

# On using the shifted minimal residual method for quantum-mechanical wave packet simulation

Hiroaki Seito<sup>1</sup>, Takeo Hoshi<sup>2</sup> and Yusaku Yamamoto<sup>1</sup>

<sup>1</sup> The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

<sup>2</sup> Tottori University, 4-101 Koyama-Minami, Tottori 680-8552, Japan

E-mail [yusaku.yamamoto@uec.ac.jp](mailto:yusaku.yamamoto@uec.ac.jp)

Received December 3, 2018, Accepted December 18, 2018

## Abstract

We apply the shifted minimum residual method to a shifted linear system  $(A + i\sigma I)\mathbf{x} = \mathbf{b}$  arising from quantum-mechanical wave packet simulation and discuss its advantages and disadvantages, comparing it with the widely used COCG method. Although the method requires more vector operations per iteration than the COCG method, it enjoys such nice properties as being free from breakdown, monotonic decrease of the residual norm and applicability to Hermitian-plus-shift systems, and can be a method of choice in some cases. Application of the method to the case of multiple shifts is also discussed.

**Keywords** shifted linear system, MINRES method, COCG method, wave packet simulation, performance evaluation

**Research Activity Group** Algorithms for Matrix / Eigenvalue Problems and their Applications

## 1. Introduction

Quantum-mechanical wave packet simulation, which follows the time evolution of a single-particle wave function described by the Schrödinger equation, is important for calculating transportation properties of materials, such as conductivity. Let us consider the space-discretized Schrödinger equation  $i\partial\mathbf{u}/\partial t = H\mathbf{u}$ , where  $\mathbf{u} = \mathbf{u}(t) \in \mathbb{C}^n$  is the wave function and  $H$  is the  $n \times n$  Hamiltonian matrix.  $H$  is real symmetric in some cases, while it is Hermitian in other cases. For time discretization of this equation, the Crank-Nicolson method is often used, because it preserves the norm of the wave function [1]. The resulting scheme can be written as

$$i\frac{\mathbf{u}(t + \Delta t) - \mathbf{u}(t)}{\Delta t} = \frac{1}{2}H\{\mathbf{u}(t) + \mathbf{u}(t + \Delta t)\}, \quad (1)$$

where  $\Delta t$  is the step size. Thus, for time evolution of one step, we have to solve the *shifted linear system*

$$(A + i\sigma I_n)\mathbf{x} = \mathbf{b}, \quad (2)$$

where  $A = H$  is either real symmetric or Hermitian,  $\sigma = -2/\Delta t$  is a real number,  $I_n$  is the identity matrix of order  $n$ ,  $\mathbf{x} = \mathbf{u}(t + \Delta t)$  and  $\mathbf{b} = (-H - (2i/\Delta t)I_n)\mathbf{u}(t)$ . By the way, the term *shifted linear system* is often used to refer to a set of linear systems of the form (2) with *multiple* values of  $\sigma$ . In this paper, however, we use this term to point to *one* linear system with *a single* value of  $\sigma$ . To refer to a set of systems, we use an expression like *shifted linear systems with multiple shifts*.

It is well known that the shift structure of (2) can be dealt with efficiently by the Krylov subspace methods and hence there have been many studies on the Krylov-based shifted linear equation solvers. Among them, the COCG (Conjugate Orthogonal Conjugate Gradient) [2] and COCR (Conjugate Orthogonal Conjugate Residual)

[3] methods are widely used when the underlying matrix  $A$  is real symmetric, because then the shifted coefficient matrix becomes complex symmetric.

On the other hand, it is also possible to construct a solver for the shifted linear system (2) based on the MINRES (minimum residual) method [4, 5]. This approach has several advantages compared with the COCG and COCR methods; it is breakdown-free, the residual norm decreases monotonically, and the matrix-vector products appearing in the algorithm involve only real arithmetic when both  $A$  and the initial residual is real. In addition, it can be applied also to the case of Hermitian  $A$ , where the COCG and COCR methods are not applicable. However, as far as we know, this approach has attracted less attention of the practitioners.

