A LOOK-BACK-TYPE RESTART FOR THE RESTARTED KRYLOV SUBSPACE METHODS TO SOLVE NON-HERMITIAN LINEAR SYSTEMS

AKIRA IMAKURA *, TOMOHIRO SOGABE †, AND SHAO-LIANG ZHANG ‡

Abstract. In this paper, we investigate the restarted Krylov subspace methods, as typified by the GMRES(m) method and the FOM(m) method, for solving non-Hermitian linear systems. We have recently focused on the restart of the GMRES(m) method and proposed the extension of the GMRES(m) method based on the error equations. The main purpose of this paper is to apply the extension to other restarted Krylov subspace methods, and propose a specific restart technique for the restarted Krylov subspace method. The specific restart technique is named as *a Look-Back-type restart*, and is based on an implicit residual polynomial reconstruction via the initial guess. The comparison analysis based on the residual polynomials and some numerical experiments indicate that the Look-Back-type restart achieves more efficient convergence results than the traditional restarted Krylov subspace methods.

Key words. non-Hermitian linear systems, restarted Krylov subspace methods, residual polynomials, a Look-Back-type restart

AMS subject classifications. 65F10

1. Introduction.

In this paper, we consider solving large and sparse linear systems of the form:

$$A\boldsymbol{x} = \boldsymbol{b}, \quad A \in \mathbb{C}^{n\times n}, \quad \boldsymbol{x}, \boldsymbol{b} \in \mathbb{C}^n, \tag{1.1}$$

where the coefficient matrix A is assumed to be non-Hermitian and nonsingular. These linear systems often arise from the discretization of partial differential equations in the fields of the computational science and engineering.

Direct methods such as the LU decomposition are generally selected in terms of accuracy and stability for the small and dense coefficient matrix A. On the other hand, for the case where the coefficient matrix A is large and sparse, iterative methods are widely used because of their computational costs and storage requirements. Recently, the Krylov subspace methods are recognized as standard algorithms for large, sparse and non-Hermitian linear systems (1.1); for details see [16, 21].

In this paper, we investigate the restarted Krylov subspace methods, as typified by the GMRES(m) method [17] and the FOM(m) method [15], for solving non-Hermitian linear systems (1.1). The Arnoldi-based Krylov subspace methods have some difficulties in terms of the computational costs and storage requirements due to the long-term recurrence. The (two-sided) Lanczos-based Krylov subspace methods also have some difficulties in terms of the pseudo convergence due to accumulations of the round-off errors. In order to remedy these difficulties, the restart is often applied to the Krylov subspace methods.

Let m be the restart frequency, and $\boldsymbol{x}_0^{(1)}$ be the initial guess of the 1st restart cycle. Then the ℓth restart cycle of the restarted Krylov subspace methods are operated as

*Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1, Tennodai Tsukuba-city, Ibaraki 305-8573, Japan (`imakura@cs.tsukuba.ac.jp`).

†Graduate School of Information Science and Technology, Aichi Prefectural University, 1522-3 Ibaragabasama, Kumabari, Nagakute-cho, Aichi-gun, Aichi, 480-1198, Japan (`sogabe@ist.aichi-pu.ac.jp`).

‡Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan (`zhang@na.cse.nagoya-u.ac.jp`).

follows:

1. Solve (approximately) $A\boldsymbol{x} = \boldsymbol{b}$ by m iterations of some Krylov subspace method with the initial guess $\boldsymbol{x}_0^{(\ell)}$, and get the approximate solution $\boldsymbol{x}^{(\ell)}$.
2. Update the initial guess of the $(\ell+1)$th restart cycle: $\boldsymbol{x}_0^{(\ell+1)} := \boldsymbol{x}^{(\ell)}$, and go to the $(\ell+1)$th restart cycle.

The restart remedies the difficulties of the Krylov subspace methods due to the long-term recurrence and/or accumulations of the round-off errors; however, the restart generally slows their convergence. Therefore, in order to improve Step 1, several improvement techniques have been proposed, such as the adaptive preconditioning techniques based on the deflation [1,6,8], the techniques based on the augmented Krylov subspace [2,4,11–13] and the techniques based on adaptively determining the restart frequency m [3,14,19,22].

On the other hand, we focused on the update of the initial guess in each restart cycle (Step 2), then we recently proposed the extension of the GMRES(m) method [9]. It is also shown that this is a natural extension in terms of the error equations and the iterative refinement scheme, and it has high potential for more efficient convergence than the traditional GMRES(m) method [9]. In this regard, however, the convergence of the extension of the GMRES(m) method did not analyzed enough, and specific algorithms for efficient convergence still have not proposed.

The main purpose of this paper is to apply the extension to other restarted Krylov subspace methods, and propose a specific restart technique for the restarted Krylov subspace method. The specific restart technique is named as *a Look-Back-type restart*, and is based on an implicit residual polynomial reconstruction via the initial guess. The performance of the Look-Back-type restart is evaluated by a comparison analysis based on the residual polynomials and some numerical experiments.

This paper is organized as follows. In Section 2, we briefly describe a general form of the restarted Krylov subspace methods. In Section 3, from the analysis based on the residual polynomials, we propose a Look-Back-type restart based on an implicit residual polynomial reconstruction. The reconstructed residual polynomials are compared with the traditional restarted Krylov subspace methods in Section 4. In Section 5, the performance of the Look-Back-type restart is evaluated by some numerical experiments. Our conclusions are summarized in Section 6.

Throughout this paper, let $\mathscr{V}, \mathscr{W}$ be the subspaces, then $\mathscr{V} + \mathscr{W}$ denotes the sum of the subspaces of $\mathscr{V}, \mathscr{W}$, i.e., $\mathscr{V} + \mathscr{W} := \{\boldsymbol{v} + \boldsymbol{w} | \boldsymbol{v} \in \mathscr{V}, \boldsymbol{w} \in \mathscr{W}\}$.

2. Restarted Krylov subspace methods. We have the following theorem for the exact solution of the linear systems (1.1); see e.g., [16].

THEOREM 2.1. *Let $\boldsymbol{x}_0$ be an initial guess and $\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}_0$ be the corresponding initial residual, respectively. We also let d be the grade of $\boldsymbol{r}_0$ with respect to A, i.e.,*

$$d := d(A, \boldsymbol{r}_0) := \min\{k | P_k(A)\boldsymbol{r}_0 = \boldsymbol{0}, P_k(\lambda) \in \mathbb{P}_k, P_k(0) = 1\},$$

where $\mathbb{P}_k$ is the set of k-degree polynomials. Then the exact solution $\boldsymbol{x}^* := A^{-1}\boldsymbol{b}$ satisfies the following relation:

$$\boldsymbol{x}^* \in \boldsymbol{x}_0 + \mathscr{K}_d(A, \boldsymbol{r}_0), \quad \mathscr{K}_d(A, \boldsymbol{r}_0) := \operatorname{span}\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{d-1}\boldsymbol{r}_0\}.$$

For a general property of the Krylov subspaces, we can also derive the following lemma.

LEMMA 2.2. *Let $A \in \mathbb{C}^{n\times n}, \boldsymbol{u}, \boldsymbol{w} \in \mathbb{C}^n$. Then, for any $\boldsymbol{v} \in \mathscr{K}_{s+1}(A, \boldsymbol{u})$,*

$$\boldsymbol{w} + \mathscr{K}_t(A, \boldsymbol{v} - A\boldsymbol{w}) \subset \mathscr{K}_{s+t}(A, \boldsymbol{u}) \tag{2.1}$$

is satisfied if and only if

$$\boldsymbol{w} \in \mathscr{K}_s(A, \boldsymbol{u}). \tag{2.2}$$

We assume, however, that $s+t < d$, where $d := d(A, \boldsymbol{u})$ is the grade of $\boldsymbol{u}$ with respect to A.

Proof. First, we prove that (2.1) is a necessary condition for (2.2). If $\boldsymbol{w} \in \mathscr{K}_s(A, \boldsymbol{u})$, then $\boldsymbol{v} - A\boldsymbol{w} \in \mathscr{K}_{s+1}(A, \boldsymbol{u})$ for any $\boldsymbol{v} \in \mathscr{K}_{s+1}(A, \boldsymbol{u})$. This leads to

$$\mathscr{K}_t(A, \boldsymbol{v} - A\boldsymbol{w}) \subset \mathscr{K}_{s+t}(A, \boldsymbol{u}).$$

Next, we prove that (2.1) is a sufficient condition for (2.2). From (2.1), we have

$$\boldsymbol{w} \in \mathscr{K}_{s+t}(A, \boldsymbol{u}), \tag{2.3}$$

$$\mathscr{K}_t(A, \boldsymbol{v} - A\boldsymbol{w}) \subset \mathscr{K}_{s+t}(A, \boldsymbol{u}). \tag{2.4}$$

Then, we introduce a condition for the vector $\boldsymbol{w}$ from (2.4), then consider the intersection with (2.3).

Since a general property of the Krylov subspaces $\mathscr{K}_k(A, \boldsymbol{x})$, i.e.,

$$\mathscr{K}_k(A, \boldsymbol{x} + \boldsymbol{y}) \subset \mathscr{K}_k(A, \boldsymbol{x}) + \mathscr{K}_k(A, \boldsymbol{y}),$$

we have

$$\mathscr{K}_t(A, A\boldsymbol{w}) \subset \mathscr{K}_t(A, \boldsymbol{v}) + \mathscr{K}_t(A, \boldsymbol{v} - A\boldsymbol{w}). \tag{2.5}$$

On the other hand, for any $\boldsymbol{v} \in \mathscr{K}_{s+1}(A, \boldsymbol{u})$, we also have

$$\mathscr{K}_t(A, \boldsymbol{v}) \subset \mathscr{K}_{s+t}(A, \boldsymbol{u}). \tag{2.6}$$

Then, substituting (2.4) and (2.6) in (2.5), the relation

$$\mathscr{K}_t(A, A\boldsymbol{w}) \subset \mathscr{K}_{s+t}(A, \boldsymbol{u})$$

is derived. This leads to the following condition for the vector $\boldsymbol{w}$:

$$A^t\boldsymbol{w} \in \mathscr{K}_{s+t}(A, \boldsymbol{u}). \tag{2.7}$$

We then consider the intersection of (2.3) and (2.7). From (2.3), the vector $\boldsymbol{w}$ can be written by the following polynomial form:

$$\boldsymbol{w} = \sum_{i=0}^{s+t-1} \rho_i A^i \boldsymbol{u} \left(= \sum_{i=0}^{s-1} \rho_i A^i \boldsymbol{u} + \sum_{i=s}^{s+t-1} \rho_i A^i \boldsymbol{u} \right).$$

Substituting the polynomial form in (2.7), we get

$$\rho_s = \rho_{s+1} = \cdots = \rho_{s+t-1} = 0.$$

This means the vector $\boldsymbol{w}$ satisfies

$$\boldsymbol{w} = \sum_{i=0}^{s-1} \rho_i A^i \boldsymbol{u} \in \mathscr{K}_s(A, \boldsymbol{u}).$$

Therefore, it is true that (2.1) is a sufficient condition for (2.2). Thus Lemma 2.2 is proved. □

In what follows, based on Theorem 2.1 and Lemma 2.2, we consider a relationship among the sequence of the iterative solutions constructed from the Krylov subspaces and describe a general form of the restarted Krylov subspace methods.

