore, when  $\|\|\mathcal{H}(\omega, \tau)\|\| < 1$ , the GSOR iteration method is convergent.

Case (b)  $B$  is rank-deficient. Let  $rank(B) = r$ . The singular value decomposition of the matrix  $\tilde{B}$  has the form

$$U\tilde{B}V^* = \Sigma = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix},$$

where  $U$  and  $V$  are unitary matrices, and  $\Lambda$  is a  $r$ -by- $r$  diagonal matrix. Define

$$\mathcal{P} = \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}.$$

Then

$$\begin{aligned} \hat{\mathcal{H}}(\omega, \tau) &= \mathcal{P}\tilde{\mathcal{H}}(\omega, \tau)\mathcal{P}^* \\ &= \begin{pmatrix} (1-\omega)I & -\omega U\tilde{B}V^* \\ (1-\omega)\tau V\tilde{B}^T U^* & I - \omega\tau V\tilde{B}^T\tilde{B}V^* \end{pmatrix} \\ &= \begin{pmatrix} (1-\omega)I & -\omega\Sigma \\ (1-\omega)\tau\Sigma^T & I - \omega\tau\Sigma^T\Sigma \end{pmatrix} \\ &= \begin{pmatrix} (1-\omega)I & 0 & -\omega\Lambda & 0 \\ 0 & (1-\omega)I & 0 & 0 \\ (1-\omega)\tau\Lambda & 0 & I - \omega\tau\Lambda^2 & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \\ &:= \begin{pmatrix} \check{\mathcal{H}}(\omega, \tau) & 0 \\ 0 & I \end{pmatrix}. \end{aligned}$$

For this case, we define the semi-contraction factor  $\|\|\mathcal{H}(\omega, \tau)\|\|$  of the GSOR method as

$$\|\|\mathcal{H}(\omega, \tau)\|\| = \|\check{\mathcal{H}}(\omega, \tau)\|_2.$$

Therefore, when  $|||\mathcal{H}(\omega, \tau)||| < 1$ , the GSOR iteration method is semi-convergent.

According to the above analyses, we may need to solve the minimization problem  $\min_{\omega, \tau} |||\mathcal{H}(\omega, \tau)|||$  to obtain the optimal iteration parameters  $\omega$  and  $\tau$ . These minimization problems can be expressed as

$$\min_{\omega, \tau} \|\hat{\mathcal{H}}(\omega, \tau)\|_2$$

and

$$\min_{\omega, \tau} \|\check{\mathcal{H}}(\omega, \tau)\|_2.$$

In fact, solving these two problems is generally very difficult and even impossible.

In next section, we will derive upper bounds for  $\|\hat{\mathcal{H}}(\omega, \tau)\|_2$  and  $\|\check{\mathcal{H}}(\omega, \tau)\|_2$ . Further, the ranges of parameters which make upper bounds be less than one are obtained.

### 3. Descriptions of Contraction Domains

In this section, we first compute the contraction and the semi-contraction factors, and then get the corresponding upper bounds. Firstly, to estimate an upper bound for the contraction factor  $|||\mathcal{H}(\omega, \tau)|||$  of the GSOR method. We rewrite the matrix  $\hat{\mathcal{H}}(\omega, \tau)$  as

$$\begin{aligned} \hat{\mathcal{H}}(\omega, \tau) &= \begin{pmatrix} (1-\omega)I & 0 & -\omega\Lambda \\ 0 & (1-\omega)I & 0 \\ -\omega\Lambda & 0 & I - \omega\tau\Lambda^2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ (\omega + \tau - \omega\tau)\Lambda & 0 & 0 \end{pmatrix} \\ &:= \mathcal{W}_1(\omega, \tau) + \mathcal{W}_2(\omega, \tau). \end{aligned}$$

Then it holds that

$$\begin{aligned} |||\mathcal{H}(\omega, \tau)||| &= |||\mathcal{H}(\omega, \tau)||| = \|\check{\mathcal{H}}(\omega, \tau)\|_2 = \|\hat{\mathcal{H}}(\omega, \tau)\|_2 \\ &\leq \|\mathcal{W}_1(\omega, \tau)\|_2 + \|\mathcal{W}_2(\omega, \tau)\|_2 \\ &= \max\{|1-\omega|, |\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\} + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}, \end{aligned} \tag{3.1}$$

where

$$\lambda_j = \frac{1}{2} \left[ 2 - \omega - \omega\tau\mu_j \pm \sqrt{\omega^2(\tau\mu_j - 1)^2 + 4\omega^2\mu_j} \right], \quad j = 1, \dots, n.$$

