# Restarted GMRES method Augmented with the Combination of Harmonic Ritz Vectors and Error Approximations

Qiang Niu  and  Linzhang Lu

*Abstract*—Restarted GMRES methods augmented with approximate eigenvectors are widely used for solving large sparse linear systems. Recently a new scheme of augmenting with error approximations is proposed. The main aim of this paper is to develop a restarted GMRES method augmented with the combination of harmonic Ritz vectors and error approximations. We demonstrate that the resulted combination method can gain the advantages of two approaches: (i) effectively deflate the small eigenvalues in magnitude that may hamper the convergence of the method and (ii) partially recover the global optimality lost due to restarting. The effectiveness and efficiency of the new method are demonstrated through various numerical examples.

*Keywords*—Arnoldi process, GMRES, Krylov subspace, Systems of Linear equations

## I. INTRODUCTION

THE generalized minimum residual (GMRES) method [24] is one of the most popular iterative methods for solving large sparse linear systems of equations

$$Ax = b, \qquad (1)$$

where $A \in \mathcal{R}^{n \times n}$ is nonsingular, and $b, x \in \mathcal{R}^n$. In order to limit the computational costs and storage requirements, the method is usually restarted after a fixed number of iterations. The resulted method is called restarted GMRES method, denoted by GMRES($m$). In the context of GMRES($m$), it is generally recognized that previous subspace information is lost due to restarting, which makes the global subspace dimension decrease, so that the global convergence of GMRES($m$) is impaired [29]. To alleviate the unwilling phenomena of convergence deterioration, several augmented or deflated restarting schemes have been proposed to accelerate the convergence, see excellent review papers [13], [31] for a comprehensive survey of acceleration techniques for Kryov subspace methods. The main idea of these methods is to utilize judiciously chosen subspace information from previous generated subspaces. The information is then used to either augment the next approximate Krylov subspace [2], [4], [8], [9], [16], [19], [20], [21], [23], [27], [28] or construct effective preconditioners [1], [3], [15], [18], [34].

R. B. Morgan [19] proposed an augmented restarting technique called GMRES-E. The author suggested to retain some fixed number of approximate eigenvectors corresponding to

Q. Niu is with Mathematics and Physics Centre, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China. e-mail:kangniu@gmail.com.
L.-Z. Lu is with School of Mathematical Sciences, Xiamen University, Xiamen 361005, China.

the smallest eigenvalues in magnitude and add them to the new subspace in the next restart cycle. It is shown that the approach is efficient if convergence of GMRES($m$) is hampered by a few well separated eigenvalues close to zero [8], [19]. R. B. Morgan [20] further show that the approximate eigenvectors can be implicitly included in the Krylov subspace, and presented an implicitly restarted GMRES (GMRES-IR) method. In M. Eiermann et al. [13], R. B. Morgan [20], [21], the convergence properties of augmented restarted GMRES method is discussed and they point out that augmenting with harmonic Ritz vectors is important; it generally leads to better convergence results than using other approximate eigenvectors.

Another kind of restarting procedure is enriching the Krylov subspace at each restart cycle such that the approximate subspace gradually includes the error approximation. In the context of inner-outer iteration, this strategy is investigated by van der Vorst and C. Vuik [33]. The method is called GMRESR, which uses GCR [14] as the outer iteration. This approach is further amended by E. Der Sturler [10], [11]. Besides, Y. Saad [26] developed a flexible preconditioned method (FGMRES). Recently, a simplified version of this accelerating approach called LGMRES is proposed in [2] by A. Baker et al. In the LGMRES method, the $(k + 1)$th approximate solution $x_{k+1}$ is constructed by using the Krylov subspace augmented with $l$ recently generated error approximations $z_j$, $j = k, \ldots, k - l + 1$, where $z_j = x_j - x_{j-1}$ and $x_j$ is the approximate solution obtained at the end of the $j$th restart cycle ($k - l + 1 \le j \le k$). Since the vectors $z_j$ is the correction vector computed by the method from the $j$th approximate subspace. So it represents the condensed useful subspace information generated during the $j$th cycle, while it is discarded in subsequent cycle of iterations. Collecting these error approximations and adding them into the next approximate subspace, then the resulted method can partially recover the global optimality of the GMRES method. Based on the polynomial expression of the approximate solution, A. Baker et al. has also showed that the practical LGMRES method can be categorized as a truncated polynomial preconditioned conjugate gradient method. A large number of experiments on a variety of problems have shown that LGMRES is a promising accelerating strategy.

The main aim of this paper is developing an accelerate the convergence of GMRES($m$) by augmenting with the combination of harmonic Ritz vectors and error approximations. We demonstrate that for a class of problems, the resulted combination method is able to gain the advantages of two approaches.

The effectiveness of the new method is demonstrated through various numerical examples. The outline of this paper is as follows. In Section 2, we review GMRES-E and LGMRES methods, and discussed the geometric properties of the residual vectors of two methods. In Section 3, we develop our new method: LGMRES-E. In Section 4, numerical examples are given to demonstrate the effectiveness of the new method. Finally, concluding remarks are given in Section 5.

Throughout the paper, denote by

$$\mathcal{K}_m(A, v_1) = \text{span}\{v_1, Av_1, \ldots, A^{m-1}v_1\}$$

the $m$-dimensional Krylov subspace generated by $A$ and $v_1$, by $\mathcal{R}^m$ the $m$-dimensional real space. The inner product $\langle \cdot, \cdot \rangle$ is the Euclidean inner product, and the norm $\|\cdot\|$ denotes both the Euclidean vector norm and the subordinate spectral matrix norm. Some MATLAB notations and inner functions are used whenever necessary.

## II. AUGMENTED RESTARTED GMRES METHODS

### A. GMRES-E

GMRES-E method proposed by R. B. Morgan [19] attempts to improve the convergence of restarted GMRES method by appending approximate eigenvectors to the Krylov subspace. In this paper, we use GMRES-E$(m, d)$ to denote the restarted GMRES method augmented with harmonic Ritz vectors. Precisely, it represents that the correction vector is found in a $(m + d)$ dimensional approximate subspace constituted by the $m$ dimensional standard Krylov subspace and $d$ harmonic Ritz vectors.