In this paper, we apply the shifted MINRES method to the linear system (2) with coefficient matrices from a real application [6, 7] and discuss its advantages and disadvantages, comparing it with the COCG method. The shifted MINRES method was proposed in [4], but there, only an outline of its derivation is given. We therefore give a detailed derivation of the algorithm, starting from the Lanczos process. Application of the shifted MINRES method to a set of shifted linear systems with multiple shifts will also be discussed.

This paper is structured as follows. Section 2 explains the derivation of the shifted MINRES method and compares it with the COCG method. In Section 3, numerical results on shifted linear systems from a real application ELES (Extra Large Scale Electronic Structure) [6, 7] will be provided. Throughout the paper, the superscript  $H$  denotes the Hermitian conjugate and  $\langle \mathbf{x}, \mathbf{y} \rangle \equiv \mathbf{y}^H \mathbf{x}$  for complex vectors  $\mathbf{x}$  and  $\mathbf{y}$ .  $\|\cdot\|$  denotes the Euclidian norm of a vector.

**Algorithmus 1** Lanczos process for generating an orthonormal basis of  $K_k(A; \mathbf{r}_0)$

---

```

1:  $\mathbf{q}_1 = \mathbf{r}_0 / \|\mathbf{r}_0\|$ 
2:  $\beta_0 = 0$ 
3: for  $j = 1, 2, \dots$ , do
4:    $\mathbf{q}_{j+1}'' = A\mathbf{q}_j - \beta_{j-1}\mathbf{q}_{j-1}$ 
5:    $\alpha_j = \langle \mathbf{q}_{j+1}'', \mathbf{q}_j \rangle$ 
6:    $\mathbf{q}_{j+1}' = \mathbf{q}_{j+1}'' - \alpha_j\mathbf{q}_j$ 
7:    $\beta_j = \|\mathbf{q}_{j+1}'\|$ 
8:    $\mathbf{q}_{j+1} = \mathbf{q}_{j+1}' / \beta_j$ 
9: end for

```

---

## 2. The shifted MINRES algorithm

### 2.1 Derivation

The shifted MINRES method is based on the Lanczos process for Hermitian (or real symmetric) matrices and the shift-invariant property of the Krylov subspace [4]. Let  $\mathbf{x}_0$  be the initial guess of the solution,  $\hat{A} \equiv A + i\sigma I$  and  $\mathbf{r}_0 \equiv \mathbf{b} - \hat{A}\mathbf{x}_0$ . We run the Lanczos process to compute the orthonormal basis  $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k$  of the Krylov subspace  $K_k(A; \mathbf{r}_0) \equiv \text{span}(\mathbf{r}_0, A\mathbf{r}_0, \dots, A^{k-1}\mathbf{r}_0)$ , which is shown in Algorithm 1.

Assume that the Lanczos process has reached the end of the  $k$ th iteration ( $j = k$ ). Then, by letting

$$Q_k = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k] \in \mathbb{C}^{n \times k}, \quad (3)$$

$$T_k = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{k-1} \\ & & \beta_{k-1} & \alpha_k \end{bmatrix} \in \mathbb{R}^{k \times k}, \quad (4)$$

$$T_k^* = \begin{bmatrix} T_k \\ \beta_k \mathbf{e}_k^\top \end{bmatrix} \in \mathbb{R}^{(k+1) \times k}, \quad (5)$$

where  $\mathbf{e}_k$  is the  $k$ th column of  $I_k$ , we have

$$AQ_k = Q_{k+1}T_k^*. \quad (6)$$

This is the matrix representation of the Lanczos process and is known as the *Lanczos decomposition* of order  $k$  of  $A$  [8]. From (6), one can readily obtain the Lanczos decomposition of  $\hat{A}$  as

$$\hat{A}Q_k = Q_{k+1}\hat{T}_k^*, \quad \hat{T}_k^* = T_k^* + i\sigma \begin{bmatrix} I_k \\ \mathbf{0}_k^\top \end{bmatrix}. \quad (7)$$