2.1. A relationship among the sequence of the iterative solutions constructed from the Krylov subspaces. For $\ell = 1, 2, \ldots$, let $\boldsymbol{x}_0^{(\ell)}$ be a sequence of initial guesses, and $\boldsymbol{r}_0^{(\ell)} := \boldsymbol{b} - A\boldsymbol{x}_0^{(\ell)}$ be the sequence of the corresponding residual vectors, respectively. Applying the sequence of $m^{(\ell)}$ dimensional Krylov subspaces $\mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)})$ for iterative solutions of the linear systems (1.1), iterative solutions $\boldsymbol{x}^{(\ell)}$ can be expressed as

$$\boldsymbol{x}^{(\ell)} = \boldsymbol{x}_0^{(\ell)} + \boldsymbol{z}^{(\ell)}, \quad \boldsymbol{z}^{(\ell)} \in \mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)}), \quad \ell = 1, 2, \ldots, \tag{2.8}$$

where the vectors $\boldsymbol{z}^{(\ell)}$ are designed by some condition, e.g., the minimal residual condition, the Ritz-Galerkin condition or the Petrov-Galerkin condition.

Note that in the case of

$$\boldsymbol{x}_0^{(\ell+1)} = \boldsymbol{x}^{(\ell)}, \quad \ell = 1, 2, \ldots, \tag{2.9}$$

if one designs the vector $\boldsymbol{z}^{(\ell)}$ using the minimal residual condition and set $m^{(\ell)} = m$, then the sequence of $\boldsymbol{x}^{(\ell)}$ is just the sequence of the approximate solutions of the GMRES(m) method. In this case, the vectors $\boldsymbol{x}^{(\ell)}$ have

$$\begin{aligned} \boldsymbol{x}^{(\ell)} &\in \boldsymbol{x}_0^{(1)} + \mathscr{K}_m(A, \boldsymbol{r}_0^{(1)}) + \mathscr{K}_m(A, \boldsymbol{r}_0^{(2)}) + \cdots + \mathscr{K}_m(A, \boldsymbol{r}_0^{(\ell)}) \\ &= \boldsymbol{x}_0^{(1)} + \mathscr{K}_{m\times\ell}(A, \boldsymbol{r}_0^{(1)}). \end{aligned}$$

This is a fortunate property for solving the linear systems because of Theorem 2.1.

Here, let us consider a more general case of

$$\boldsymbol{x}_0^{(\ell+1)} = \boldsymbol{x}^{(\ell)} + \boldsymbol{y}^{(\ell+1)}, \quad \ell = 1, 2, \ldots. \tag{2.10}$$

Then in this section, we discuss a condition for the vector $\boldsymbol{y}^{(\ell+1)}$ in which all the vectors $\boldsymbol{x}^{(\ell)}$ on the affine space:

$$\boldsymbol{x}^{(\ell)} \in \boldsymbol{x}_0^{(1)} + \mathscr{K}_{M^{(\ell)}}(A, \boldsymbol{r}_0^{(1)}), \quad M^{(\ell)} = \sum_{i=1}^{\ell} m^{(i)}, \quad \ell = 1, 2, \ldots,$$

as well as the case of (2.9).

LEMMA 2.3. *For any $\boldsymbol{z}^{(\ell)} \in \mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)})$ and $\boldsymbol{z}^{(\ell+1)} \in \mathscr{K}_{m^{(\ell+1)}}(A, \boldsymbol{r}_0^{(\ell+1)})$,*

$$\boldsymbol{x}^{(\ell+1)} \in \boldsymbol{x}_0^{(\ell)} + \mathscr{K}_{m^{(\ell)}+m^{(\ell+1)}}(A, \boldsymbol{r}_0^{(\ell)})$$

is satisfied if and only if

$$\boldsymbol{y}^{(\ell+1)} \in \mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)}).$$

We assume, however, that $m^{(\ell)} + m^{(\ell+1)} < d$, where $d := d(A, \boldsymbol{r}_0^{(\ell)})$ is the grade of $\boldsymbol{r}_0^{(\ell)}$ with respect to A.

Proof. From the definition of the vector $\boldsymbol{x}^{(\ell+1)}$ and $\boldsymbol{r}_0^{(\ell+1)} = \boldsymbol{r}^{(\ell)} - A\boldsymbol{y}^{(\ell+1)}$,

$$\boldsymbol{x}^{(\ell+1)} \in \boldsymbol{x}_0^{(\ell)} + \mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)}) + \boldsymbol{y}^{(\ell+1)} + \mathscr{K}_{m^{(\ell+1)}}(A, \boldsymbol{r}^{(\ell)} - A\boldsymbol{y}^{(\ell+1)}),$$

for any $\boldsymbol{z}^{(\ell)} \in \mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)})$. Here, from Lemma 2.2,

$$\boldsymbol{y}^{(\ell+1)} + \mathscr{K}_{m^{(\ell+1)}}(A, \boldsymbol{r}^{(\ell)} - A\boldsymbol{y}^{(\ell+1)}) \subset \mathscr{K}_{m^{(\ell)}+m^{(\ell+1)}}(A, \boldsymbol{r}_0^{(\ell)}) \Leftrightarrow \boldsymbol{y}^{(\ell+1)} \in \mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)}),$$

Algorithm 1 A general form of the restarted Krylov subspace method

1: Choose the initial guess $\boldsymbol{x}_0^{(1)}$
2: **For** $\ell = 1, 2, \ldots,$ until convergence **Do:**
3: Set the restart frequency $m^{(\ell)}$ and the KS$^{(\ell)}$ method
4: Solve (approximately) $A\boldsymbol{x} = \boldsymbol{b}$ by $m^{(\ell)}$ iterations of the KS$^{(\ell)}$ method with the initial guess $\boldsymbol{x}_0^{(\ell)}$, and get the approximate solution $\boldsymbol{x}^{(\ell)}$
5: Set the vector $\boldsymbol{y}^{(\ell+1)} \in \mathscr{K}_{M^{(\ell)}}(A, \boldsymbol{r}_0^{(1)})$, where $M^{(\ell)} = \sum_{i=1}^{\ell} m^{(i)}$
6: Update the initial guess $\boldsymbol{x}_0^{(\ell+1)} := \boldsymbol{x}^{(\ell)} + \boldsymbol{y}^{(\ell+1)}$
7: **End For**

because of $\boldsymbol{r}^{(\ell)} \in \mathscr{K}_{m^{(\ell)}+1}(A, \boldsymbol{r}_0^{(\ell)})$ for any $\boldsymbol{z}^{(\ell)} \in \mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)})$. Therefore, also using the following obvious relation:

$$\mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)}) \subset \mathscr{K}_{m^{(\ell)}+m^{(\ell+1)}}(A, \boldsymbol{r}_0^{(\ell)}),$$

Lemma 2.3 is proved. □

Proposition 2.4. *For any $\boldsymbol{z}^{(\ell)} \in \mathscr{K}_{m^{(\ell)}}(A, \boldsymbol{r}_0^{(\ell)}), \ell = 1, 2, \ldots,$*

$$\boldsymbol{x}^{(\ell)} \in \boldsymbol{x}_0^{(1)} + \mathscr{K}_{M^{(\ell)}}(A, \boldsymbol{r}_0^{(1)}), \quad \ell = 1, 2, \ldots,$$

is satisfied if and only if

$$\boldsymbol{y}^{(\ell+1)} \in \mathscr{K}_{M^{(\ell)}}(A, \boldsymbol{r}_0^{(1)}), \quad \ell = 1, 2, \ldots,$$

where $M^{(\ell)} = \sum_{i=1}^{\ell} m^{(i)}$. We assume, however, that $M^{(\ell)} < d$, where $d := d(A, \boldsymbol{r}_0^{(1)})$ is the grade of $\boldsymbol{r}_0^{(1)}$ with respect to A.

Proof. Proposition 2.4 is proved by using Lemma 2.3 recursively, starting from $\ell = 1$. □

2.2. A general form of the restarted Krylov subspace method. We consider the sequential process which refines the approximate solution of the linear system (1.1) by Eqs. (2.8) and (2.10). This sequential process can be regarded as a general form of the restarted Krylov subspace method; see Algorithm 1, which can dynamically set the restart frequency $m^{(\ell)}$ and the using Krylov subspace method KS$^{(\ell)}$ in each restart cycle, and which also update the initial guess by Eq. (2.10). Note that the algorithms of the extension of the GMRES(m) method proposed in [9] can be introduced from the Algorithm 1 [9, Algorithm 3.2].

In this regard, however it is expected that the approximate solution $\boldsymbol{x}^{(\ell)}$ obtained by Algorithm 1 satisfies

$$\boldsymbol{x}^{(\ell)} \in \boldsymbol{x}_0^{(1)} + \mathscr{K}_{M^{(\ell)}}(A, \boldsymbol{r}_0^{(1)}),$$

because of Theorem 2.1. Therefore, in this paper, we impose

$$\boldsymbol{y}^{(\ell+1)} \in \mathscr{K}_{M^{(\ell)}}(A, \boldsymbol{r}_0^{(1)})$$

on the vector $\boldsymbol{y}^{(\ell+1)}$ in each restart cycle of Algorithm 1; see Proposition 2.4.

3. A Look-Back-type restart. In this section, from the analysis based on the residual polynomials, we propose a Look-Back-type restart based on an implicit residual polynomial reconstruction via the initial guess.

In Section 3.1, we analyze residual polynomials of the general form of the restarted Krylov subspace method (Algorithm 1) and describe an implicit residual polynomial reconstruction via the initial guess. In Section 3.2, we propose a Look-Back-type restart based on an implicit residual polynomial reconstruction, and its implementation details are shown in Section 3.3.