At the end of $(k+1)$th restart cycle, GMRES-E$(m, d)$ seeks the approximate solution $x_{k+1}$ of the form

$$x_{k+1} = x_k + z_{k+1}$$

with

$$z_{k+1} \in \mathcal{K}_m(A, r_k) + \{\varphi_j\}_{j=1}^d,$$

where $\{\varphi_j\}_{j=1}^d$ is a group of vectors constituted by $d$ harmonic Ritz vectors associated with $d$ smallest (in magnitude) harmonic Ritz values. If $\varphi_j$ is complex, then $\bar{\varphi}_j$ is also an harmonic Ritz vector of $A$. In this case, we should take the real part and imaginary part of $\varphi_j$ instead of $\varphi_j$ and $\bar{\varphi}_j$ for practical applications.

Suppose columns of $V_{m+1}$ are the basis vectors constructed by the Arnoldi process with starting vector $v_1 = r_k/\|r_k\|$, and $s = m + d$. Let

$$\tilde{W}_s = [v_1, \ldots, v_m, \varphi_1, \ldots, \varphi_d]$$

and

$$\tilde{Q}_{s+1} = [v_1, \ldots, v_m, v_{m+1}, \ldots, v_{s+1}],$$

where $v_{j+m+1}$, $1 \le j \le d$, is computed by orthogonalizing the vector $A\varphi_j$ against the first $m+j$ columns of $\tilde{Q}_{s+1}$. Then

$$A\tilde{W}_s = \tilde{Q}_{s+1}\tilde{H}_s,$$

where $\tilde{H}_s$ is an $(s+1)$-by-$s$ upper Hessenberg matrix with its elements constructed by the orthogonalization process. Then the $(k+1)$th approximate solution can be formed by

$$x_{k+1} = x_k + \tilde{W}_s y_{k+1},$$

where $y_{k+1}$ is computed by solving

$$\|r_{k+1}\| = \|\beta e_1 - \tilde{H}_s y_{k+1}\| = \min_{y \in \mathcal{R}^s} \|\beta e_1 - \tilde{H}_s y\|. \quad (2)$$

It is easy to see that (2) is equivalent to

$$r_{k+1} \perp A\tilde{W}_s. \quad (3)$$

The GMRES-E$(m, d)$ method can be described as follows.

*Algoritm* 1. *The kth cycle of GMRES-E$(m, d)$ method.*
1) $r_k = b - Ax_k$, $\beta = \|r_k\|$, $v_1 = r_k/\beta$, $s = m + d$;
2) *for* $j = 1, 2, \ldots, s$
3) *if* $j \le m$, *then* $w = Av_j$, *else* $w = A\varphi_{j-m}$;
4) *for* $i=1:j$,
5) $\tilde{h}_{ij} = (w, v_i)$;
6) $w = w - \tilde{h}_{ij}v_i$;
7) *end*
8) $\tilde{h}_{j+1,j} = \|w\|$;
9) *if* $\tilde{h}_{j+1,j} = 0$, *then stop*;
10) $v_{j+1} = w/\tilde{h}_{j+1,j}$;
11) *end*
12) *Compute* $y_{k+1} = arg\min_{y \in \mathcal{C}^s} \|\beta e_1 - \tilde{H}_s y\|$, *form* $x_{k+1} = x_k + \tilde{W}_s y_{k+1}$;
13) *Compute the harmonic Ritz vectors* $\varphi_j$, $j = 1, \ldots, d$;
14) *Compute* $r_{k+1} = b - Ax_{k+1}$, *if* $\|r_{k+1}\| < tol$, *then stop*;
15) $k = k + 1$, *goto 1.*

In practical implementation, reorthogonalization is needed in the orthonormalization process of the algorithm, We refer to [19] for the implementation details.

In the context of the GMRES$(m)$ method, the geometric relationships between sequential and skip residual vectors are investigated in [2], [29]. Following the definition given in [2], we refer to the angles between two consecutive residual vectors as the sequential angle, denoted by $\angle(r_{k+1}, r_k)$; and the angles between every other residual vectors as the skip angles, denoted by $\angle(r_{k+1}, r_k)$. Subsequently, we investigate the geometry relationships between GMRES-E residual vectors. These studies can be regarded as the straightforward generalizations of the results in [2], and their proof are patterned on the proof of Theorem 4-5 in [2]. We included them for completeness.

*Proposition 1: Let $r_{k+1}$ and $r_k$ be the $(k+1)$th and the kth residual vectors from GMRES-E method. Then their sequential angles are given by*

$$\cos \angle(r_{k+1}, r_k) = \frac{\|r_{k+1}\|}{\|r_k\|}. \quad (4)$$

*Proof:* At the $(k + 1)$th cycle of GMRES-E method, we seek the approximate solution of form $x_{k+1} = x_k + z_{k+1}$ with $z_{k+1} \in \mathcal{K}_m(A, r_k) + \{\varphi_j\}_{j=1}^d$. Therefore the corresponding residual vectors satisfies

$$r_{k+1} = r_k - Az_{k+1}.$$

From (3) we have $r_{k+1} \perp Az_{k+1}$, so

$$\langle r_{k+1}, r_k \rangle = \|r_{k+1}\|^2.$$

TABLE I
SOME RECORDED SEQUENTIAL AND SKIP ANGLES OF GMRES-E(12,3)
TESTED ON MATRIX SHERMAN1

| k | 20 | 30 | 40 | 50 | 60 | 70 |
|---|----|----|----|----|----|----|
| $\angle(r_k, r_{k-2})$ | 8.1 | 4.5 | 4.8 | 3.4 | 0.8 | 0.5 |
| $\angle(r_k, r_{k-1})$ | 39.1 | 41.2 | 39.4 | 36.9 | 40.5 | 42.6 |