Once the Lanczos decomposition has been computed, the minimal residual solution in the affine space  $\mathbf{x}_0 + K_k(\hat{A}; \mathbf{b})$  can be obtained as in the GMRES method. Let  $\mathbf{x} = \mathbf{x}_0 + Q_k \mathbf{y}_k$ , where  $\mathbf{y}_k \in \mathbb{C}^k$ , be a vector in this space. Then, its residual norm can be computed as

$$\begin{aligned} \|\mathbf{b} - \hat{A}\mathbf{x}\| &= \|\mathbf{b} - \hat{A}\mathbf{x}_0 - \hat{A}Q_k \mathbf{y}_k\| \\ &= \|\|\mathbf{r}_0\| \mathbf{q}_1 - \hat{A}Q_k \mathbf{y}_k\| \\ &= \|\|\mathbf{r}_0\| Q_{k+1} \mathbf{e}_1 - Q_{k+1} \hat{T}_k^* \mathbf{y}_k\| \\ &= \|Q_{k+1} (\|\mathbf{r}_0\| \mathbf{e}_1 - \hat{T}_k^* \mathbf{y}_k)\| \\ &= \|\|\mathbf{r}_0\| \mathbf{e}_1 - \hat{T}_k^* \mathbf{y}_k\|, \end{aligned} \quad (8)$$

where  $\mathbf{e}_1$  is the 1st column vector of  $I_{k+1}$ . Hence, the

problem of finding  $\mathbf{y}_k$  that minimizes the residual is the least squares problem. Using the QR decomposition  $\hat{T}_k^* = U_k R_k$ , where  $U_k \in \mathbb{C}^{(k+1) \times k}$  and  $R_k \in \mathbb{C}^{k \times k}$ , its solution can be written as  $\mathbf{y}_k = \|\mathbf{r}_0\| R_k^{-1} U_k^\top \mathbf{e}_1$ . Thus, we obtain the minimal residual solution at step  $k$  as

$$\mathbf{x}_k = \mathbf{x}_0 + \|\mathbf{r}_0\| Q_k R_k^{-1} U_k^\top \mathbf{e}_1. \quad (9)$$

However, using (9) to compute  $\mathbf{x}_k$  requires all of  $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k$  to be stored in the memory. To derive a practical algorithm, we need to transform it to a recursive formula, which requires only the three newest basis vectors  $\mathbf{q}_{k-2}, \mathbf{q}_{k-1}, \mathbf{q}_k$  to be stored (along with some auxiliary vectors). The details of this transformation, which is omitted in [4], will be given below.

First, we express  $Q_k R_k^{-1}$  in terms of  $Q_{k-1} R_{k-1}^{-1}$ . Letting  $\hat{T}_k = T_k + i\sigma I_k$ , we have

$$\hat{T}_k^* = \left[ \begin{array}{c|c} \hat{T}_{k-1} & \beta_{k-1} \\ \hline & \beta_{k-1} \\ \hline \mathbf{0}_{k-1} & \beta_k \end{array} \right]. \quad (10)$$

In the QR decomposition of  $\hat{T}_k^*$ , we apply the Givens rotations

$$G_j = \begin{bmatrix} c_j & s_j \\ -\bar{s}_j & c_j \end{bmatrix} = \begin{bmatrix} \cos \theta_j & \sin \theta_j e^{i\phi_j} \\ -\sin \theta_j e^{-i\phi_j} & \cos \theta_j \end{bmatrix} \quad (j = 1, 2, \dots, k) \quad (11)$$

to annihilate the  $(2, 1), (3, 2), \dots, (k+1, k)$  elements in this order. As a result,  $\hat{T}_k^*$  is transformed into

$$R_k = \left[ \begin{array}{c|c} R_{k-1} & r_{k-2,k} \\ \hline & r_{k-1,k} \\ \hline \mathbf{0}_{k-1} & r_{kk} \\ \hline \mathbf{0}_{k-1} & 0 \end{array} \right]. \quad (12)$$