3.1. An implicit residual polynomial reconstruction via the initial guess. Recall that for the general form of the restarted Krylov subspace method (Algorithm 1), the initial guess and the corresponding residual are updated by

$$\boldsymbol{x}_0^{(\ell+1)} = \boldsymbol{x}^{(\ell)} + \boldsymbol{y}^{(\ell+1)}, \quad \boldsymbol{r}_0^{(\ell+1)} = \boldsymbol{r}^{(\ell)} - A\boldsymbol{y}^{(\ell+1)}, \quad \boldsymbol{y}^{(\ell+1)} \in \mathscr{K}_{M^{(\ell)}}(A, \boldsymbol{r}_0^{(1)}), \tag{3.1}$$

in each restart cycle. Here, let us focus on the residual vector $\boldsymbol{r}^{(\ell+1)}$ of the $(\ell+1)$th restart cycle, and consider the corresponding residual polynomial.

Because of Proposition 2.4, we have $\boldsymbol{r}^{(\ell)} \in \mathscr{K}_{M^{(\ell)}+1}(A, \boldsymbol{r}_0^{(1)})$. Thus, there exists a polynomial $Q_{M^{(\ell)}}^{(\ell)}(\lambda) \in \mathbb{P}_{M^{(\ell)}}$ such that

$$\boldsymbol{r}^{(\ell)} = Q_{M^{(\ell)}}^{(\ell)}(A)\boldsymbol{r}_0^{(1)}, \quad Q_{M^{(\ell)}}(0) = 1.$$

From this observation, let $P_{m^{(\ell+1)}}^{(\ell+1)}(\lambda) \in \mathbb{P}_{m^{(\ell+1)}}, P_{m^{(\ell+1)}}^{(\ell+1)}(0) = 1$ be the residual polynomial corresponding to $m^{(\ell+1)}$ iterations of the $\mathrm{KS}^{(\ell+1)}$ method of the $(\ell+1)$th restart cycle, then the residual vector $\boldsymbol{r}^{(\ell+1)}$ can be shown by using $Q_{M^{(\ell)}}^{(\ell)}(\lambda)$ like

$$\begin{aligned}
\boldsymbol{r}^{(\ell+1)} &= P_{m^{(\ell+1)}}^{(\ell+1)}(A)\boldsymbol{r}_0^{(\ell+1)} \\
&= P_{m^{(\ell+1)}}^{(\ell+1)}(A)(\boldsymbol{r}^{(\ell)} - A\boldsymbol{y}^{(\ell+1)}) \\
&= P_{m^{(\ell+1)}}^{(\ell+1)}(A)(Q_{M^{(\ell)}}^{(\ell)}(A)\boldsymbol{r}_0^{(1)} - A\boldsymbol{y}^{(\ell+1)}).
\end{aligned}$$

For any vectors $\boldsymbol{y}^{(\ell+1)} \in \mathscr{K}_{M^{(\ell)}}(A, \boldsymbol{r}_0^{(1)})$, from Proposition 2.4, there is also a polynomial $\widetilde{Q}_{M^{(\ell)}}^{(\ell)}(\lambda) \in \mathbb{P}_{M^{(\ell)}}, \widetilde{Q}_{M^{(\ell)}}^{(\ell)}(0) = 1$ such that

$$\widetilde{Q}_{M^{(\ell)}}^{(\ell)}(A)\boldsymbol{r}_0^{(1)} = Q_{M^{(\ell)}}^{(\ell)}(A)\boldsymbol{r}_0^{(1)} - A\boldsymbol{y}^{(\ell+1)}. \tag{3.2}$$

Therefore, the residual vector $\boldsymbol{r}^{(\ell+1)}$ can be rewritten as follows:

$$\boldsymbol{r}^{(\ell+1)} = Q_{M^{(\ell+1)}}^{(\ell+1)}(A)\boldsymbol{r}_0^{(1)} = P_{m^{(\ell+1)}}^{(\ell+1)}(A)\widetilde{Q}_{M^{(\ell)}}^{(\ell)}(A)\boldsymbol{r}_0^{(1)}.$$

As a consequence, we can observe that the vector $\boldsymbol{y}^{(\ell+1)}$ of the general form of the restarted Krylov subspace method reconstructs the residual polynomial by Eq. (3.2). In this paper we call this an implicit residual polynomial reconstruction via the initial guess.

3.2. Proposal for a Look-Back-type restart. Determining the reconstructed polynomial $\widetilde{Q}_{M^{(\ell)}}^{(\ell)}(\lambda)$ particularly, the corresponding vector $\boldsymbol{y}^{(\ell+1)}$ can be computed uniquely as follows:

$$\boldsymbol{y}^{(\ell+1)} = A^{-1}\left(Q_{M^{(\ell)}}^{(\ell)}(A) - \widetilde{Q}_{M^{(\ell)}}^{(\ell)}(A)\right)\boldsymbol{r}_0^{(1)}, \tag{3.3}$$

from Eq. (3.2).

In terms of computational cost, the polynomial $\widetilde{Q}^{(\ell)}_{M^{(\ell)}}(\lambda)$ should be determined such that Eq. (3.3) is easy to compute. From this observation, in this paper, we express the reconstructed residual polynomial $\widetilde{Q}^{(\ell)}_{M^{(\ell)}}(\lambda)$ as the linear combination of the current residual polynomial $Q^{(\ell)}_{M^{(\ell)}}(\lambda)$ and at most $M^{(\ell)}$-degree polynomial $R^{(\ell)}_{M^{(\ell)}}(\lambda) \in \mathbb{P}_{M^{(\ell)}}, R^{(\ell)}_{M^{(\ell)}}(0) = 1$, i.e.,

$$\widetilde{Q}^{(\ell)}_{M^{(\ell)}}(\lambda) := \tau^{(\ell)} Q^{(\ell)}_{M^{(\ell)}}(\lambda) + (1 - \tau^{(\ell)}) R^{(\ell)}_{M^{(\ell)}}(\lambda),$$

where $\tau^{(\ell)} \in \mathbb{C}$. Then, we define the polynomial $R^{(\ell)}_{M^{(\ell)}}(\lambda)$ by using the dth previous polynomial before $Q^{(\ell)}_{M^{(\ell)}}(\lambda)$ as follows:

$$R^{(\ell)}_{M^{(\ell)}}(\lambda) := \begin{cases} Q^{(\ell_d)}_{M^{(\ell_d)}}(\lambda) & (d : \text{even}) \\ \widetilde{Q}^{(\ell_d)}_{M^{(\ell_d)}}(\lambda) & (d : \text{odd}) \end{cases}, \quad \ell_d := \begin{cases} \ell - \frac{d}{2} & (d : \text{even}) \\ \ell - \frac{d+1}{2} & (d : \text{odd}) \end{cases} \tag{3.4}$$

with the (fixed) positive integer parameter $d \in \mathbb{N}$. Here we note that one can naturally extend it by adaptively determining the parameter d in each restart cycle.

In the case of the definition (3.4), Eq. (3.3) can be rewritten by

$$\boldsymbol{y}^{(\ell+1)} = (1 - \tau^{(\ell)}) A^{-1} \left(Q^{(\ell)}_{M^{(\ell)}}(A) - R^{(\ell)}_{M^{(\ell)}}(A) \right) \boldsymbol{r}_0^{(1)}, \tag{3.5}$$

and we have

$$\left(Q^{(\ell)}_{M^{(\ell)}}(A) - R^{(\ell)}_{M^{(\ell)}}(A) \right) \boldsymbol{r}_0^{(1)} = \begin{cases} \boldsymbol{r}^{(\ell)} - \boldsymbol{r}^{(\ell_d)} & (d : \text{even}) \\ \boldsymbol{r}^{(\ell)} - \boldsymbol{r}_0^{(\ell_d)} & (d : \text{odd}) \end{cases}.$$

Let $\Delta \boldsymbol{x}^{(\ell)}$ and $\Delta \boldsymbol{r}^{(\ell)}$ be

$$\Delta \boldsymbol{x}^{(\ell)} := \begin{cases} \boldsymbol{x}^{(\ell)} - \boldsymbol{x}^{(\ell_d)} & (d : \text{even}) \\ \boldsymbol{x}^{(\ell)} - \boldsymbol{x}_0^{(\ell_d)} & (d : \text{odd}) \end{cases}, \quad \Delta \boldsymbol{r}^{(\ell)} := \begin{cases} \boldsymbol{r}^{(\ell)} - \boldsymbol{r}^{(\ell_d)} & (d : \text{even}) \\ \boldsymbol{r}^{(\ell)} - \boldsymbol{r}_0^{(\ell_d)} & (d : \text{odd}) \end{cases},$$

respectively, then from $\Delta \boldsymbol{r}^{(\ell)} = -A \Delta \boldsymbol{x}^{(\ell)}$, Eq. (3.5) can be computed by

$$\boldsymbol{y}^{(\ell+1)} = \mu^{(\ell)} \Delta \boldsymbol{x}^{(\ell)}, \tag{3.6}$$

where $\mu^{(\ell)} := \tau^{(\ell)} - 1 \in \mathbb{C}$. In this paper, we name the restart technique based on Eqs. (3.1) and (3.6) as a Look-Back-type restart.

Based on the Look-Back-type restart, we can efficiently achieve the implicit residual polynomial reconstruction without explicit computations for the polynomials such as $Q^{(\ell)}_{M^{(\ell)}}(\lambda)$ and $R^{(\ell)}_{M^{(\ell)}}(\lambda)$.

3.3. Implementation details. For any $\mu^{(\ell)} \in \mathbb{C}$ the Look-Back-type restart does not satisfy the monotonic decrease of the residual 2-norm $\|\boldsymbol{r}^{(\ell)}\|_2$ even if it is applied to the GMRES(m) method. We now have the following proposition for the monotonic decrease of the residual 2-norm of the Look-Back GMRES(m) method.