TABLE II
SOME RECORDED SEQUENTIAL AND SKIP ANGLES OF GMRES-E(12,3)
TESTED ON MATRIX E05R0000

| k | 20 | 30 | 40 | 50 | 60 | 70 |
|---|----|----|----|----|----|----|
| $\angle(r_k, r_{k-2})$ | 7.6 | 3.4 | 3.1 | 2.5 | 2.2 | 1.8 |
| $\angle(r_k, r_{k-1})$ | 26.2 | 24.4 | 24.7 | 25.4 | 28.1 | 26.0 |

Based on the definition of cosine, then relationship (4) follows. ∎

*Proposition 2:* Let $r_{k+1}$ and $r_{k-1}$ be the $(k+1)$th and the $(k-1)$th residual vectors from GMRES-E method. Then their skip angles are given by

$$\cos \angle (r_{k+1}, r_{k-1}) = \frac{||r_{k+1}||}{||r_{k-1}||} - \frac{\langle Az_k, Az_{k+1}\rangle}{||r_{k+1}||||r_{k-1}||}, \quad (5)$$

*Proof:* Let $x_{k+1} = x_k + z_{k+1}$ and $x_k = x_{k-1} + z_k$, then we have

$$r_{k+1} = r_k - Az_{k+1} \quad (6)$$

and

$$r_k = r_{k-1} - Az_k. \quad (7)$$

From (7), it follows that

$$\langle r_{k+1}, r_k \rangle = \langle r_{k+1}, r_{k-1}\rangle - \langle Az_k, r_{k+1}\rangle.$$

As *Proposition 1* reveals that $\langle r_{k+1}, r_k \rangle = \langle r_{k+1}, r_{k+1}\rangle$, so we have

$$
\begin{aligned}
\langle r_{k+1}, r_{k-1}\rangle &= \langle r_{k+1}, r_{k+1}\rangle + \langle Az_k, r_{k+1}\rangle \\
&= \langle r_{k+1}, r_{k+1}\rangle + \langle Az_k, r_k - Az_{k+1}\rangle \\
&= \langle r_{k+1}, r_{k+1}\rangle - \langle Az_k, Az_{k+1}\rangle.
\end{aligned}
\quad (8)
$$

The proof is completed. ∎

From *Proposition 1* we can see that convergence rate of the GMRES-E method can be learned from the cosine values of the sequential angles. Particularly, relationship (4) indicates that if the sequential residual vectors preserve at a large degree, then fast convergence behavior will appear. However, relationship (5) does not reveal such kinds of information in a simple way, as $\langle Az_k, Az_{k+1}\rangle$ also has influence on the reduction of the residual norm. In the context of restarted GMRES method, similar phenomena are discovered and well investigated in [2], [29]. In Section 3, we will present a property between the skip angles of our new algorithm LGMRES-E, and give some comparisons between the above propositions.

In order to study the influence of term $\langle Az_k, Az_{k+1}\rangle$ on skip angles in (5), we record some sequential and skip angles of GMRES-E tested on two matrices "sherman1" and "e05r0000" from [22]. For a sequence of indexes $k$, the corresponding skip and sequence angles are recorded and listed in Table I and Table II. In the test, the residual norms of "sherman1" and

"e05r0000" both reach the $10^{-9}$ after 236 and 148 restarts, respectively. From Tables I- II, we can see that the sequential angles are preserved at a reasonable large degree while the skip angles become quite small, i.e., the directions of the residual vectors often alternate in a cyclic fashion. As the residual vector will be normalized as the starting vector for next cycle, so small skip angles implies that the every other approximate subspaces are close to each other. This phenomena is discovered in [2] in the restarted GMRES method. The occurrence of small skip angles implies that the convergence can be accelerated if this alternating behavior of the residual vectors is corrected. Inspired by the idea of the LGMRES method, we will present a combinative augmentation technique such that both the sequential angles and the skip angles can be kept at reasonable large degrees.

*B. LGMRES*

Implementation of the LGMRES method is similar to the GMRES-E method. At the end of each restart cycle, the LGMRES$(m, l)$ method finds the correction vector in an $(m + l)$ dimensional subspace constituted by the $m$ dimensional standard Krylov subspace and $l$ recently generated error approximations.

Let $\hat{x}$ be the exact solution of (1), and $x_k$ be the approximate solution obtained at the end of $k$th restart cycle. Then the $k$th error defined by

$$e_k = \hat{x} - x_k.$$

It is obvious that if the next approximate subspace incorporates this error vector, then the linear system (1) will be solved in the next restart cycle. The main problem is that $e_k = A^{-1}r_k$ is unknown. So it is natural to come forth the GMRESR family of methods [33], or the schemes of constructing a vector $M_k^{-1}r_k$ to approximate $e_k$ at every restart cycle, which leads to the flexible preconditioned GMRES method [8], [26] (if $M_k$ changes with $k$).

As the correction is originated from last approximate subspace, so it represents in certain sense the approximate subspace information generated in the last cycle. However, the subspace is discarded due to restarting. It is proposed in [2] that $z_k$ should be retained in the subsequent approximate subspace. If $p$ previously generated error approximations are collected, then the $(k + 1)$th cycle of the LGMRES method will seek an approximate solution of the form

$$x_{k+1} = x_k + q_{m-1}(A)r_k + \sum_{j=k-l+1}^{k} \alpha_{kj} z_j,$$

where $q_{m-1}$ is a polynomial of degree at most $m - 1$, which is determined along with $\alpha_{kj}$ such that the norm of $r_{k+1}$ is minimized. The $(k + 1)$th cycle of the LGMRES$(m, l)$ algorithm can be described as follows.