Assuming that  $G_1, G_2, \dots, G_{k-1}$  and  $R_{k-1}$  have been computed by the end of the last step, we only need to compute  $G_k$  and the last column of  $R_k$ .  $r_{k-2,k}$  and  $r_{k-1,k}$  are determined from the following updating formulas:

$$r_{k-2,k} = 0, \quad r_{k-1,k} = \beta_{k-1}, \quad r_{kk} = \alpha_k + i\sigma, \quad (13)$$

$$[r_{k-2,k}, r_{k-1,k}]^\top := G_{k-2} [r_{k-2,k}, r_{k-1,k}]^\top, \quad (14)$$

$$[r_{k-1,k}, r_{kk}]^\top := G_{k-1} [r_{k-1,k}, r_{kk}]^\top. \quad (15)$$

Now, for two scalars  $a, b$ , let  $(r, G) = \text{Givens}(a, b)$  be a function that returns a pair of a scalar  $r$  and a Givens rotation  $G$  such that  $G[a, b]^\top = [r, 0]^\top$ . Then,  $r_{kk}$  and  $G_k$  are computed from  $r_{kk}$  given by (15) and  $\beta_k$  as

$$(r_{kk}, G_k) := \text{Givens}(r_{kk}, \beta_k). \quad (16)$$

Now, let  $Q_k R_k^{-1} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_k]$ . Then, from the structure of  $R_k$ , it is easy to show that

$$Q_k R_k^{-1} = [Q_{k-1} R_{k-1}^{-1}, \mathbf{p}_k], \quad (17)$$

$$\mathbf{p}_k = (\mathbf{q}_k - r_{k-2,k} \mathbf{p}_{k-2} - r_{k-1,k} \mathbf{p}_{k-1}) / r_{kk},$$

where we define  $r_{-1,1} = r_{0,1} = r_{0,2} = 0$ .

Next, we consider computing  $U_k^\top \mathbf{e}_1 \in \mathbb{C}^k$  by recursion.

**Algorithmus 2** Shifted MINRES algorithm

---

```

1:  $\mathbf{r}_0 = \mathbf{b} - \hat{A}\mathbf{x}_0$ ,  $\mathbf{q}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|$ ,  $\mathbf{p}_{-1} = \mathbf{p}_0 = \mathbf{0}$ 
2:  $\beta_0 = 0$ ,  $f_1^{(0)} = 1$ 
3: for  $j = 1, 2, \dots$ , do
4:    $\mathbf{q}_{j+1}'' = A\mathbf{q}_j - \beta_{j-1}\mathbf{q}_{j-1}$ 
5:    $\alpha_j = \langle \mathbf{q}_{j+1}'', \mathbf{q}_j \rangle$ 
6:    $\mathbf{q}_{j+1}'' = \mathbf{q}_{j+1}'' - \alpha_j\mathbf{q}_j$ 
7:    $\beta_j = \|\mathbf{q}_{j+1}'\|$ 
8:    $r_{j-2,j} = 0$ ,  $r_{j-1,j} = \beta_{j-1}$ ,  $r_{jj} = \alpha_k + i\sigma$ 
9:   if  $j \geq 3$  then update  $r_{j-2,j}$  and  $r_{j-1,j}$  by (14)
10:  if  $j \geq 2$  then update  $r_{j-1,j}$  and  $r_{jj}$  by (15)
11:  compute  $G_j$  and update  $r_{jj}$  by (16)
12:   $\mathbf{p}_j = (\mathbf{q}_j - r_{j-2,j}\mathbf{p}_{j-2} - r_{j-1,j}\mathbf{p}_{j-1})/r_{jj}$ 
13:   $\mathbf{x}_j = \mathbf{x}_{j-1} + \|\mathbf{r}_0\|c_j f_j^{(j-1)}\mathbf{p}_j$ 
14:   $f_{j+1}^{(j)} = -\bar{s}_j f_j^{(j-1)}$ 
15:   $\mathbf{q}_{j+1} = \mathbf{q}_{j+1}''/\beta_j$ 
16: end for

```

---

Let the  $k+1$  dimensional vector obtained by applying the Givens rotation  $G_j$  to the  $j$ th and the  $(j+1)$ st elements of  $\mathbf{e}_1 \in \mathbb{C}^{k+1}$  for  $j = 1, 2, \dots, k$  in this order be  $\mathbf{f}^{(k)} = [f_1^{(k)}, f_2^{(k)}, \dots, f_{k+1}^{(k)}]^\top$ . Then,  $U_k^\mathsf{H}\mathbf{e}_1$  is obtained by deleting the last element of  $\mathbf{f}^{(k)}$ . Noting that