PROPOSITION 3.1. *The Look-Back GMRES(m) method satisfies the monotonic decrease of the residual 2-norm:*

$$\|\boldsymbol{r}^{(\ell+1)}\|_2 \le \|\boldsymbol{r}_0^{(\ell+1)}\|_2 = \|\boldsymbol{r}^{(\ell)} - A\boldsymbol{y}^{(\ell+1)}\|_2 \le \|\boldsymbol{r}^{(\ell)}\|_2, \quad \ell = 1, 2, \ldots, \tag{3.7}$$

Algorithm 2 A restarted Krylov subspace method with a Look-Back-type restart

1: Choose the parameter $d \geq 2$ and the initial guess $\boldsymbol{x}_0^{(1)}$
2: **For** $\ell = 1, 2, \ldots,$ until convergence **Do:**
3: Set the restart frequency $m^{(\ell)}$ and the KS$^{(\ell)}$ method
4: Solve (approximately) $A\boldsymbol{x} = \boldsymbol{b}$ by $m^{(\ell)}$ iterations of the KS$^{(\ell)}$ method with the initial guess $\boldsymbol{x}_0^{(\ell)}$, and get the approximate solution $\boldsymbol{x}^{(\ell)}$
5: Compute the vector $\boldsymbol{y}^{(\ell+1)}$ as follows:
  If $\ell = 1$ **then** $\boldsymbol{y}^{(\ell+1)} = \boldsymbol{0}$
  If $\ell \geq 2$ **then**
    If $(\ell = d = 2)$ **or** $(d : \text{even}, \ell \leq \frac{d}{2})$ **or** $(d : \text{odd}, \ell \leq \frac{d-1}{2})$ **then**
      $\Delta\boldsymbol{x}^{(\ell)} := \boldsymbol{x}^{(\ell)} - \boldsymbol{x}_0^{(1)}$
    Else
      $\Delta\boldsymbol{x}^{(\ell)} := \begin{cases} \boldsymbol{x}^{(\ell)} - \boldsymbol{x}^{(\ell_d)} & (d : \text{even}) \\ \boldsymbol{x}^{(\ell)} - \boldsymbol{x}_0^{(\ell_d)} & (d : \text{odd}) \end{cases}, \quad \ell_d := \begin{cases} \ell - \frac{d}{2} & (d : \text{even}) \\ \ell - \frac{d-1}{2} & (d : \text{odd}) \end{cases}$
    End If
    $\boldsymbol{y}^{(\ell+1)} = \mu^{(\ell)}\Delta\boldsymbol{x}^{(\ell)}, \mu^{(\ell)} = \arg\min_{\mu\in\mathbb{C}} \|\boldsymbol{r}^{(\ell)} - \mu A\Delta\boldsymbol{x}^{(\ell)}\|_2$
  End If
6: Update the initial guess $\boldsymbol{x}_0^{(\ell+1)} := \boldsymbol{x}^{(\ell)} + \boldsymbol{y}^{(\ell+1)}$
7: **End For**

as well as the GMRES(m) method if and only if $\mu^{(\ell)}$ *is set by*

$$|\mu^{(\ell)} - \mu_{\min}^{(\ell)}| \leq |\mu_{\min}^{(\ell)}|, \quad \mu_{\min}^{(\ell)} = \arg\min_{\mu\in\mathbb{C}} \|\boldsymbol{r}^{(\ell)} - \mu A\Delta\boldsymbol{x}^{(\ell)}\|_2, \quad \ell = 1, 2, \ldots.$$

Proof. The first inequality of (3.7) is satisfied for any $\mu^{(\ell)}$ because of the minimal residual condition of the GMRES(m) method. Then we prove the necessary and sufficient condition of the second inequality of (3.7).

Since the definition of $\boldsymbol{y}^{(\ell+1)}$ (3.6), the inequality $\|\boldsymbol{r}^{(\ell)} - A\boldsymbol{y}^{(\ell+1)}\|_2 \leq \|\boldsymbol{r}^{(\ell)}\|_2$ can be rewritten as $\|\boldsymbol{r}^{(\ell)}\|_2 - \|\boldsymbol{r}^{(\ell)} - \mu^{(\ell)}A\Delta\boldsymbol{x}^{(\ell)}\|_2 \geq 0$, and this is equivalence to

$$\mu^{(\ell)}(\boldsymbol{r}^{(\ell)}, A\Delta\boldsymbol{x}^{(\ell)}) + \overline{\mu}^{(\ell)}(A\Delta\boldsymbol{x}^{(\ell)}, \boldsymbol{r}^{(\ell)}) - \mu^{(\ell)}\overline{\mu}^{(\ell)}(A\Delta\boldsymbol{x}^{(\ell)}, A\Delta\boldsymbol{x}^{(\ell)}) \geq 0.$$

Then, using $\mu_{\min}^{(\ell)} = (A\Delta\boldsymbol{x}^{(\ell)}, \boldsymbol{r}^{(\ell)})/(A\Delta\boldsymbol{x}^{(\ell)}, A\Delta\boldsymbol{x}^{(\ell)})$, we have

$$|\mu_{\min}^{(\ell)}| - |\mu^{(\ell)} - \mu_{\min}^{(\ell)}| \geq 0.$$

Therefore Proposition 3.1 is proved. ◻

The algorithm of the restarted Krylov subspace method with the Look-Back-type restart is shown in Algorithm 2, where $\mu^{(\ell)}$ is set at $\mu^{(\ell)} = \mu_{\min}^{(\ell)}$. The incremental operations per restart cycle and storage requirements for the Look-Back-type restart compared with the traditional restart (2.9) are also shown in Table 3.1, where Mat-Vec and AXPY denote the incremental number of matrix-vector multiplications and additions of scaled vectors respectively, and Storage means main part of incremental storage requirements.

We note that these incremental operations and storage requirements correspond to our current implementation. One can decrease the incremental Mat-Vec of the Look-Back-type restart; however, such implementation may increase the effect of round-off error.

Table 3.1: The incremental operations per restart cycle and storage requirements for the Look-Back-type restart compared with the traditional restart (2.9).

Mat-Vec	AXPY	Inner-Product	Storage	
			d:even	d:odd
1	3	2	$\frac{d}{2}n$	$\frac{d+1}{2}n$

The parameter d depends on the degree of the polynomial $R^{(\ell)}_{M^{(\ell)}}(\lambda)$. It is expected that quite large parameter d leads to decrease the effect of the Look-Back-type restart, because the degree of the corresponding polynomial $R^{(\ell)}_{M^{(\ell)}}(\lambda)$ is very small, and also leads to increase the storage requirements; see Table 3.1. Thus the quite large d is not recommended. In this paper, we set $d = 3$.

4. Comparison analysis based on the residual polynomials. Let the coefficient matrix A be diagonalizable: $A = X\Lambda X^{-1}$, where $\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$ is the diagonal matrix of the eigenvalues of A, and $\boldsymbol{r}_0$ and $\boldsymbol{r}_k$ be the initial residual and the residual vector obtained by k iterations of the Krylov subspace method, respectively. We also let $P_k(\lambda) \in \mathbb{P}_k$ be the residual polynomial of the Krylov subspace method, i.e., $\boldsymbol{r}_k = P_k(A)\boldsymbol{r}_0, P_k(0) = 1$. Then we have

$$\|\boldsymbol{r}_k\|_2 \leq \max_{i=1,2,\ldots,n} |P_k(\lambda_i)| \|X\|_2 \|X^{-1}\|_2 \|\boldsymbol{r}_0\|_2.$$

Thus, the convergence rate of the residual 2-norm depends strongly on the magnitude of the residual polynomial corresponding to each eigenvalue λ_i, and it is expected that $P_m(\lambda)$ converge uniformly with respect to each eigenvalue λ_i for efficient convergence.

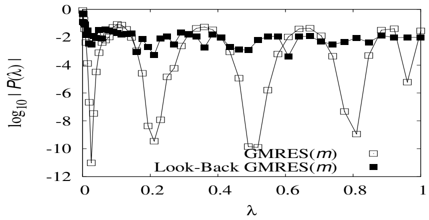
In the following this section, we apply the typical restarted Krylov subspace methods and their Look-Back-type methods: the GMRES(m) method; the FOM(m) method; the Look-Back GMRES(m) method; and the Look-Back FOM(m) method, to the 50 dimensional symmetric positive definite linear system of the form:

$$A\boldsymbol{x} = \boldsymbol{b}, \quad A = \mathrm{diag}\{0.02^2, 0.04^2, \ldots, 1.00^2\}, \quad \boldsymbol{b} = [1, 1, \ldots, 1]^{\mathrm{T}}. \tag{4.1}$$

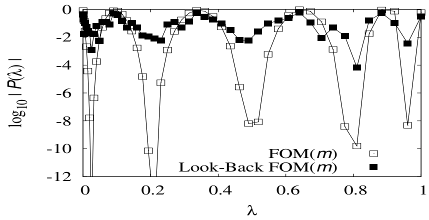
Then we analyze comparison between the magnitude of the residual polynomial $|P_{m\times\ell}(\lambda_i)|$ with respect to each eigenvalue of A and the convergence rate of each method. For solving the model problem (4.1), we set $m = 5$ as the restart frequency, $d = 3$ as the parameter for the Look-Back-type restart and $\boldsymbol{x}_0^{(1)} = [0, 0, \ldots, 0]^{\mathrm{T}}$ for the initial guess.

The log-plot of $|P_{m\times\ell}(\lambda_i)|$ for 10th, 20th, ..., 50th restart cycle of each method are shown in Fig. 4.1, and the residual 2-norm histories are also shown in Fig. 4.2.

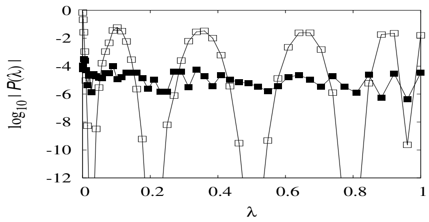
We firstly consider the residual polynomials and the convergence rate of the GMRES(m) method and the FOM(m) method. From the white square □ in Fig. 4.1, we can observe that the residual polynomials $P^{\mathrm{G}}_{m\times\ell}(\lambda), P^{\mathrm{F}}_{m\times\ell}(\lambda)$ have at most 5 (multiple) roots even $\ell \geq 2$. Therefore $|P^{\mathrm{G}}_{m\times\ell}(\lambda_i)|, |P^{\mathrm{F}}_{m\times\ell}(\lambda_i)|$ corresponding to few eigenvalues near the roots are quite small even at early iterations; however, $|P^{\mathrm{G}}_{m\times\ell}(\lambda_i)|$, $|P^{\mathrm{F}}_{m\times\ell}(\lambda_i)|$ corresponding to other eigenvalues show relatively large values even $\ell = 50$. As the results of this non-uniform convergence manner of the residual polynomials $P^{\mathrm{G}}_{m\times\ell}(\lambda_i)$, $P^{\mathrm{F}}_{m\times\ell}(\lambda_i)$ with respect to each eigenvalue, the poor convergence of the GMRES(m) method and the FOM(m) method are caused; see Fig. 4.2.