*Algorithm 2. The kth cycle of the LGMRES$(m, l)$ method*
1) *Given $r_k$, $\beta = ||r_k||$, $v_1 = r_k/\beta$, $s = m + l$;*
2) *Orthonormalization process: the same as steps 2 to 11 in Algorithm 1, except revising the third step as: if $j \leq m$, then $w = Av_j$, else $w = Az_{k-(j-m-1)}$;*

3) *Find* $y_{k+1} = arg\min_{y \in \mathcal{R}^s} \|\beta e_1 - \hat{H}_s y\|$, *form* $x_{k+1} = x_k + \hat{W}_s y_{k+1}$; *where elements of $\hat{H}_s$ are generated by the orthogonalization process in step 2 and* $\hat{Q}_{s+1} = [v_1, \ldots, v_m, v_{m+1}, \ldots, v_{s+1}]$, $\hat{W}_s = [v_1, \ldots, v_m, z_k, \ldots, z_{k-l+1}]$;

4) *Form* $z_{k+1} = \hat{W}_s y_{k+1}$; *(also* $Az_{k+1} = \hat{Q}_{s+1}\hat{H}_s y_{k+1}$*)*;

5) *Compute* $r_{k+1} = b - Ax_{k+1}$, *if* $\|r_{k+1}\| < tol$, *then stop;*

6) $k = k + 1$, *goto 1.*

A few implementation details are given in [2], we list them as follows:

- Only $m$ matrix-vector multiplications are required per cycle of iterations if $l$ pairs of $z_j$ and $Az_j$ are stored.
- For $k < l$, there are no enough error vectors, so it is recommended using Arnoldi vectors instead of $z_j$ when $j < 1$ such that the dimension of the approximation subspace can be fixed as $m + l$.
- The optimal values for $l$ are typically very small, generally $l \le 3$.

The next proposition is given in [2], it indicates that the convergence of LGMRES method also correlates the skip angles, and fast convergence implies large skip angles. Experiments in [2] showed that the LGMRES method can preserve both sequential and skip angles at reasonable large degrees, so accelerated convergence behavior is exhibited.

*Proposition 3:* [2]
*Let $r_k$, $r_{k+1}$ and $r_{k-1}$ be the kth, the $(k+1)$th and the $(k-1)$th residual vectors respectively, formed by the LGMRES method. Then the angles between these residual vectors are given by*

$$\cos \angle(r_{k+1}, r_k) = \frac{\|r_{k+1}\|}{\|r_k\|}, \tag{9}$$

$$\cos \angle(r_{k+1}, r_{k-1}) = \frac{\|r_{k+1}\|}{\|r_{k-1}\|}. \tag{10}$$

## III. THE PROPOSED METHOD

From the numerical results in Section 2.1, we can see that the sequential angles of GMRES-E can be retained at moderate degrees while the skip angles are usually pretty small, i.e., every other residual vectors of GMRES-E point to nearly the same direction. According to the analysis of GMRES($m$) in [2], the phenomena of low skip angles implies that some subspace information is lost due to restarting, such that the global optimal property of the full GMRES method is lost. Thereby a faster convergence rate can be achieved if previously generated subspace information can be included in the next approximate subspace.

In this section, we develop a restarted GMRES method augmented with a combination of harmonic Ritz vectors and error approximations. Our idea is to append error approximations to the augmented subspace at the end of every restarted cycle of GMRES-E. We call the new algorithm LGMRES-E and use LGMRES-E($m, d, l$) to denote that at the end of each restart cycle of the method, a new approximate solution is found in an $(m + d + l)$ dimensional subspace constructed by the Krylov subspace augmented with $d$ harmonic Ritz vectors and

$l$ error approximations. For example, the $(k+1)$th approximate solution $x_{k+1}$ is of the form

$$x_{k+1} = x_k + z_{k+1}$$

and $z_{k+1}$ is computed from

$$span\{r_k, Ar_k, \ldots, A^{m-1}r_k, \varphi_1, \ldots, \varphi_d, z_{k-l+1}, \ldots, z_k\},$$

where $\varphi_j$, $1 \le j \le d$, are the harmonic Ritz vectors calculated at the end of the kth cycle of iterations, and $z_j$, $k-l+1 \le j \le k$, are the most recently generated error approximations. When $k < l$, error approximations are not generated, and thereby GMRES-E($m, d + l$) is used instead.

Let $s = m + d + l$, and

$$W_s = [v_1, \ldots, v_m, \varphi_1, \ldots, \varphi_d, z_{k-l+1}, \ldots, z_k],$$

where the first $m$ columns are the orthonormalized Arnoldi basis vectors (obtained by the orthogonalization process of Algorithm 1). Let

$$Q_{s+1} = [v_1, \ldots, v_{m+1}, v_{m+2}, \ldots, v_{s+1}]$$

be $n \times (s+1)$ matrix whose first $m+1$ columns are the Arnoldi basis vectors and whose last $(d + l)$ columns are formed by orthogonalizing the vectors $A\varphi_i$ and $Az_j$, for $1 \le i \le d$ and $k - l + 1 \le j \le k$, against their previous columns of $Q_{s+1}$, respectively. The above procedure can be demonstrated by the following algorithm,

*Algorithm 1:* The $k$th cycle of LGMRES-E($m, d, l$)

1) Given $r_k$, $\beta = \|r_k\|$, $v_1 = r_k/\beta$, $s = m + l + d$;

2) for $j = 1, 2, \ldots, s$;

3) $w = \begin{cases} Av_j, & j \le m \\ A\varphi_{j-m}, & m + 1 \le j \le m + d \\ Az_{j-m-d+k-l}, & m + d + 1 \le j \le m + d + l \end{cases}$

4) for i=1:j,

5) $h_{ij} = (w, v_i)$;

6) $w = w - h_{ij}v_i$;

7) end

8) $h_{j+1,j} = \|w\|$;

9) if $h_{j+1,j} = 0$, then stop;

10) $v_{j+1} = w/h_{j+1,j}$;

11) end

12) Find $y_{k+1} = arg\min_{y \in \mathcal{R}^s} \|\beta e_1 - H_s y\|$, form $x_{k+1} = x_k + W_s y_{k+1}$;

13) Record $z_{k+1} = W_s y_{k+1}$; (also $Az_{k+1} = Q_{s+1}H_s y_{k+1}$);

14) Compute the harmonic Ritz vectors $\varphi_j$, $j = 1, \ldots, d$;

15) Compute $r_{k+1} = b - Ax_{k+1}$, if $\|r_{k+1}\| < tol$, then stop;

16) $k = k + 1$, goto 1.

The above procedures can be summarized as the following relationship

$$AW_s = Q_{s+1}H_s, \tag{11}$$

where $H_s$ is an $(s + 1)$-by-$s$ upper Hessenberg matrix whose nonzero elements $h_{i,j}$ are defined by the orthogonalization process of Algorithm 3.