Here,  $\mu_j (j = 1, 2, \dots, n)$  are the eigenvalues of the matrix  $Q^{-1}B^T A^{-1}B$ . Denote by  $\mu_{max}$  and  $\mu_{min}$  the largest and the smallest eigenvalues of the matrix  $Q^{-1}B^T A^{-1}B$ , respectively. Because  $\tilde{B}^T \tilde{B}$  is similar to  $Q^{-1}B^T A^{-1}B$ , we know that  $\tilde{B}^T \tilde{B}$  and  $Q^{-1}B^T A^{-1}B$  have the same eigenvalues. Hence, to guarantee the convergence of the GSOR method, we need to have  $|||\mathcal{H}(\omega, \tau)||| < 1$ . Now, the following theorem gives such a condition.

**Theorem 3.1.** *Let  $A \in \mathbb{R}^{m \times m}$  and  $Q \in \mathbb{R}^{n \times n}$  be symmetric positive definite, and  $B \in \mathbb{R}^{m \times n}$  be of full column-rank. Denote the largest and the smallest eigenvalues of matrix  $Q^{-1}B^T A^{-1}B$  by  $\mu_{max}$  and  $\mu_{min}$ , respectively. Then the contraction factor  $|||\mathcal{H}(\omega, \tau)|||$  of the GSOR method is less than one, provided that*

(a) the parameters  $\omega$  and  $\tau$  satisfy  $2 - \omega - \omega\tau\mu_{min} \leq 0$  and

$$\omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} + 2|\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < 4 - \omega - \omega\tau\mu_{max};$$

(b) the parameters  $\omega$  and  $\tau$  satisfy  $2 - \omega - \omega\tau\mu_{max} \geq 0$  and

$$\begin{cases} \omega\sqrt{(\tau\mu_{min} - 1)^2 + 4\mu_{min}} + 2|\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < \omega + \omega\tau\mu_{min}, \\ \omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} + 2|\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < \omega + \omega\tau\mu_{max}; \end{cases}$$

(c) the parameters  $\omega$  and  $\tau$  satisfy  $2 - \omega - \omega\tau\mu_{min} > 0$ ,  $2 - \omega - \omega\tau\mu_{max} < 0$ , and

$$\begin{cases} \omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} + 2|\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < 4 - \omega - \omega\tau\mu_{max}, \\ \omega\sqrt{(\tau\mu_{min} - 1)^2 + 4\mu_{min}} + 2|\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < \omega + \omega\tau\mu_{min}. \end{cases}$$

*Proof.* From formula (3), we know that the parameters  $\omega$  and  $\tau$  satisfy

$$|1 - \omega| < 1 \quad \text{and} \quad |\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < 1. \tag{3.2}$$

By assuming  $\tau > 0$ ,  $\mu_{min} > 0.125$  and  $\mu_{max} > 1$ , from (3.2) we see that the parameters must satisfy

$$\begin{cases} 0 < \omega < 1, \\ \frac{-1 - \omega\sqrt{\mu_{max}}}{\sqrt{\mu_{max}} - \omega\sqrt{\mu_{max}}} < \tau < \frac{1 - \omega\sqrt{\mu_{max}}}{\sqrt{\mu_{max}} - \omega\sqrt{\mu_{max}}} < 1 \end{cases} \tag{3.3}$$

or

$$\begin{cases} 1 < \omega < 2, \\ \frac{-1 - \omega\sqrt{\mu_{max}}}{\sqrt{\mu_{max}} - \omega\sqrt{\mu_{max}}} > \tau > \frac{1 - \omega\sqrt{\mu_{max}}}{\sqrt{\mu_{max}} - \omega\sqrt{\mu_{max}}} > 1. \end{cases} \tag{3.4}$$

Let

$$f(\omega, \tau, \mu) = \frac{1}{2} \left( |2 - \omega - \omega\tau\mu| + \omega\sqrt{(\tau\mu - 1)^2 + 4\mu} \right), \quad \mu \in [\mu_{min}, \mu_{max}].$$