$$\mathbf{f}^{(k)} = \left[ \begin{array}{c|c} I_{k-1} & \\ \hline & G_k \end{array} \right] \left[ \begin{array}{c} \mathbf{f}^{(k-1)} \\ 0 \end{array} \right] = \left[ \begin{array}{c} f_1^{(k-1)} \\ \vdots \\ f_{k-1}^{(k-1)} \\ c_k f_k^{(k-1)} \\ -\bar{s}_k f_k^{(k-1)} \end{array} \right],$$

we can derive the recursion for  $U_k^\mathsf{H}\mathbf{e}_1$  and  $f_{k+1}^{(k)}$  as

$$U_k^\mathsf{H}\mathbf{e}_1 = \left[ \begin{array}{c} U_{k-1}^\mathsf{H}\mathbf{e}_1' \\ c_k f_k^{(k-1)} \end{array} \right], \quad f_{k+1}^{(k)} = -\bar{s}_k f_k^{(k-1)}, \quad (18)$$

where  $\mathbf{e}_1'$  is the 1st column of  $I_k$ .

Plugging (17) and the first equation of (18) into (9) yields the recursion formula for  $\mathbf{x}_k$ :

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}_0 + \|\mathbf{r}_0\|Q_{k-1}R_{k-1}^{-1}U_{k-1}^\mathsf{H}\mathbf{e}_1' + \|\mathbf{r}_0\|c_k f_k^{(k-1)}\mathbf{p}_k \\ &= \mathbf{x}_{k-1} + \|\mathbf{r}_0\|c_k f_k^{(k-1)}\mathbf{p}_k. \end{aligned} \quad (19)$$

In summary, the algorithm of the shifted MINRES method [4, Section 3] can be stated as in Algorithm 2.

## 2.2 Comparison with the COCG method

Now we discuss the advantages and disadvantages of the shifted MINRES method. First, it is breakdown-free in exact arithmetic for nonsingular  $\hat{A}$ . In fact, breakdown can occur only if  $\beta_k = 0$  or  $r_{kk} = 0$ . But  $\beta_k = 0$  means that  $K_{k+1}(A; \mathbf{r}_0) = K_k(A; \mathbf{r}_0)$ , or equivalently,  $K_{k+1}(\hat{A}; \mathbf{r}_0) = K_k(\hat{A}; \mathbf{r}_0)$ . In that case, it is easy to show that  $\hat{A}^{-1}\mathbf{r}_0 \in K_k(\hat{A}; \mathbf{r}_0)$ . So  $\mathbf{x} = A^{-1}\mathbf{b} \in \mathbf{x}_0 + K_k(\hat{A}; \mathbf{r}_0)$  and the solution has already been reached. Also,  $r_{kk} = 0$  cannot occur because it means that the rank of  $\hat{T}_k^*$  is smaller than  $k$  and this contradicts the fact that the left-hand side of the first equation of (7) is of rank  $k$ . Second, from the construction of the algorithm, the residual norm decreases monotonically. Third, if  $A$  and the ini-

tial residual  $\mathbf{r}_0$  are real (as is the case where  $\mathbf{b}$  is real and we choose  $\mathbf{x}_0 = \mathbf{0}$ ), all the matrix-vector products can be done in real arithmetic, which greatly reduces the computational work. Fourth, it can be used not only for real symmetric  $A$  but also for complex Hermitian  $A$ .

None of these properties are shared by the COCG and COCR methods. Theoretically, they have a possibility of breakdown because they are based on the complex symmetric Lanczos process, in which quantities like  $\mathbf{r}_k^\top \mathbf{r}_k$  appear in the denominators. They do not have minimization property of the residual norm or other quantities. Also, even when  $A$  and  $\mathbf{r}_0$  are real, the matrix-vector products in the COCG and COCR methods involve complex multiplicand vectors. Finally, they cannot be used when  $A$  is complex Hermitian, because the shifted matrix  $\hat{A}$  is not complex symmetric. In that case, some practitioners use the (shifted) Bi-CG method, which does not exploit the symmetry of the coefficient matrix [9].