By using (11), the $(k + 1)$th approximate solution of form

$$x_{k+1} = x_k + W_s y_{k+1},$$

TABLE III
SHERMAN1

| k | 20 | 30 | 40 | 50 |
|---|---|---|---|---|
| $\angle(r_k, r_{k-2})$ | 69.3 | 56.6 | 69.1 | 66.5 |
| $\angle(r_k, r_{k-1})$ | 55.7 | 42.7 | 52.4 | 52.2 |

TABLE IV
E05R0000

| k | 20 | 30 | 40 | 50 |
|---|---|---|---|---|
| $\angle(r_k, r_{k-2})$ | 59.7 | 68.1 | 64.6 | 62.6 |
| $\angle(r_k, r_{k-1})$ | 48.2 | 52.3 | 45.9 | 47.8 |

can be derived by solving the following minimization problem,

$$\|r_{k+1}\| = \|\beta e_1 - H_s y_{k+1}\| = \min_{y \in \mathcal{R}^s} \|\beta e_1 - H_s y\|. \quad (12)$$

It is easy to verify that (12) is equivalent to

$$r_{k+1} \perp AW_s. \quad (13)$$

*Proposition 4:* Let $r_{k+1}$, $r_k$, and $r_{k-1}$ be the residual vectors from LGMRES-E. Then the angles between these residual vectors are given by

$$\cos \angle(r_{k+1}, r_k) = \frac{\|r_{k+1}\|}{\|r_k\|}, \quad (14)$$

$$\cos \angle(r_{k+1}, r_{k-1}) = \frac{\|r_{k+1}\|}{\|r_{k-1}\|}. \quad (15)$$

*Proof:* Define $z_k = W_s y_k$, then it is easy to see that the residual vectors satisfy $r_{k+1} = r_k - Az_{k+1}$ and $r_k = r_{k-1} - Az_k$. As (13) implies $r_{k+1} \perp Az_{k+1}$, so

$$\langle r_{k+1}, r_{k+1} \rangle = \langle r_{k+1}, r_k \rangle.$$

Moreover, from (13) we also have $r_{k+1} \perp Az_k$, so

$$\langle r_{k+1}, r_{k-1} - r_k \rangle = 0.$$

Thus relationships (14) and (15) are proved by the definition of cosine. ∎

The results (14) and (15) show that the residual vectors of LGMRES-E have similar geometric properties with that of LGMRES discussed in [2]. LGMRES-E(12,2,1) is tested on the same matrices as the ones in Table I and Table II. The tested results are shown in Table III and Table IV. By using the same stopping criterion as before, the first linear system "sherman1" converges after 58 restarts and the second linear system "e05r0000" converges after 51 restarts. From Table III and Table IV, it is easy to see that the skip angles of LGMRES-E are generally larger than the corresponding sequence angles, and both kinds of angles can kept at reasonable degrees. This implies that LGMRES-E can avoid the alternating behavior observed in the GMRES-E method, and therefore a faster convergence behavior is expected. In the next section, numerical examples are given to demonstrate the improvement of LGMRES-E method for a class of problems.

## IV. EXPERIMENTAL RESULTS

In this section, we report some numerical results tested on a variety of matrices from Matrix Market Collection [22] and Harwell-Beoing sparse matrix collection [12] and University of Florida Sparse Matrix Collection [7]. All the tests are run on a desktop machine with INTEL PENTIUM IV 2.4GHz CPUs and 256M of main memory, by using MATLAB 7.0.4. Without special explanation, we always use a random vector as the right hand side, and the initial approximation $x_0$ is chosen as a zero vector. We terminated the iterations once the residual vectors $r_k$ satisfies

$$\frac{\|r_k\|}{\|r_0\|} \leq 10^{-8}.$$

All the matrices used in this paper is outlined in Table V. The columns labeled $size$ and $nnz$ are for matrix dimension and number of non-zeros entries. $\lambda_{max}$ refers to the estimate of the largest (in magnitude) eigenvalue of $A$, and $\lambda_{-i}$, $i = 1, 2, 3$, refer to the estimation of three smallest eigenvalues in magnitude, respectively.

In [2], A. Baker et.al has demonstrated that the optimal number $l$ of error approximations used in LGMRES at every cycle of the iterations is typically small, generally less than 3, and the distinction of the performances is actually close to each other for $l \leq 3$. We observe in this paper that the optimal value for $l$ is also small. The number of restarts to make the relative residual norm below $10^{-5}$ is recorded in Table VI, from which we can see that LGMRES-E($m, d, l$) with $l = 1$ produces better results on most of the tested matrices. Therefore we set $l = 1$ for LGMRES-E($m, d, l$) and LGMRES($m, l$) in the following examples. We remark that if larger subspaces are used, then there are indeed cases that LGMRES-E($m, d, l$) with $l > 1$ produces better results. Morover, we observe that GMRES-E and LGMRES-E generally exhibit certain superlinear convergence behavior for most of the tested problems, especially for the problems with well-separated small eigenvalues. This is because both methods can periodically deflate small eigenvalues that hamper the convergence, which improves the convergence of GMRES($m$) [19]. However, the linear independence between every other approximate subspaces is impaired, which can be learned from the small skip angles displayed in Table I and Table II. This phenomena indicates that every other subspace have similar information such that the GMRES-E gain little progress in every other restart cycles. By retaining error approximations, LGMRES-E can partially preserve the linear independence between the sequential and every other approximate subspaces, so that LGMRES-E can improve the performance of GMRES-E. As far as LGMRES method is concerned, we observe that the residual norms of LGMRES decreases steadily, often with an alternating behavior of local slow and local superlinear convergence. The reason may be that small eigenvalues produce continual influence on the convergence rate at the early phase of every cycle. To show the superiority of the new method, we mainly compare the convergence behaviors between LGMRES-E and GMRES-E or LGMRES-E and LGMRES in the following discussions. The convergence curves of three algorithms will be outlined in each figure.