Then

$$\begin{aligned} f(\omega, \tau, \mu) &\geq \frac{1}{2} [|2 - \omega - \omega\tau\mu| + |\omega\tau\mu - \omega|] \\ &\geq \frac{1}{2} |2 - \omega - \omega\tau\mu + \omega\tau\mu - \omega| = |1 - \omega|, \end{aligned}$$

Which gives that

$$\begin{aligned} |||\mathcal{H}(\omega, \tau)||| &\leq \max_{\mu \in [\mu_{min}, \mu_{max}]} \{f(\omega, \tau, \mu)\} + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}} \\ &:= g(\omega, \tau). \end{aligned}$$

After direct calculations, we obtain

$$f(\omega, \tau, \mu) = \begin{cases} \frac{1}{2} [2 - \omega - \omega\tau\mu + \omega\sqrt{(\tau\mu - 1)^2 + 4\mu}], & \text{for } 2 - \omega - \omega\tau\mu \geq 0, \\ \frac{1}{2} [-(2 - \omega - \omega\tau\mu) + \omega\sqrt{(\tau\mu - 1)^2 + 4\mu}], & \text{for } 2 - \omega - \omega\tau\mu < 0. \end{cases}$$

Consequently, we have

$$\frac{\partial f(\omega, \tau, \mu)}{\partial \mu} = \begin{cases} \frac{-\omega\tau}{2} \left[ 1 - \frac{(\tau\mu + \frac{2}{\tau} - 1)}{\sqrt{(\tau\mu - 1)^2 + 4\mu}} \right], & \text{for } 2 - \omega - \omega\tau\mu > 0, \\ \frac{\omega\tau}{2} \left[ 1 + \frac{(\tau\mu + \frac{2}{\tau} - 1)}{\sqrt{(\tau\mu - 1)^2 + 4\mu}} \right], & \text{for } 2 - \omega - \omega\tau\mu < 0 \end{cases}$$

and

$$\left(\tau\mu + \frac{2}{\tau} - 1\right)^2 - (\tau\mu - 1)^2 - 4\mu = \frac{4}{\tau^2}(1 - \tau).$$

And under the condition (3.3), it holds that

$$\frac{\partial f(\omega, \tau, \mu)}{\partial \mu} \begin{cases} > 0, & \text{for } 2 - \omega - \omega\tau\mu > 0, \\ > 0, & \text{for } 2 - \omega - \omega\tau\mu < 0. \end{cases}$$

Hence, we get

$$g(\omega, \tau) = f(\omega, \tau, \mu_{max}) + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}.$$

Under the condition (3.4), it holds that

$$\frac{\partial f(\omega, \tau, \mu)}{\partial \mu} \begin{cases} < 0, & \text{for } 2 - \omega - \omega\tau\mu > 0, \\ > 0, & \text{for } 2 - \omega - \omega\tau\mu < 0. \end{cases}$$

Hence, we get

$$g(\omega, \tau) = \max\{f(\omega, \tau, \mu_{min}), f(\omega, \tau, \mu_{max})\} + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}.$$

Now, we define the functions  $g_1(\omega, \tau)$  and  $g_2(\omega, \tau)$  by

$$g_1(\omega, \tau) = \frac{1}{2} \left[ |2 - \omega - \omega\tau\mu_{min}| + \omega\sqrt{(\tau\mu_{min} - 1)^2 + 4\mu_{min}} \right] + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}},$$

and

$$g_2(\omega, \tau) = \frac{1}{2} \left[ |2 - \omega - \omega\tau\mu_{max}| + \omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} \right] + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}.$$

Easily, it holds that

$$g(\omega, \tau) = \max\{g_1(\omega, \tau), g_2(\omega, \tau)\}.$$

Now we discuss the contraction conditions on the parameters  $\omega$  and  $\tau$ .