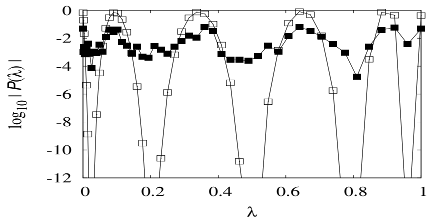
(a) (Look-Back) GMRES(m), $\ell = 10$



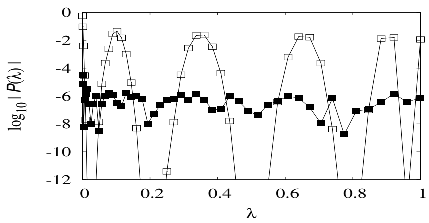
(b) (Look-Back) FOM(m), $\ell = 10$



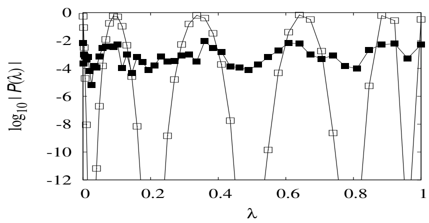
(c) (Look-Back) GMRES(m), $\ell = 20$



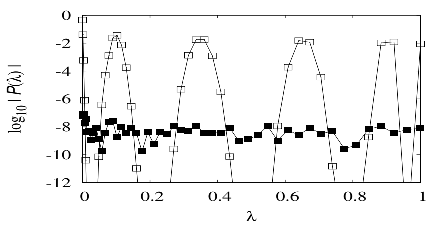
(d) (Look-Back) FOM(m), $\ell = 20$



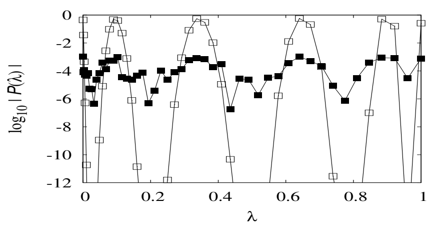
(e) (Look-Back) GMRES(m), $\ell = 30$



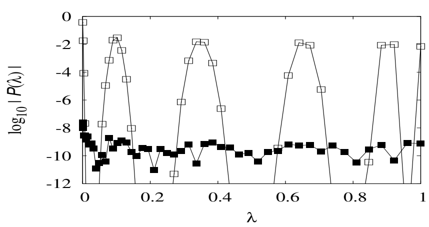
(f) (Look-Back) FOM(m), $\ell = 30$



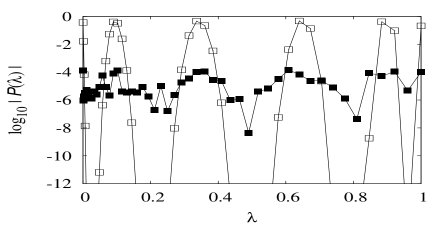
(g) (Look-Back) GMRES(m), $\ell = 40$



(h) (Look-Back) FOM(m), $\ell = 40$

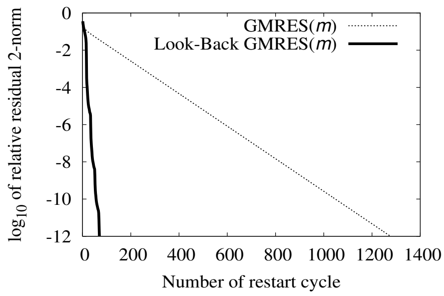


(i) (Look-Back) GMRES(m), $\ell = 50$

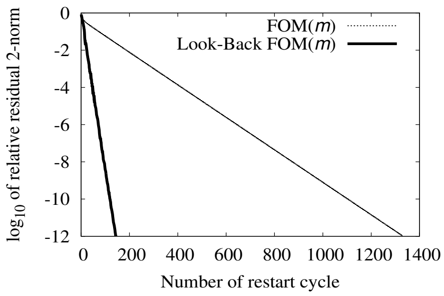


(j) (Look-Back) FOM(m), $\ell = 50$

Fig. 4.1: The graphs of $\log_{10}|P^{\mathrm{G}}_{m\times\ell}(\lambda_i)|$, $\log_{10}|P^{\mathrm{F}}_{m\times\ell}(\lambda_i)|$, $\log_{10}|P^{\mathrm{LB-G}}_{m\times\ell}(\lambda_i)|$ and $\log_{10}|P^{\mathrm{LB-F}}_{m\times\ell}(\lambda_i)|$ for the restart cycle $\ell = 10, 20, \ldots, 50$.



(a) (Look-Back) GMRES(m)



(b) (Look-Back) FOM(m)

Fig. 4.2: The relative residual 2-norm history of the GMRES(m) method, the FOM(m) method, Look-Back GMRES(m) method and the Look-Back FOM(m) method.

Next we consider the residual polynomials and the convergence rate of the Look-Back GMRES(m) method and the Look-Back FOM(m) method. Unlike the non-uniform convergence manner for the GMRES(m) method and the FOM(m) method, the Look-Back-type restart reconstructs the residual polynomial such that $|P^{\mathrm{LB-G}}_{m\times\ell}(\lambda_i)|$ and $|P^{\mathrm{LB-F}}_{m\times\ell}(\lambda_i)|$ show almost the uniform convergence manner in all eigenvalues; see black square ■ in Fig. 4.1. As the results of the uniform convergence manner of the residual polynomials, the Look-Back GMRES(m) method and the Look-Back FOM(m) method have the better convergence for the model problem (4.1) than the GMRES(m) method and the FOM(m) method shown in Fig. 4.2.

We note that we also have almost the same results for other symmetric positive definite linear systems with the different eigenvalue distribution.

5. Numerical experiments and results. In this section, we test the performance of the GMRES(m) method and the Look-Back GMRES(m) method with no preconditioners in Section 5.1 and with the ILU(0) preconditioner [10] in Section 5.2 respectively. We also examine, in Section 5.3, the relationship between the symmetry property of the coefficient matrix and the convergence rate of these methods.

All the numerical experiments were carried out in double precision arithmetic on OS: CentOS 64bit, CPU: Intel Xeon X5550 2.67GHz (1 core), Memory: 48GB, Compiler: GNU Fortran ver. 4.1.2, Compile option: -O3.

5.1. Numerical experiments I. In this section, we test the performance of the GMRES(m) method and the Look-Back GMRES(m) method without preconditioners. The performance of these methods is evaluated by the test problems from The University of Florida Sparse Matrix Collection [7].

The characteristics of the coefficient matrices of the test problems are shown in Table 5.1. (R) or (C) denotes the matrix type: Real nonsymmetric or Complex non-Hermitian respectively. n, Nnz and $Ave.Nnz$ denote the number of dimension, the number of nonzero elements and the average nonzero elements per row or column respectively.

We set $\boldsymbol{b} = [1, 1, \dots, 1]^{\mathrm{T}}$ as the right-hand side, $\boldsymbol{x}_0^{(1)} = [0, 0, \dots, 0]^{\mathrm{T}}$ for the initial guess. We also set $m = 30, 100$ as the restart frequency, $d = 3$ as the parameter for the Look-Back-type restart and stopping criterion was set as $\|\boldsymbol{r}_k\|_2/\|\boldsymbol{b}\|_2 \le 10^{-10}$.

Table 5.1: Characteristics of the coefficient matrices of the test problems for the GMRES(m) method and the Look-Back GMRES(m) method.

(Type) Matrix name	n	Nnz	$Ave.Nnz$	Application area
(R) CAVITY10	2597	76367	29.41	Computational fluid dynamics
(R) CHIPCOOL0	20082	281150	14.00	Model reduction problem
(R) COUPLED	11341	98523	8.69	Circuit simulation
(R) EPB1	14734	95053	6.45	Thermal problem
(R) EX28	2603	77781	29.88	Computational fluid dynamics
(R) FLOWMETER5	9669	67391	6.97	Model reduction problem
(C) KIM1	38415	933195	24.29	2D/3D problem
(C) LIGHT_IN_TISSUE	29282	406084	13.87	Electromagnetics problem
(R) NS3DA	20414	1679599	82.28	Computational fluid dynamics
(R) RAEFSKY1	3242	294276	90.77	Computational fluid dynamics
(R) RAJAT03	7602	32653	4.30	Circuit simulation
(R) RDB5000	5000	29600	5.92	Computational fluid dynamics
(C) WAVEGUIDE3D	21036	303468	14.43	Electromagnetics problem
(R) XENON1	48600	1181120	24.30	Materials problem
(R) XENON2	157464	3866688	24.56	Materials problem
(C) YOUNG1C	841	4089	4.86	Acoustics problem

Numerical results. The numerical results are presented in Tables 5.2–5.3. In these tables, a † denotes that the methods did not converge within 100000 iterations. We also present in Fig. 5.1 the relative residual 2-norm histories of both methods with $m = 30$ for KIM1, LIGHT_IN_TISSUE, NS3DA, RAJAT03, RDB5000 and XENON2. We analyze the results in terms of three aspects: convergence rate; computation time per restart cycle (m iterations); and total computation time.

We firstly consider the convergence rate of both methods. In terms of the number of iterations (Iter), the Look-Back GMRES(m) method shows almost the same or lower Iter than the GMRES(m) method in most cases. Especially, for CAVITY10 ($m = 30$), CHIPCOOL0 ($m = 30, 100$), COUPLED($m = 30$), EX28 ($m = 30, 100$), FLOWMETER5 ($m = 30, 100$) and RAJAT03 ($m = 30$), the GMRES(m) method did not converge within 100000 iterations; on the other hand, the Look-Back GMRES(m) method converged to the solution satisfying the required accuracy $\|\boldsymbol{r}_k\|_2/\|\boldsymbol{b}\|_2 \leq 10^{-10}$; see TRR of Tables 5.2–5.3. We can see from these results that the Look-Back-type restart has high potential to improve significantly the convergence rate of the GMRES(m) method. We can also see from the comparison between the numerical results for $m = 30, 100$ that the smaller restart frequency m leads to the larger improvements in terms of Iter.

We can also see from Fig. 5.1 that the Look-Back GMRES(m) method shows the monotonic decrease in the residual 2-norm as well as the GMRES(m) method. This means that Proposition 3.1 is experimentally supported by these results. For KIM1, NS3DA and RDB5000, both methods show the same level of convergence throughout the whole iteration; see Fig. 5.1 (a), (c), (e). On the other hand, the Look-Back-type restart well played for LIGHT_IN_TISSUE, RAJAT03 and XENON2, and then the Look-Back GMRES(m) method shows a very well convergence property compared with the GMRES(m) method throughout the whole iteration; see Fig. 5.1 (b), (d), (f).

Next, we consider the computation time per restart cycle (m iterations); see

Table 5.2: Convergence results (Iter : number of iterations, t_{Total} : total computation time, t_{Restart} : computation time per restart cycle, TRR : explicitly computed relative residual 2-norm) of the GMRES(m) method and the Look-Back GMRES(m) method for $m = 30$.