TABLE V
THE MATRICES INFORMATION USED IN THE NUMERICAL TESTS

| $matrix$ | $size$ | $nnz$ | $\lambda_{-1}$ | $\lambda_{-2}$ | $\lambda_{-3}$ |
|---|---|---|---|---|---|
| sherman1 | 1000 | 3750 | $-3.0 \times 10^{-4}$ | $-1.0 \times 10^{-3}$ | $-1.1 \times 10^{-3}$ |
| sherman4 | 1104 | 3786 | 0.03 | 0.09 | 0.28 |
| sherman5 | 3312 | 20793 | 0.05 | 0.13 | 0.40 |
| cavity05 | 1182 | 32747 | $2.3 \times 10^{-5}$ | $9.7 \times 10^{-4}$ | $9.7 \times 10^{-4}$ |
| cavity10 | 2597 | 76367 | $4.4 \times 10^{-6}$ | $4.3 \times 10^{-4}$ | $4.3 \times 10^{-4}$ |
| e05r0000 | 236 | 5856 | $-4 \times 10^{-4}$ | $-3.9 \times 10^{-4}$ | $-3.9 \times 10^{-4}$ |
| saylr4 | 3654 | 22316 | $-1.9 \times 10^{-3}$ | $-2.7 \times 10^{-2}$ | -0.1 |
| memplus | 17758 | 126150 | $1.16 \times 10^{-5}$ | $1.17 \times 10^{-5}$ | $1.18 \times 10^{-5}$ |
| orsreg1 | 2205 | 14133 | -4.91 | -4.98 | -5.2 |
| wang4 | 26068 | 177196 | $4.6 \times 10^{-6}$ | $5.2 \times 10^{-6}$ | $1.5 \times 10^{-5}$ |

TABLE VI
THE PERFORMANCE OF LGMRES-E$(21, 4 - l, l)$ FOR DIFFERENT $l$.

| $l$ | sherman1 | sherman4 | sherman5 | cavity05 | cavity10 | saylr4 |
|---|---|---|---|---|---|---|
| 1 | 19 | 5 | 87 | 38 | 61 | 124 |
| 2 | 18 | 5 | 116 | 52 | 81 | 167 |
| 3 | 21 | 5 | 140 | 51 | 89 | 178 |

**Example 1.**

Three matrices "sherman1", "sherman4" and "sherman5" from Harwell-Beoing sparse matrix collection [12] are tested, and the corresponding right hand sides given in the matrix collection are used. These matrices are all real nonsymmetric, generated in oil reservoir simulation. Table VII shows the tested results for "sherman1" and "sherman4". The matrix "sherman4" can be handled easily even by GMRES(20). It is easy to see that LGMRES-E(17,2,1) behaves slightly better than GMRES-E(17,3) while considerably faster than LGMRES(19,1). The matrix "sherman1" is a relatively difficult one. LGMRES-E(17,2,1) is slightly superior to LGMRES(19,1) while far better than GMRES-E(17,3). As noted in [20], the matrix "sherman5" is a difficult matrix without preconditioning. If the dimension of the approximate subspace is set to be 20, only LGMRES-E(17,2,1) converges within 500 restarts. When enlarging the subspace dimension to be 30, all the algorithms converges as are shown in Figures 1. GMRES-E(27,3) converges within 208 restarts in 71.9 seconds, and LGMRES-E(27,2,1) converges within 117 restarts in 40.1 seconds. LGMRES(29,1) converges steadily but rather more slowly, and the minimal residual norm reaches 3.4e-6 within 300 restarts.

TABLE VII
NUMERICAL RESULTS FOR "SHERMAN1" AND "SHERMAN4"

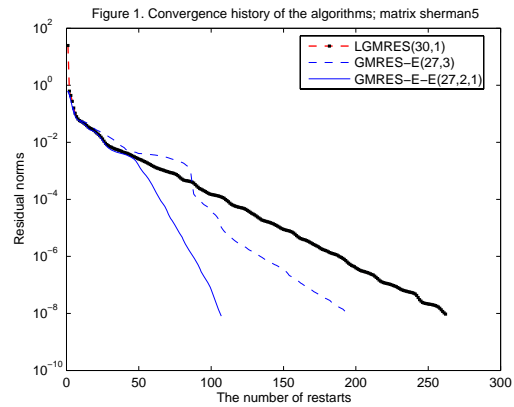| | sherman1 | | sherman4 | |
|---|---|---|---|---|
| algorithms | $iter$ | $time$ | $iter$ | $time$ |
| LGMRES-E(17,2,1) | 35 | 0.71 | 10 | 0.45 |
| GMRES-E(17,3) | 77 | 1.46 | 11 | 0.61 |
| LGMRES(19,1) | 36 | 0.72 | 21 | 0.95 |
| GMRES(20) | 230 | 2.28 | 40 | 1.10 |



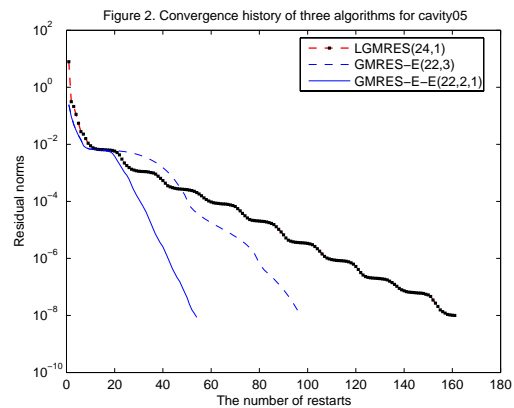Fig. 1. The convergence curves of matrix sherman5.



Fig. 2. The convergence curves of matrix cavity05.

**Example 2.**

We consider two matrices "cavity5" and "cavity10" from [22] arising from finite element modeling. Both matrices are real nonsymmetric with a relative small eigenvalue in magnitude as shown in Table V. The normalized right hand side corresponding to each matrix is used. The convergence results are reported in Figure 2 and Figure 3. It is obvious to see that LGMRES-E exhibits superlinear convergence and always converges faster than the other two methods. LGMRES presents a pattern of alternating between local slow and local steep decrease in the residual norms. As far as CPU timings in seconds are concerned, LGMRES, GMRES-E and LGMRES-E spend 14.1s, 5.5s, 4.2s on "cavity05", and 62.6s, 57.5s, 16.8s on "cavity10", respectively.