(a) Assume  $2 - \omega - \omega\tau\mu_{min} \leq 0$ . For this case, it holds that

$$2 - \omega - \omega\tau\mu_{max} \leq 0.$$

Consequently,

$$g_1(\omega, \tau) = \frac{1}{2} \left[ -(2 - \omega - \omega\tau\mu_{min}) + \omega\sqrt{(\tau\mu_{min} - 1)^2 + 4\mu_{min}} \right] + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}$$

and

$$g_2(\omega, \tau) = \frac{1}{2} \left[ -(2 - \omega - \omega\tau\mu_{max}) + \omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} \right] + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}.$$

Let  $h(\tau, \mu) = (\tau\mu - 1)^2 + 4\mu$ . Because

$$\frac{\partial h(\tau, \mu)}{\partial \mu} = 2\tau(\tau\mu - 1) + 4 = 2\tau^2\mu - 2\tau + 4 > 0, \quad \text{for } \mu \in [\mu_{min}, \mu_{max}],$$

we see that the function  $h(\tau, \mu)$  is increasing with respect to  $\mu$  when  $\mu \in [\mu_{min}, \mu_{max}]$ . Further, it holds that

$$h(\tau, \mu_{max}) > h(\tau, \mu_{min}) > 0.$$

Nothing that

$$g_1(\omega, \tau) - g_2(\omega, \tau) = \frac{1}{2}\tau\omega(\mu_{min} - \mu_{max}) + \frac{1}{2}\omega[\sqrt{h(\tau, \mu_{min})} - \sqrt{h(\tau, \mu_{max})}] < 0,$$

we obtain  $g(\omega, \tau) = g_2(\omega, \tau)$ . Therefore,  $g(\omega, \tau) < 1$  if

$$\omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} + 2|\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < 4 - \omega - \omega\tau\mu_{max}.$$

(b) Assume  $2 - \omega - \omega\tau\mu_{max} \geq 0$ . For this case, it holds that

$$2 - \omega - \omega\tau\mu_{min} \geq 0.$$

Consequently,

$$g_1(\omega, \tau) = \frac{1}{2} \left[ (2 - \omega - \omega\tau\mu_{min}) + \omega\sqrt{(\tau\mu_{min} - 1)^2 + 4\mu_{min}} \right] + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}$$

and

$$g_2(\omega, \tau) = \frac{1}{2} \left[ (2 - \omega - \omega\tau\mu_{max}) + \omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} \right] + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}.$$

By direct calculations we get

$$g_1(\omega, \tau) - g_2(\omega, \tau) = \frac{1}{2}\omega \left[ \tau(\mu_{max} - \mu_{min}) + \sqrt{h(\tau, \mu_{min})} - \sqrt{h(\tau, \mu_{max})} \right].$$

Note that

$$\begin{aligned} & \left[ \tau(\mu_{max} - \mu_{min}) + \sqrt{h(\tau, \mu_{min})} \right]^2 - \left[ \sqrt{h(\tau, \mu_{max})} \right]^2 \\ &= 2(\mu_{max} - \mu_{min}) \left[ \tau\sqrt{h(\tau, \mu_{min})} - (\tau^2\mu_{min} - \tau + 2) \right] \end{aligned}$$

and

$$\tau^2h(\tau, \mu_{min}) - (\tau^2\mu_{min} - \tau + 2)^2 = 4(\tau - 1).$$



Therefore,  $g(\omega, \tau) < 1$  if

$$\begin{cases} \omega\sqrt{(\tau\mu_{min} - 1)^2 + 4\mu_{min}} + 2|\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < \omega + \omega\tau\mu_{min}, \\ \omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} + 2|\omega + \tau - \omega\tau|\sqrt{\mu_{max}} < \omega + \omega\tau\mu_{max}. \end{cases}$$

(c) Assume  $2 - \omega - \omega\tau\mu_{min} > 0$  and  $2 - \omega - \omega\tau\mu_{max} < 0$ . For this case, it holds that

$$g_1(\omega, \tau) = \frac{1}{2} \left[ (2 - \omega - \omega\tau\mu_{min}) + \omega\sqrt{(\tau\mu_{min} - 1)^2 + 4\mu_{min}} \right] + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}$$

and

$$g_2(\omega, \tau) = \frac{1}{2} \left[ -(2 - \omega - \omega\tau\mu_{max}) + \omega\sqrt{(\tau\mu_{max} - 1)^2 + 4\mu_{max}} \right] + |\omega + \tau - \omega\tau|\sqrt{\mu_{max}}.$$

It follows that

$$\begin{aligned} &g_1(\omega, \tau) - g_2(\omega, \tau) \\ &= (2 - \omega) - \frac{1}{2}\tau\omega(\mu_{max} + \mu_{min}) + \frac{1}{2}\omega \left[ \sqrt{h(\tau, \mu_{min})} - \sqrt{h(\tau, \mu_{max})} \right] \\ &= (2 - \omega) + \frac{1}{2}\omega\sqrt{h(\tau, \mu_{min})} - \frac{1}{2}\tau\omega(\mu_{max} + \mu_{min}) - \frac{1}{2}\omega\sqrt{h(\tau, \mu_{max})}. \end{aligned}$$

Therefore,  $g(\omega, \tau) < 1$  if  $g_1(\omega, \tau) < 1$  and  $g_2(\omega, \tau) < 1$ .