Matrix	Method	Iter	Time[sec.]		TRR
			t_{total}	t_{restart}	
CAVITY10	GMRES(m)	†	†	7.79×10^{-3}	-9.37
	LB-GMRES(m)	27961	7.40×10^{0}	7.94×10^{-3}	-10.00
CHIPCOOL0	GMRES(m)	†	†	5.00×10^{-2}	-0.43
	LB-GMRES(m)	34321	5.80×10^{1}	5.05×10^{-2}	-10.00
COUPLED	GMRES(m)	†	†	2.43×10^{-2}	-6.36
	LB-GMRES(m)	78392	6.43×10^{1}	2.46×10^{-2}	-10.00
EPB1	GMRES(m)	2826	2.53×10^{0}	2.69×10^{-2}	-10.00
	LB-GMRES(m)	1993	1.85×10^{0}	2.81×10^{-2}	-10.00
EX28	GMRES(m)	†	†	7.85×10^{-3}	-0.82
	LB-GMRES(m)	25321	6.93×10^{0}	8.23×10^{-3}	-10.01
FLOWMETER5	GMRES(m)	†	†	1.78×10^{-2}	-0.38
	LB-GMRES(m)	78091	4.68×10^{1}	1.80×10^{-2}	-10.00
KIM1	GMRES(m)	2820	2.79×10^{1}	3.02×10^{-1}	-10.01
	LB-GMRES(m)	3963	4.02×10^{1}	3.06×10^{-1}	-10.01
LIGHT_IN_TISSUE	GMRES(m)	2964	1.80×10^{1}	1.82×10^{-1}	-10.00
	LB-GMRES(m)	938	5.81×10^{0}	1.86×10^{-1}	-10.00
NS3DA	GMRES(m)	2330	1.49×10^{1}	1.92×10^{-1}	-10.00
	LB-GMRES(m)	2317	1.51×10^{1}	1.96×10^{-1}	-10.00
RAEFSKY1	GMRES(m)	5036	3.18×10^{0}	1.89×10^{-2}	-10.00
	LB-GMRES(m)	2551	1.62×10^{0}	1.90×10^{-2}	-10.04
RAJAT03	GMRES(m)	†	†	1.30×10^{-2}	-0.55
	LB-GMRES(m)	16621	7.41×10^{0}	1.32×10^{-2}	-10.02
RDB5000	GMRES(m)	962	2.98×10^{-1}	9.27×10^{-3}	-10.00
	LB-GMRES(m)	1021	3.08×10^{-1}	9.05×10^{-3}	-10.02
WAVEGUIDE3D	GMRES(m)	34991	1.85×10^{2}	1.58×10^{-1}	-10.00
	LB-GMRES(m)	29685	1.57×10^{2}	1.59×10^{-1}	-10.00
XENON1	GMRES(m)	11521	6.20×10^{1}	1.61×10^{-1}	-10.00
	LB-GMRES(m)	1879	1.06×10^{1}	1.70×10^{-1}	-10.00
XENON2	GMRES(m)	15691	2.81×10^{2}	5.37×10^{-1}	-10.00
	LB-GMRES(m)	2372	4.26×10^{1}	5.41×10^{-1}	-10.00
YOUNG1C	GMRES(m)	6026	7.16×10^{1}	4.67×10^{-3}	-10.00
	LB-GMRES(m)	5564	6.66×10^{1}	4.33×10^{-3}	-10.00

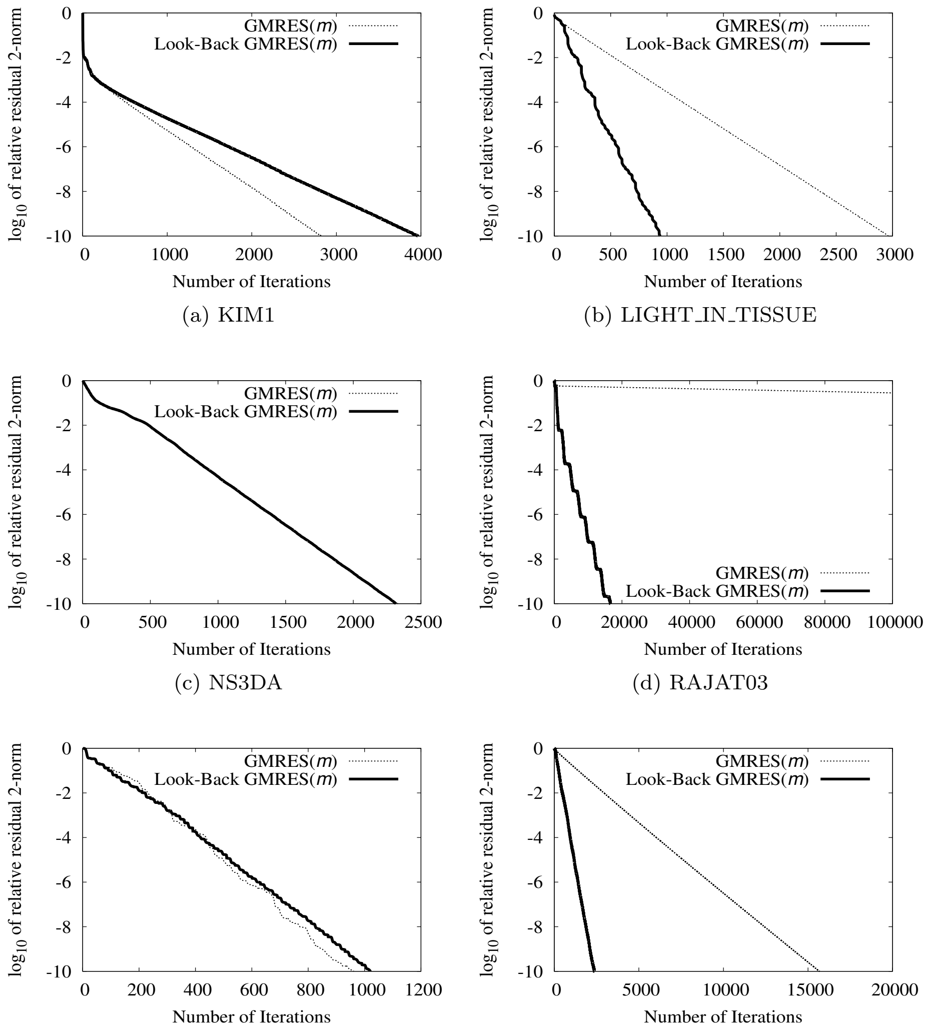
t_{Restart} of Tables 5.2–5.3. In terms of computation time per restart cycle, the Look-Back GMRES(m) method requires at most 10% more time than the GMRES(m) method for both restart frequency m, and the difference becomes smaller with increasing the restart frequency m. This depends on the fact that, in our implementation, the required additional operations for the Look-Back-type restart are just one matrix-vector multiplication and few AXPY and inner-products per restart cycle; see Table 3.1.

Table 5.3: Convergence results (Iter : number of iterations, t_{Total} : total computation time, t_{Restart} : computation time per restart cycle, TRR : explicitly computed relative residual 2-norm) of the GMRES(m) method and the Look-Back GMRES(m) method for $m = 100$.

Matrix	Method	Iter	Time[sec.]		TRR
			t_{total}	t_{restart}	
CAVITY10	GMRES(m)	56098	2.56×10^{1}	4.55×10^{-2}	-10.00
	LB-GMRES(m)	5001	2.26×10^{0}	4.52×10^{-2}	-10.01
CHIPCOOL0	GMRES(m)	†	†	3.56×10^{-1}	-1.62
	LB-GMRES(m)	25301	9.10×10^{1}	3.58×10^{-1}	-10.00
COUPLED	GMRES(m)	13333	2.36×10^{1}	1.77×10^{-1}	-10.00
	LB-GMRES(m)	6115	1.09×10^{1}	1.78×10^{-1}	-10.00
EPB1	GMRES(m)	1781	3.91×10^{0}	2.21×10^{-1}	-10.00
	LB-GMRES(m)	1479	3.24×10^{0}	2.21×10^{-1}	-10.01
EX28	GMRES(m)	†	†	4.63×10^{-2}	-2.19
	LB-GMRES(m)	19601	8.96×10^{0}	4.57×10^{-2}	-10.02
FLOWMETER5	GMRES(m)	†	†	1.42×10^{-1}	-6.08
	LB-GMRES(m)	62901	8.67×10^{1}	1.38×10^{-1}	-10.00
KIM1	GMRES(m)	519	9.38×10^{0}	1.85×10^{0}	-10.03
	LB-GMRES(m)	514	9.29×10^{0}	1.83×10^{0}	-10.00
LIGHT_IN_TISSUE	GMRES(m)	888	1.10×10^{1}	1.24×10^{0}	-10.02
	LB-GMRES(m)	828	1.02×10^{1}	1.25×10^{0}	-10.00
NS3DA	GMRES(m)	1983	1.64×10^{1}	8.26×10^{-1}	-10.01
	LB-GMRES(m)	1959	1.58×10^{1}	8.08×10^{-1}	-10.00
RAEFSKY1	GMRES(m)	2775	2.41×10^{0}	8.72×10^{-2}	-10.01
	LB-GMRES(m)	1703	1.49×10^{0}	8.71×10^{-2}	-10.01
RAJAT03	GMRES(m)	93476	9.56×10^{1}	1.02×10^{-1}	-10.00
	LB-GMRES(m)	5304	5.43×10^{0}	1.03×10^{-1}	-10.00
RDB5000	GMRES(m)	387	2.53×10^{-1}	6.55×10^{-2}	-10.00
	LB-GMRES(m)	380	2.48×10^{-1}	6.54×10^{-2}	-10.03
WAVEGUIDE3D	GMRES(m)	25845	2.58×10^{2}	9.99×10^{-1}	-10.00
	LB-GMRES(m)	23826	2.39×10^{2}	1.00×10^{0}	-10.00
XENON1	GMRES(m)	3888	3.84×10^{1}	9.91×10^{-1}	-10.00
	LB-GMRES(m)	1912	1.90×10^{1}	1.00×10^{0}	-10.00
XENON2	GMRES(m)	4933	1.58×10^{2}	3.20×10^{0}	-10.00
	LB-GMRES(m)	2273	7.28×10^{1}	3.24×10^{0}	-10.00
YOUNG1C	GMRES(m)	1636	4.59×10^{1}	2.80×10^{-2}	-10.00
	LB-GMRES(m)	1493	4.20×10^{1}	2.80×10^{-2}	-10.00

In terms of the total computation time (t_{Total}), from the results of the smaller Iter and almost the same t_{Restart}, we can see that the Look-Back GMRES(m) method can obtain the solution within much smaller computation time than the GMRES(m) method except the case for KIM1 ($m = 30$), NS3DA ($m = 30$) and RDB5000 ($m = 30$).