**Example 3.**

The 3D oil reservoir simulation matrix "saylr4" from Harwell-Boeing collection [12] is tested in this example. It is a real symmetric matrix. Without preconditioning, this problem is very difficult for GMRES-E. The minimal relative residual norm of GMRES-E(21,4) and GMRES-E(27,3) are $7.8 \times 10^{-4}$ and $6.2 \times 10^{-8}$ within 500 restarts. The tested results of LGMRES and LGMRES-E are listed in Table VIII. We can
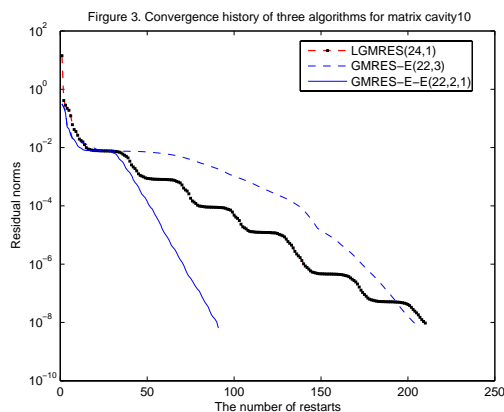
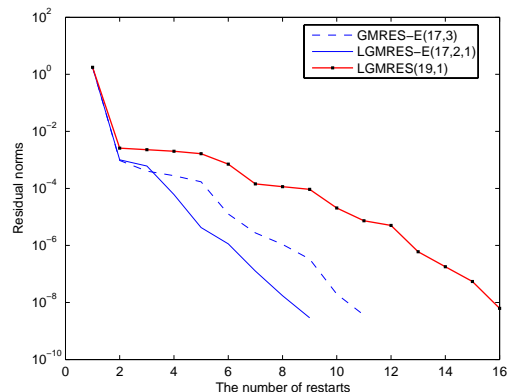Fig. 3.    The convergence curves of matrix cavity10.

see that that LGMRES-E is more efficient, in terms of both the number of restarts and CPU timings.

TABLE VIII
COMPARISON OF LGMRES-E WITH LGMRES

| Algorithm | $iter$ | $time$ | $res$ |
|---|---|---|---|
| LGMRES-E(12,2,1) | 347 | 41.1 | $9.9 \times 10^{-9}$ |
| LGMRES(14,1) | 473 | 87.7 | $9.8 \times 10^{-9}$ |
| LGMRES-E(17,2,1) | 223 | 44.3 | $9.6 \times 10^{-9}$ |
| LGMRES(19,1) | 297 | 78.9 | $9.6 \times 10^{-9}$ |
| LGMRES-E(22,2,1) | 164 | 48.7 | $9.4 \times 10^{-9}$ |
| LGMRES(24,1) | 220 | 83.9 | $9.6 \times 10^{-9}$ |
| LGMRES-E(27,2,1) | 136 | 56.5 | $8.9 \times 10^{-9}$ |
| LGMRES(29,1) | 176 | 92.7 | $9.4 \times 10^{-9}$ |

Example 4.
In this example, we tested a relatively large matrix "wang4" from the University of Florida Sparse Matrix Collection [7]. This matrix is of order 26080 generated from semiconductor device problem. The right hand side $b$ is set to be $A * ones(n, 1)$ such that the exact solution be $ones(n, 1)$. Taking sparse incomplete LU factorization of $A$ with $droptol = 0.1$ as a preconditioner [25], i.e., the preconditioner $luinc(A, 0.1)$ in MATLAB. In Figure 4 we display the convergence curves of three preconditioned methods. In this example the new method exhibits faster convergence behavior with preconditioning. It can be calculated using $eigs$ that the three smallest eigenvalues of the preconditioned system is about 0.0023, 0.0029 and 0.003. So the LGMRES-E and GMRES-E benefit from preconditioning. However, if there are no small eigenvalues close to zero, then augmenting approximate eigenvectors may not be helpful, while adding error approximations can still produce active effect, as is shown by the next example.

Example 5.
Two real nonsymmetric matrices "memplus" and "orsreg1" used in this example are taken from [12]. Figure 5 and Figure 6 illustrate the convergence curves of the three methods. From the two figures we can see that compared with LGMRES and LGMRES-E, the GMRES-E has the worst convergence behavior. LGMRES works best and these results show the



Fig. 4.    The convergence curves of matrix wang4.
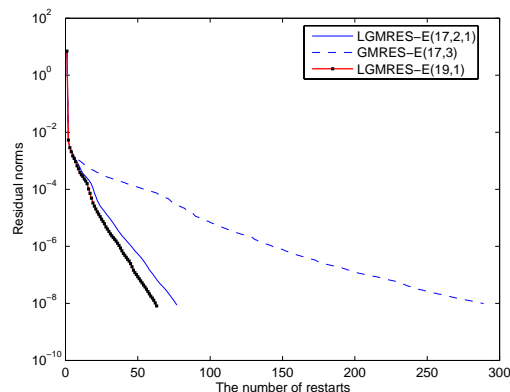


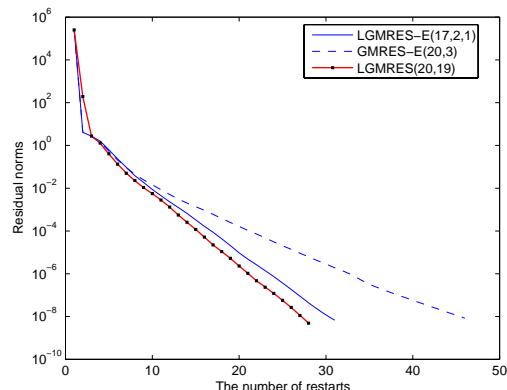Fig. 5.    The convergence curves of matrix memplus.



Fig. 6.    The convergence curves of matrix orsreg1.

effect of using error approximations. In this case, LGMRES-E works slightly worse than LGMRES, but much better than GMRES-E. The explanation for this phenomena is that augmenting approximate eigenvectors does not produce active effect on improving the convergence rate, whereas adding error approximations does. We can see from Table V that the smallest eigenvalues are not well separated for matrix memplus, and there are no the small eigenvalues near zero for matrix "orsreg1".