The above analysis directly leads to the results in this theorem. □

According to Theorem 3.1, similar to the analysis in [14, 19], we can obtain semi-contraction factor of the GSOR method for the singular but consistent augmented linear system.

**Theorem 3.2.** *Let  $A \in \mathbb{R}^{m \times m}$  and  $Q \in \mathbb{R}^{n \times n}$  be symmetric positive definite, and  $B \in \mathbb{R}^{m \times n}$  be rank deficient. Denote the nonzero largest and smallest eigenvalues of matrix  $Q^{-1}B^T A^{-1}B$  by  $\mu_{max}$  and  $\mu_{min}$ , respectively. Then the semi-contraction factor of the GSOR method is less than one under the same conditions in Theorem 3.1.*

### 4. Numerical Experiments

In this section, we use two examples to show the correctness of the estimates about the contraction and semi-contraction factors. In actual computations, our examples are run in MATLAB with machine precision  $10^{-16}$ . The right-hand-side vector is chosen such that the exact solution of the augmented linear system is  $(1, 1, \dots, 1)^T \in \mathbb{R}^{m+n}$ .

**Example 4.1.** ([6]) Consider the augmented linear system (1.1) with its coefficient matrix being of the matrix blocks

$$A = \begin{pmatrix} I \otimes T + T \otimes I & 0 \\ 0 & I \otimes T + T \otimes I \end{pmatrix} \in \mathbb{R}^{2l^2 \times 2l^2} \tag{4.1}$$

and

$$B = \begin{pmatrix} I \otimes F \\ F \otimes I \end{pmatrix} \in \mathbb{R}^{2l^2 \times l^2}, \tag{4.2}$$

where

$$T = \frac{1}{h^2} \cdot \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{l \times l} \quad \text{and} \quad F = \frac{1}{h} \cdot \text{tridiag}(-1, 1, 0) \in \mathbb{R}^{l \times l}, \quad (4.3)$$

with  $\otimes$  being the Kronecker product symbol and  $h = \frac{1}{l+1}$  the discretization meshsize.

For this example,  $m = 2l^2$  and  $n = l^2$ . Hence, the total number of variables in the augmented linear system (1.1) is  $m + n = 3l^2$ . The choices of the matrix  $Q$ , an approximation to the Schur complement  $B^T A^{-1} B$ , are listed in Table 4.1, where  $\lambda_{min}$  and  $\lambda_{max}$  are used to denote the smallest and the largest eigenvalues of the matrix  $A$ , respectively.

Table 4.1: Choices of the matrix  $Q$  for Example 4

Case No.	Matrix $Q$	Description
I	$\frac{1}{v} B^T B$	$v = \sqrt{\lambda_{min} \lambda_{max}}$
II	$\text{tridiag}(B^T \hat{A}^{-1} B)$	$\hat{A} = \text{tridiag}(A)$
III	$\text{tridiag}(B^T A^{-1} B)$	

Table 4.2: Optimal parameters versus contraction factor for Example 4

$m$		128	512	1152	2048
$n$		64	256	576	1024
$m + n$		192	768	1728	3072
Case I	$\omega_*$	0.31	0.19	0.14	0.11
	$\tau_*$	0.40	0.23	0.16	0.12
	$   \mathcal{H}   _*$	0.98732	0.99777	0.99924	0.99967
Case II	$\omega_*$	0.63	0.63	0.63	0.62
	$\tau_*$	1.13	1.10	1.08	1.08
	$   \mathcal{H}   _*$	0.91699	0.95480	0.96885	0.97614
Case III	$\omega_*$	0.68	0.69	0.71	0.70
	$\tau_*$	1.32	1.41	1.42	1.46
	$   \mathcal{H}   _*$	0.88841	0.93172	0.95040	0.96077

In this section, we use  $\omega_*$  and  $\tau_*$  to denote the optimal parameters in the computations. The optimal parameters  $\omega_*$  and  $\tau_*$ , and the corresponding contraction factor  $|||\mathcal{H}(\omega_*, \tau_*)|||$  (denoted as  $|||\mathcal{H}|||_*$  for short) of the GSOR method for Examples 4 and 4 are listed in Tables 4.2 and 4.4, respectively, for different problem sizes  $(m, n)$ .