One of the disadvantages of the shifted MINRES method is that it requires 5 AXPYs (operations of the form  $\mathbf{z} := \alpha\mathbf{x} + \mathbf{y}$ ) per iteration, while COCG and COCR methods require 3 and 4, respectively [3]. The effect of this is mitigated when  $A$  has many nonzero elements per row and the total computational time is dominated by the time of matrix-vector products. The number of inner products or norm computations per iteration is two for all the three algorithms. Another shortcoming is that the shifted MINRES method cannot be combined with efficient incomplete factorization type preconditioners, for the preconditioning destroys the shift structure. In contrast, the COCG and COCR methods can be used with this type of preconditioners, if the preconditioned matrix is constructed to be complex symmetric.

## 2.3 Shifted linear systems with multiple shifts

In some applications like the computation of Green's function for many-particle Hamiltonian, a set of shifted linear systems with multiple shifts need to be solved [10]. The shifted MINRES method can be easily adapted to this case as well. Since the orthonormal basis vectors  $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k$  are common to all shifts, one only needs to add equations for  $T_k^*$ , its QR decomposition and the vectors  $\mathbf{p}_k$  and  $\mathbf{x}_k$  for each value of the shift. In Algorithm 2, this amounts to making a copy of lines 8–14 for each shift. The shift switching technique, devised for the shifted COCG and COCR methods, is not needed.

## 3 Numerical results

In this section, we compare the convergence properties of the shifted MINRES method with that of the COCG method. As test matrices, we use matrices generated by an electronic structure calculation code ELSES. The matrices used are PPE3594 and PPE7194, which are sparse real symmetric matrices coming from the simulation of organic polymers named poly-(phenylene ethynylene) [11]. PPE3594 is of order  $n = 3,594$  and has 41,781 nonzero elements, while PPE7194 is of order  $n = 7,194$  and has 84,026 nonzero elements. These matrices are shifted by  $i\sigma$ , where  $\sigma = 0, 0.01, 0.1$  and  $1$ , and the shifted matrices  $\hat{A} = A + i\sigma I$  are generated. Note that as  $\sigma$  increases,  $\hat{A}$  approaches a multiple of the identity

matrix and therefore its conditioning gets improved.

We compare three types of solvers, namely, the shifted MINRES method without a preconditioner, the COCG method without a preconditioner and the COCG method with a complex symmetric incomplete Cholesky (CSIC(0)) decomposition preconditioner. In the last approach, we first compute the complex symmetric incomplete Cholesky decomposition  $\hat{A} \simeq \hat{L}\hat{L}^\top$  without fill-ins and then apply the COCG method to the preconditioned system  $(\hat{L}^{-1}\hat{A}\hat{L}^{-\top})(\hat{L}^\top \mathbf{x}) = \hat{L}^{-1}\mathbf{b}$ . Note that the preconditioned matrix  $\hat{L}^{-1}\hat{A}\hat{L}^{-\top}$  is complex symmetric but no longer has the shift structure (2). Hence, while the COCG method can be used, the shifted MINRES method is not applicable. The solver codes were written in C using double precision arithmetic. All the computations were performed on one node of Fujitsu FX10 installed at the Educational Center on Computational Science and Engineering, Kobe University.

Convergence histories of the three methods for PPE3594 are shown for the cases of  $\sigma = 0.01$  and  $\sigma = 0.1$  in Figs. 1 and 2, respectively. Here, the horizontal axis is the number of iterations and the vertical axis is  $\log_{10}$  of the relative residual norm,  $\log_{10}(\|\mathbf{b} - \hat{A}\mathbf{x}_k\|/\|\mathbf{b}\|)$ . Note that we used the true residual instead of the vector  $\mathbf{r}_k$  appearing in the algorithms. The graphs clearly show the monotonic decrease of the residual norm of the shifted MINRES method. The convergence history of the COCG method without preconditioning is close to that of the shifted MINRES method, but oscillation of the residual can be seen due to the lack of the minimal residual property. We observed that this oscillation becomes severer as  $\sigma$  decreases, that is, as  $\hat{A}$  deviates from the identity matrix. The preconditioning was effective in the case of  $\sigma = 0.1$  and succeeded in reducing the number of iterations of the COCG method by about 2.5 times. However, in the  $\sigma = 0.01$  case, the preconditioning was harmful and led to non-convergence of the COCG method. In this case, it was found that a very small pivot was generated during the CSIC(0) decomposition and it resulted in huge norm of the preconditioned matrix. Thus, the shifted MINRES method was the best in terms of the number of iterations in this case. The results for the PPE7194 matrix were similar to those for PPE3594.