5.2. Numerical experiments II. In this section, we test the performance of the GMRES(m) method and the Look-Back GMRES(m) method with the ILU(0)



(a) KIM1 (b) LIGHT_IN_TISSUE (c) NS3DA (d) RAJAT03 (e) RDB5000 (f) XENON2

Fig. 5.1: The relative residual 2-norm history of GMRES(m) and Look-Back GMRES(m) of $m = 30$ without preconditioners for KIM1, LIGHT_IN_TISSUE, NS3DA, RAJAT03, RDB5000 and XENON2.

preconditioner, where the right preconditioning was used. The performance of these methods is evaluated by the test problems from the discretization of two types of Partial Differential Equations (PDEs).

For the first example, the test problems are obtained from the discretization of the following PDE:

$$-(Au_x)_x - (Au_y)_y + \alpha \exp(2(x^2 + y^2))u_x = F(x, y),$$

with $\alpha = 0.0, 0.5, 1.0, 2.0$, over the unit square $(x, y) \in [0, 1] \times [0, 1]$ with the Dirichlet boundary condition. The boundary condition and functions A and F are shown in

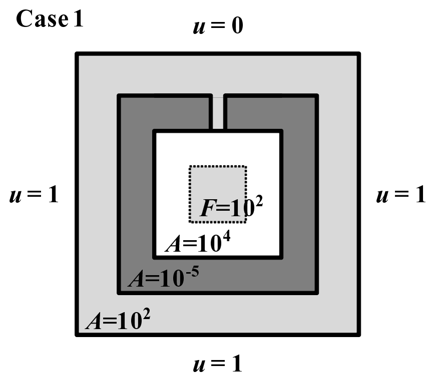


Fig. 5.2: The boundary condition and the functions A and F of van der Vorst's PDE for numerical experiments II. (case 1)

Fig. 5.2. These problems were taken from van der Vorst's paper [20], and some authors have also used to test the performance of the Krylov subspace methods, e.g., [18]. We discretized them by five point central differences over the square grid with size $1/201$ in each directions, and the obtained linear systems have 200^2 unknowns and 199598 nonzero elements.

For the second example, the test problems are obtained from the discretization of the Helmholtz equation [5]

$$u_{xx} + u_{yy} + \sigma^2 u = 0$$

with $\sigma = 0.69, 1.39$, over the square region $(x, y) \in [0, \pi] \times [0, \pi]$. The boundary conditions are defined by

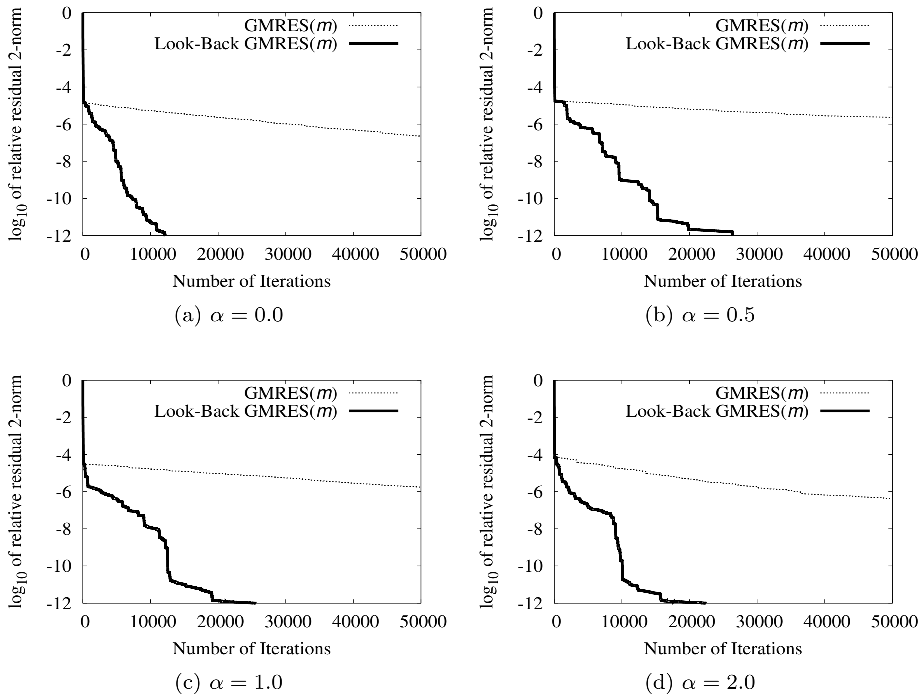
$$\begin{array}{ll} u_x|_{x=0} = \mathbf{i}\left(\sigma^2 - \frac{1}{4}\right)^{1/2} \cos\frac{y}{2}, & \text{(Neumann condition)} \\ u_x - \mathbf{i}\left(\sigma^2 - \frac{1}{4}\right)^{1/2} u|_{x=\pi} = 0, & \text{(Radiation condition)} \\ u_y|_{y=0} = 0, & \text{(Neumann condition)} \\ u|_{y=\pi} = 0, & \text{(Dirichlet condition)} \end{array}$$

where $\mathbf{i}$ denotes $\mathbf{i}^2 = -1$. Note that the exact solution is written by $u(x, y) = \exp\{\mathbf{i}(\sigma^2 - \frac{1}{4})^{1/2} x\} \cos\frac{y}{2}$. We discretized them by five point central differences over the square grid with size $1/201$ in each directions, and the obtained linear systems have (200×201) unknowns and 200596 nonzero elements.

For first and second examples, we set $\boldsymbol{x}_0^{(1)} = [0, 0, \ldots, 0]^{\mathrm{T}}$ for the initial guess, $m = 30$ as the restart frequency, $d = 3$ as the parameter for the Look-Back GMRES(m) method and stopping criterion was set as $\|\boldsymbol{r}_k\|_2/\|\boldsymbol{b}\|_2 \leq 10^{-12}$.

Numerical results. The relative residual 2-norm histories of the ILU(0) preconditioned GMRES(m) method and the ILU(0) preconditioned Look-Back GMRES(m) method for the first example are presented in Fig. 5.3 and for the second example are shown in Fig. 5.4.

We can see from Fig. 5.3 for the van der Vorst's PDE that the residual 2-norm histories of the ILU(0) preconditioned GMRES(m) method stagnated at $\|\boldsymbol{r}_k\|_2/\|\boldsymbol{b}\|_2 \approx 10^{-5}$; on the other hand, the ILU(0) preconditioned Look-Back GMRES(m) method



(a) $\alpha = 0.0$ (b) $\alpha = 0.5$

(c) $\alpha = 1.0$ (d) $\alpha = 2.0$

Fig. 5.3: The relative residual 2-norm history of the GMRES(m) method and the Look-Back GMRES(m) method with the ILU(0) preconditioner for van der Vorst's PDE (case 1) with $\alpha = 0.0, 0.5, 1.0, 2.0$.

converged smoothly for all parameters α. These results show that the Look-Back-type restart has high potential to remedy the stagnation of the residual 2-norm of the GMRES(m) method.

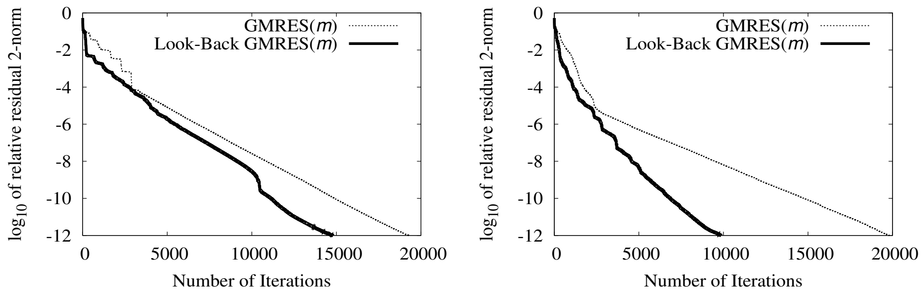
For the Helmholtz equations, the ILU(0) preconditioned Look-Back GMRES(m) method also shows much higher convergence rate than the ILU(0) preconditioned GMRES(m) method; see Fig. 5.4. Especially for the case of $\sigma = 1.39$, the convergence rate of the ILU(0) preconditioned Look-Back GMRES(m) method is more than two times faster as compared with the ILU(0) preconditioned GMRES(m) method.

5.3. Numerical experiments III. In this section, we examine the relationship between the symmetry property of the coefficient matrix and the convergence rate of the GMRES(m) method and the Look-Back GMRES(m) method without preconditioner.

For the test problems, we recall van der Vorst's PDE:

$$-(Au_x)_x - (Au_y)_y + \alpha \exp(2(x^2 + y^2))u_x = F(x, y),$$

with $\alpha = 0.0, 0.1, 0.2, \ldots, 10.0$, over the unit square $(x, y) \in [0, 1] \times [0, 1]$, where we consider two type of the boundary conditions and functions A and F shown in Fig. 5.5. The test problems are obtained from five point central differences over the square grid with size $1/51$ in each directions, and obtained linear systems have 50^2 unknowns and



(a) $\sigma = 0.69$ (b) $\sigma = 1.39$

Fig. 5.4: The relative residual 2-norm history of the GMRES(m) method and the Look-Back GMRES(m) method with the ILU(0) preconditioner for the Helmholtz equation with $\sigma = 0.69, 1.39$.

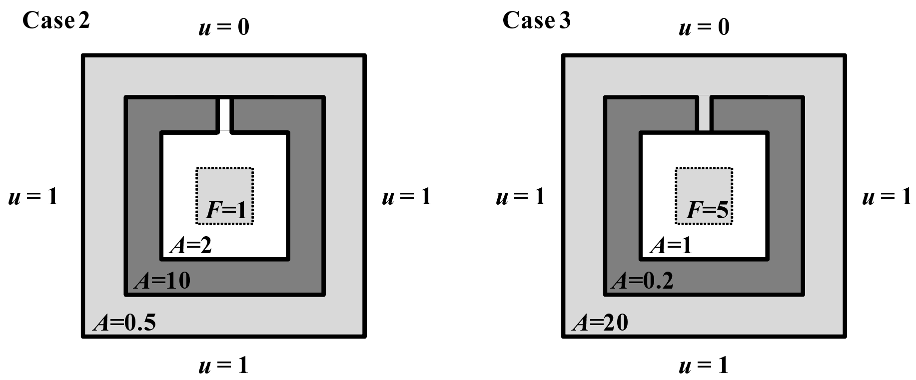


Fig. 5.5: The boundary condition and the functions A and F of van der Vorst's PDE for the numerical experiments III (case 2 and case 3).