In Table 4.2, for the same choices of the matrix  $Q$ , when  $m$  and  $n$  increase, the corresponding  $|||\mathcal{H}|||_*$  also increase. According to the matrix  $Q$  in Table 4.1, Case III is the best choice for the GSOR method in our test for Example 4.

**Example 4.2.** ([19]) Consider the augmented linear system (1.1) in which the matrix  $B$  is rank deficient, with its coefficient matrix of the matrix blocks

$$A = \begin{pmatrix} I \otimes T + T \otimes I & 0 \\ 0 & I \otimes T + T \otimes I \end{pmatrix} \in \mathbb{R}^{2l^2 \times 2l^2}$$

and

$$B = \begin{pmatrix} \hat{B} & b_1 & b_2 \end{pmatrix},$$

where  $T = 1h^{-2} \cdot \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{l \times l}$ ,  $\hat{B}$  is of the form (4.2), and

$$b_1 = \hat{B} \begin{pmatrix} e \\ 0 \end{pmatrix}, \quad b_2 = \hat{B} \begin{pmatrix} 0 \\ e \end{pmatrix}, \quad e = (1, 1, \dots, 1)^T \in \mathbb{R}^{l^2/2},$$

with  $F = 1h^{-1} \cdot \text{tridiag}(-1, 1, 0) \in \mathbb{R}^{l \times l}$ .

For Example 4,  $m = 2l^2$  and  $n = l^2 + 2$ . Hence, the total number of variables in the augmented linear system (1.1) is  $m + n = 3l^2 + 2$ . In fact, Example 4 is a technical modification of Example 4. In Table 4.3, we list the four cases of the matrix  $Q$ , where  $\text{Diag}(M, N)$  denotes the block diagonal matrix

$$\text{Diag}(M, N) = \begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix}.$$

We list the results of Example 4 in Table 4.4. For the same choices of the matrix  $Q$ , when  $m$  and  $n$  are increasing, the corresponding  $\|\mathcal{H}\|_*$  are also increasing. According to the matrix  $Q$  in Table 4.3, Case IV is the best choice for the GSOR method in our test.

Table 4.3: Choices of the matrix  $Q$  for Example 4, with  $\hat{Q} = \text{Diag}(\hat{B}^T \hat{A}^{-1} \hat{B}, \tilde{B}^T \tilde{B})$

Case No.	Matrix $Q$	Description
I	$\hat{Q}$	$\hat{A} = \text{tridiag}(A)$
II	$\hat{Q}$	$\hat{A} = \text{diag}(A)$
III	$\text{tridiag}(\hat{Q})$	$\hat{A} = \text{tridiag}(A)$
IV	$\text{tridiag}(\hat{Q})$	$\hat{A} = A$

Table 4.4: Optimal parameters versus semi-contraction factor for Example 4

$m$		128	512	1152	2048
$n$		64	256	576	1024
$m + n$		192	768	1728	3072
Case I	$\omega_*$	0.21	0.07	0.03	0.02
	$\tau_*$	0.24	0.07	0.03	0.02
	$\ \mathcal{H}\ _*$	0.98544	0.99876	0.99977	0.99990
Case II	$\omega_*$	0.13	0.04	0.02	0.01
	$\tau_*$	0.13	0.04	0.02	0.01
	$\ \mathcal{H}\ _*$	0.99567	0.99960	0.99990	0.99997
Case III	$\omega_*$	0.63	0.62	0.63	0.62
	$\tau_*$	1.12	1.11	1.08	1.08
	$\ \mathcal{H}\ _*$	0.91419	0.95361	0.96823	0.97577
Case IV	$\omega_*$	0.68	0.69	0.71	0.70
	$\tau_*$	1.31	1.41	1.42	1.46
	$\ \mathcal{H}\ _*$	0.88477	0.93005	0.94943	0.96017

**Acknowledgments.** Part of this work was done while the first author was visiting the State Key Laboratory of Scientific/Engineering Computing of Chinese Academy of Sciences in 2009. She is very much indebted to Prof. Zhong-Zhi Bai for his kind invitation. This work is supported

by State Key Laboratory of Scientific/Engineering Computing, Chinese Academy of Sciences; The National Key Basic Research Program of China (973 program) under Grant 2005CB321701; The National Natural Science Foundation (No. 10771168); The Natural Science and Technology Development Plan Research Project of Shaanxi Province (No. 2008K01-33); and The Special Plan Research Project of Shaanxi Education Department (No. 09JK716), P.R. China.