Finally, we generated a random real skew-symmetric matrix  $B$  and applied the shifted MINRES method to  $(A + iB + i\sigma I)\mathbf{x} = \mathbf{b}$ , where  $A$  is the PPE3594 matrix. This matrix has an Hermite-plus-shift structure, which cannot be dealt with by the COCG method. The shifted MINRES method was able to solve the equation without problem and exhibited smooth and monotonic decrease of the residual norm.

These results show that the shifted MINRES method is another reliable and smoothly converging solver for the shifted system (2) and can be a method of choice in some applications.

## Acknowledgment

We thank the anonymous referee and the editor for helpful comments and Prof. Youhei Yamaji for valuable discussion. This study is supported by JSPS KAKENHI

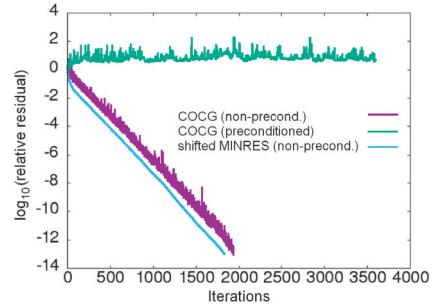


Fig. 1. Convergence history for the PPE3594 matrix ( $\sigma = 0.01$ ).

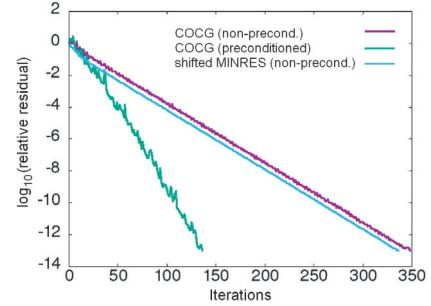


Fig. 2. Convergence history for the PPE3594 matrix ( $\sigma = 0.1$ ).

Grant Nos. 16KT0016, 17H02828 and 17K19966. The computational experiments in this paper were performed using FX10 at the Education Center on Computational Science and Engineering of Kobe University.

## References

- [1] T. Tada, T. Yamamoto and S. Watanabe, Molecular orbital concept on spin-flip transport in molecular junctions, *Theor. Chem. Acc.*, **130** (2011), 775–788.
- [2] H. A. Van der Vorst and J. B. M. Melissen, A Petrov-Galerkin type method for solving  $A\mathbf{x} = \mathbf{b}$ , where  $A$  is symmetric complex, *IEEE Trans. Magn.*, **26** (1990), 706–708.
- [3] T. Sogabe and S. -L. Zhang, A COCR method for solving complex symmetric linear systems, *J. Comput. Appl. Math.*, **199** (2007), 297–303.
- [4] R. W. Freund, On conjugate gradient type methods and polynomial preconditioners for a class of complex non-hermitian matrices, *Numer. Math.*, **57** (1990), 285–312.
- [5] G. Gu and G. Liu, On convergence property of MINRES method for solving a complex shifted Hermitian linear system, *Int. J. Math. Comput. Sci.*, **7** (2013), 290–294. <https://waset.org/Publication/16770>
- [6] ELSES (Extra Large Scale Electronic Structure). <http://www.elses.jp>
- [7] T. Hoshi, S. Yamamoto, T. Fujiwara, T. Sogabe and S.-L. Zhang, An order- $N$  electronic structure theory with generalized eigenvalue equations and its application to a ten-million-atom system, *J. Phys.: Condens. Matter*, **24** (2012), 165502, 5pp.
- [8] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.
- [9] Youhei Yamaji, private communication.
- [10] S. Yamamoto, T. Sogabe, T. Hoshi, S. -L. Zhang and T. Fujiwara, Shifted conjugate-orthogonal-conjugate-gradient method and its application to double orbital extended Hubbard model, *J. Phys. Soc. Jpn.*, **77** (2008), 114713.
- [11] H. Imachi, S. Yokoyama, T. Kaji, Y. Abe, T. Tada and T. Hoshi, One-hundred-nm-scale electronic structure and transport calculations of organic polymers on the K computer, *AIP Conference Proceedings*, **1790** (2016), 020010.