12398 nonzero elements. Notice that the obtained linear systems are real symmetric for $\alpha = 0.0$, and the larger α leads to the stronger non-symmetry for the obtained linear systems.

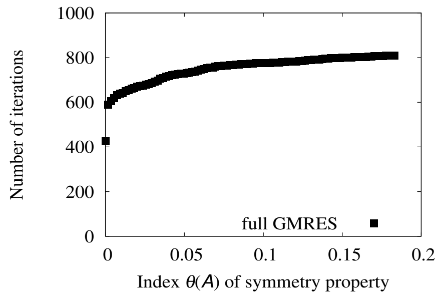
For this experiment, we set $\boldsymbol{x}_0^{(1)} = [0, 0, \ldots, 0]^{\mathrm{T}}$ for the initial guess, $m = 10, 30$ as the restart frequency, $d = 3$ as the parameter for Look-Back GMRES(m) and stopping criterion was set as $||\boldsymbol{r}_k||_2/||\boldsymbol{b}||_2 \leq 10^{-12}$.

Numerical results. We introduce the index $\theta(A)$ for the symmetry property of the matrix $A \in \mathbb{R}^{n\times n}$ defined by

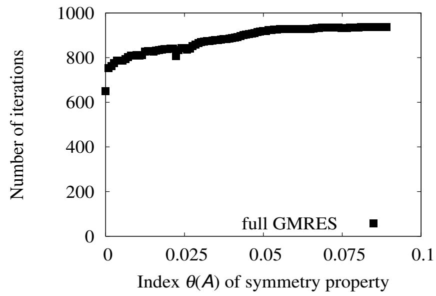
$$\theta(A) := \frac{||(A - A^{\mathrm{T}})/2||_2}{||A||_2},$$

where $0 \leq \theta(A) \leq 1, \theta(A) \in \mathbb{R}$. We note that $\theta(A) = 0$ for the symmetric matrices: $A = A^{\mathrm{T}}$ and $\theta(A) = 1$ for the skew-symmetric matrices: $A = -A^{\mathrm{T}}$.

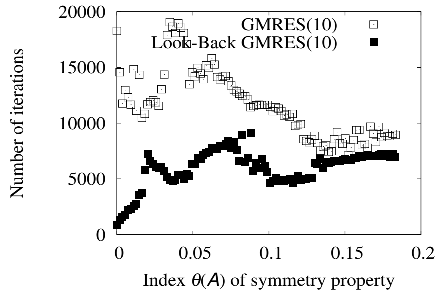
For the numerical results, the graphs of the comparison between the symmetry



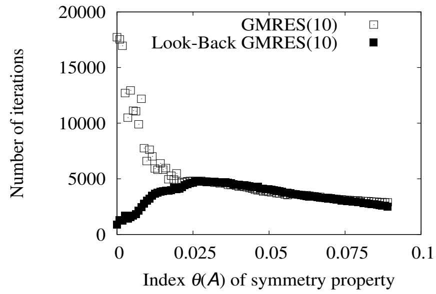
(a) case 1: full GMRES method



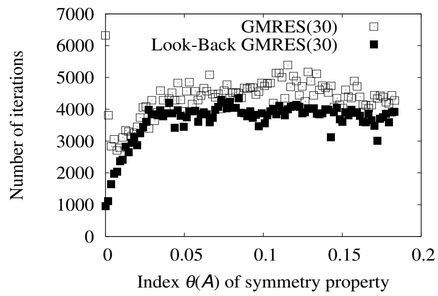
(b) case 2: full GMRES method



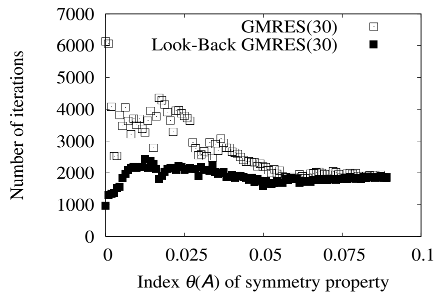
(c) case 1: GMRES(10) method



(d) case 2: GMRES(10) method



(e) case 1: GMRES(30) method



(f) case 2: GMRES(30) method

Fig. 5.6: The graphs of the number of iterations versus the index $\theta(A)$ of the symmetry property for the full GMRES method , the GMRES(m) method and the Look-Back GMRES(m) method for van der Vorst's PDE (case 2 and case 3) with $\alpha = 0, 0.1, 0.2, \ldots, 10.0$.

property $\theta(A)$ of the coefficient matrix and the number of iterations of both methods are presented in Fig. 5.6 (c)–(f). The white square □ denotes the results of the GMRES(m) method and the black square ■ denotes the results of the Look-Back GMRES(m) method respectively. We also show the numerical results for the full GMRES method (without restart) in Fig. 5.6 (a), (b), for comparison. The left column of Fig. 5.6 shows the results for the case 2 of the van der Vorst's PDE and the right column shows for the case 3 respectively.

It is generally expected that the smaller $\theta(A)$ (near symmetric matrices) leads to the better convergence rate. This is experimentally supported by the numerical results of the full GMRES method shown in Fig. 5.6 (a), (b). The number of iterations of the full GMRES method increases almost monotonically with increasing $\theta(A)$.

However, the convergence rate of the GMRES(m) method becomes worse especially for the case of the smaller $\theta(A)$ which leaded to the better convergence for the full GMRES method; see □ of Fig. 5.6 (c)–(f). This means that the restart slows the convergence rate especially for near symmetric cases.

On the other hand, we can obtained the results that the Look-Back-type restart remedies this difficulty. The number of iterations of the Look-Back GMRES(m) method decreases in many cases especially for the smaller $\theta(A)$ as compared with the GMRES(m) method, then the Look-Back GMRES(m) method shows the better convergence with decreasing $\theta(A)$ as well as the full GMRES method; see ■ of Fig. 5.6 (c)–(f).

6. Conclusion. In this paper, we investigated the restarted Krylov subspace methods, as typified by the GMRES(m) method and the FOM(m) method, for solving non-Hermitian linear systems. We proposed the Look-Back-type restart based on the implicit residual polynomial reconstruction via the initial guess for the restarted Krylov subspace method.

From our comparison analysis and numerical experiments, we leaned that the Look-Back-type restart can reconstruct the residual polynomial uniformly and it modifies the convergence property of the restarted Krylov subspace methods in many problems. Especially for the near symmetric cases, the Look-Back-type restart significantly improves the convergence rate of the restarted Krylov subspace methods.

Acknowledgments. This work is supported in part by Strategic Programs for Innovative Research (SPIRE) Field 5 "The origin of matter and the universe" and KAKENHI (Grant No. 25870099).

REFERENCES

[1] James Baglama, Daniela Calvetti, Gene H. Golub, and Lothar Reichel, *Adaptively preconditioned GMRES algorithms*, SIAM Journal on Scientific Computing, 20 (1999), pp. 243–269.

[2] James Baglama and Lothar Reichel, *Augmented GMRES-type methods*, Numerical Linear Algebra with Applications, 14 (2007), pp. 337–350.

[3] Allison H. Baker, Elizabeth R. Jessup, and Tzanio V. Kolev, *A simple strategy for varying the parameter in GMRES(m)*, Journal of Computational and Applied Mathematics, 230 (2009), pp. 751–761.

[4] Allison H. Baker, Elizabeth R. Jessup, and Thomas Manteuffel, *A technique for accelerating the convergence of restarted GMRES*, SIAM Journal on Matrix Analysis and Applications, 26 (2005), pp. 962–984.

[5] Alvin Bayliss, Charles I. Goldstein, and Eli Turkel, *An iterative method for the Helmholtz equation*, Journal of Computational Physics, 49 (1983), pp. 443–457.

[6] Kevin Burrage and Jocelyne Erhel, *On the performance of various adaptive preconditioned GMRES strategies*, Numerical Linear Algebra with Applications, 5 (1998), pp. 101–121.

[7] Timothy A. Davis, *The University of Florida Sparse Matrix Collection.*

[8] Jocelyne Erhel, Kevin Burrage, and Bert Pohl, *Restarted GMRES preconditioned by deflation*, Journal of Computational and Applied Mathematics, 69 (1996), pp. 303–318.

[9] Akira Imakura, Tomohiro Sogabe, and Shao-Liang Zhang, *An efficient variant of the GMRES(m) method based on the error equations*, East Asian Journal on Applied Mathematics, 2 (2012), pp. 1–22.

[10] J. A. Meijerink and Henk A. van der Vorst, *An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix*, Mathematics of Computation, 31 (1977), pp. 148–162.

[11] Ronald B. Morgan, *A restarted GMRES method augmented with eigenvectors*, SIAM Journal on Matrix Analysis and Applications, 16 (1995), pp. 1154–1171.
[12] ———, *Implicitly restarted GMRES and Arnoldi methods for nonsymmetric systems of equations*, SIAM Journal on Matrix Analysis and Applications, 21 (2000), pp. 1112–1135.
[13] ———, *GMRES with deflated restarting*, SIAM Journal on Scientific Computing, 24 (2002), pp. 20–37.
[14] Kentaro Moriya and Takashi Nodera, *The DEFLATED-GMRES(m, k) method with switching the restart frequency dynamically*, Numerical Linear Algebra with Applications, 7 (2000), pp. 569–584.
[15] Yousef Saad, *Krylov subspace methods for solving large unsymmetric linear systems*, Mathematics of Computation, 37 (1981), pp. 105–126.
[16] ———, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, second ed., 2003.
[17] Yousef Saad and Martin H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM Journal on Scientific and Statistical Computing, 7 (1986), pp. 856–869.
[18] Gerard L. G. Sleijpen and Diederik R. Fokkema, *BiCGStab(l) for linear equations involving unsymmetric matrices with complex spectrum*, Electronic Transactions on Numerical Analysis, 1 (1993), pp. 11–32.
[19] Maria Sosonkina, Layne T. Watson, Rakesh K. Kapania, and Homer F. Walker, *A new adaptive GMRES algorithm for achieving high accuracy*, Numerical Linear Algebra with Applications, 5 (1998), pp. 275–297.
[20] Henk A. van der Vorst, *Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM Journal on Scientific and Statistical Computing, 13 (1992), pp. 631–644.
[21] ———, *Iterative Krylov Methods for Large Linear Systems*, Cambridge University Press, New York, NY, 2003.
[22] Lingie Zhang and Takashi Nodera, *A new adaptive restart for GMRES(m) method*, Australian & New Zealand Industrial and Applied Mathematics Journal, 46 (2005), pp. 409–426.