## References

- [1] K. Arrow, L. Hurwicz and H. Uzawa, *Studies in Nonlinear Programming*, Stanford University Press, Stanford, 1958.
- [2] Z.-Z. Bai, Structured preconditioners for nonsingular matrices of block two-by-two structures, *Math. Comput.*, **75** (2006), 791-815.
- [3] Z.-Z. Bai and G.H. Golub, Accelerated Hermitian and skew-Hermitian splitting iteration methods for saddle-point problems, *IMA J. Numer. Anal.*, **27** (2007), 1-23.
- [4] Z.-Z. Bai, G.H. Golub and C.-K. Li, Convergence properties of preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite matrices, *Math. Comput.*, **76** (2006), 287-298.
- [5] Z.-Z. Bai, G.H. Golub and M.K. Ng, Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, *SIAM J. Matrix Anal. Appl.*, **24** (2003), 603-626.
- [6] Z.-Z. Bai, G.H. Golub and J.-Y. Pan, Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite linear systems, *Numer. Math.*, **98** (2004), 1-32.
- [7] Z.-Z. Bai and G.-Q. Li, Restrictively preconditioned conjugate gradient methods for systems of linear equations, *IMA J. Numer. Anal.*, **23** (2003), 561-580.
- [8] Z.-Z. Bai, B.N. Parlett and Z.-Q. Wang, On generalized successive overrelaxation methods for augmented linear systems, *Numer. Math.*, **102** (2005), 1-38.
- [9] Z.-Z. Bai and Z.-Q. Wang, Restrictive preconditioners for conjugate gradient methods for symmetric positive definite linear systems, *J. Comput. Appl. Math.*, **187** (2006), 202-226.
- [10] Z.-Z. Bai and Z.-Q. Wang, On parameterized inexact Uzawa methods for generalized saddle point problems, *Linear Algebra Appl.*, **428** (2008), 2900-2932.
- [11] Z.-Z. Bai, L. Wang and J.-Y. Yuan, Weak-convergence theory of quasi-nonnegative splittings for singular matrices, *Appl. Numer. Math.*, **47** (2003), 75-89.
- [12] M. Benzi and G.H. Golub, A preconditioner for generalized saddle point problems, *SIAM J. Matrix Anal. Appl.*, **26**(2004), 20-41.
- [13] J.H. Bramble, J.E. Pasciak and A.T. Vassilev, Analysis of the inexact Uzawa algorithm for saddle point problems, *SIAM J. Numer. Anal.*, **34** (1997), 1072-1092.
- [14] F. Chen and Y.-L. Jiang, A generalization of the inexact parameterized Uzawa methods for saddle point problems, *Appl. Math. Comput.*, **206** (2008), 765-771.
- [15] H.C. Elman and G.H. Golub, Inexact and preconditioned Uzawa algorithms for saddle point problems, *SIAM J. Numer. Anal.*, **31**(1994), 1645-1661.
- [16] G.H. Golub, X. Wu and J.-Y. Yuan, SOR-like methods for augmented systems, *BIT*, **41** (2001), 71-85.
- [17] C.-J. Li, Z. Li, D.J. Evans and T. Zhang, A note on an SOR-like method for augmented systems, *IMA J. Numer. Anal.*, **23** (2003), 581-592.
- [18] J.-F. Yin and Z.-Z. Bai, The restrictively preconditioned conjugate gradient methods on normal residual for block two-by-two linear systems, *J. Comput. Math.*, **26** (2008), 240-249.
- [19] B. Zheng, Z.-Z. Bai and X. Yang, On semi-convergence of parameterized Uzawa methods for singular saddle point problems, *Linear Algebra Appl.*, **431**(2009), 808-817.