## V. Conclusions

In this paper we have proposed a combined augmented restarted GMRES method: LGMES-E. The new method can be viewed as an accelerated GMRES-E by using the accelerating technique of LGMRES. With a combined accelerating strategy, the new method can not only effectively deflate the smallest eigenvalues by augmenting harmonic Ritz vectors, but also has the property of recovering global optimality of the full GMRES method by augmenting error approximations. Numerical examples have shown that the new method is efficient on problems with well separated with small eigenvalues.

## References

[1] K. Baglama, D. Calvetti, G.H. Golub, and L. Reichel, *Adaptively preconditioned GMRES algorithms,* SIAM J. Sci. Comput., 20 (1998) 243-269.

[2] A. Baker, L. Jessup, and T. Manteuffel, *A technique for accelerating the convergence of restarted GMRES*, SIAM J. Matrix Anal. Appl., 26 (2005) 962-984.

[3] K. Burrage and J. Erhel, *On the performance of various adaptive preconditioned GMRES strategies,* Numer. Linear Algebra Appl., 5 (1998) 101-121.

[4] C. Le Calvez and B. Molina, *Implicitly restarted and deflated GMRES*, Numer. Algo., 21 (1999) 261-285.

[5] Z.-H. Cao, *A note on the convergence behavior of GMRES,* Appl. Numer. Math., 25 (1997) 13-20.

[6] B. Carpentieri, I.S. Duff, L. Giraud, *A class of spectral two-level preconditioners*, SIAM Journal on Scientific Computing, 25 (2003) 749-765.

[7] T. Davis, University of Florida sparse matrix collection, http://www.cise.ufl.edu/research/ sparse/matrices, 2002.

[8] A. Chapman and Y. Saad, *Deflated and augmented Krylov subspace techniques*, Numer. Linear Algebra Appl., 4 (1997) 43-66.

[9] D. Darnell, R. B. Morgan, and W. Wilcox, *Deflation of eigenvalues for iterative methods in lattice QCD*, Nucl. Phys. B (Proc. Suppl.), 129 (2004) 856-858.

[10] E. De Sturler, *Nested Krylov methods based on GCR*, J. Comput. Appl. Math., 67 (1996) 15-41.

[11] E. De Sturler, *Truncation strategy for optimal Krylov subspace methods*, SIAM J. Numer. Anal., 36 (1999) 864-889.

[12] I. S. Duff, R. G. Grimes, and J. G. Lewis, *Sparse matrix test problems*, ACM Trans. Math. Soft., 15 (1989) 1-14.

[13] M. Eiermann, O. G. Ernst, and O. Schneider, *Analysis of Acceleration Strategies for Restarted Minimal Residual Methods,* J. Comput. Appl. Math., 123 (2000) 261-292.

[14] S. C.Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear ,* SIAM J. Numer. Anal., 20 (1983) 345-357

[15] J. Erhel, K. Burrage, and B. Pohl, *Restarted GMRES preconditioned by deflation*, J. Comput. Appl. Math., 69 (1996) 303-318.

[16] J. Erhel and F. Guyomarc'h, *An Augmented Conjugate Gradient Method for solving consecutive symmetric positive definite systems*, SIAM. J. Matrix Anal. Appl., 21 (2000) 1279-1299.

[17] L. Giraud, S. Gratton, E. Martin, *Incremental spectral preconditioners for sequences of linear systems*, Applied Numerical Mathematics, In Press.

[18] S.A. Kharchenko, A. Yeramin, *Eigenvalue translation based preconditioners for the GMRES(k) method*, Numer. Linear Algebra Appl. 2, (1995) 51-77.

[19] R. B. Morgan, *A restarted GMRES method augmented with eigenvectors*, SIAM J. Matrix Anal. Appl., 16 (1995) 1154-1171.

[20] R. B. Morgan, *Implicitly restarted GMRES and Arnoldi methods for nonsymmetric systems of equations*, SIAM J. Matrix Anal. Appl., 21 (2000) 1112-1135.

[21] R. B. Morgan, *GMRES with deflated restarting*, SIAM J. Sci. Comput., 24 (2002) 20-37.

[22] National Institute of Standards and Technology, Mathematical and Computatinal Sciences Division, *Matrix Market*, Available at http://math.nist.gov/MatrixMarket/

[23] S. Rollin and W. Fichtner, *Improving the accurancy of GMRES with deflated restarting*, SIAM J. Sci. Comput., 30 (2007) 232-245.

[24] Y. Saad and M. H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986) 856-869.

[25] Y. Saad, Iterative Methods for Sparse Linear Systems, PWS Publishing Company: Boston, MA, 1996.

[26] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993) 461-469.

[27] Y. Saad, *Analysis of augmented Krylov subspace methods*, SIAM J. Matrix Anal. Appl., 18 (1997) 435-449.

[28] Y. Saad, *A deflated version of the conjugate gradient algorithm*, SIAM J. Sci. Comput., 21 (2000) 1909-1926.

[29] V. Simoncini, *On the convergence of restarted Krylov subspace methods*, SIAM J. Matrix Anal. Appl., 22 (2000) 430-452.

[30] V. Simoncini and D. B. Szyld, *On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods*, SIAM Review., 47 (2005) 247-272.

[31] V. Simoncini and D. B. Szyld. *Recent computational developments in Krylov subspace methods for linear systems*, Numer. Linear Algebra Appl., 14 (2007) 1-59.

[32] H. A. van der Vorst and C. Vuik, *The superlinear convergence behaviour of GMRES*, J. Comput. Appl. Math., 48 (1993) 327-341.

[33] H. A. van der Vorst and C. Vuik, *GMRESR: A family of nested GMRES methods*, Numer. Linear Algebra Appl., 1 (1994) 369-386.

[34] J. R. Wallis, R. P. Kendall and T. E. Little *Constraint residual acceleration of conjugate residual methods,* presented at the SPE 1985 Reservoir Simulation Symmposium held in Dallas, Texas, 1985.