Received: 26 January 2016 | Revised: 26 October 2017 | Accepted: 7 December 2017

DOI: 10.1002/nla.2148

RESEARCH ARTICLE

WILEY

A block GMRES method with deflated restarting for solving linear systems with multiple shifts and multiple right-hand sides

Dong-Lin Sun[1,2] | Ting-Zhu Huang[1] | Yan-Fei Jing[1] | Bruno Carpentieri[3]

[1]School of Mathematical Sciences/Institute of Computational Science, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China

[2]Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Nijenborgh 9 Groningen, PO Box 407, 9747 AG, The Netherlands

[3]Faculty of Computer Science, Free University of Bozen-Bolzano, piazza Domenicani 3 Bozen-Bolzano, 39100, Italy

Correspondence
Ting-Zhu Huang, School of Mathematical Sciences/Institute of Computational Science University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China.
Email: tingzhuhuang@126.com
Yan-Fei Jing, School of Mathematical Sciences/Institute of Computational Science University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China.
Email: yanfeijing@uestc.edu.cn

Funding information
National Natural Science Foundation of China (NSFC), Grant/Award Number: 61772003; University of Groningen Ubbo Emmius scholarship

Summary

The restarted block generalized minimum residual method (BGMRES) with deflated restarting (BGMRES-DR) was proposed by Morgan to dump the negative effect of small eigenvalues from the convergence of the BGMRES method. More recently, Wu et al. introduced the shifted BGMRES method (BGMRES-Sh) for solving the sequence of linear systems with multiple shifts and multiple right-hand sides. In this paper, a new shifted block Krylov subspace algorithm that combines the characteristics of both the BGMRES-DR and the BGMRES-Sh methods is proposed. Moreover, our method is enhanced with a seed selection strategy to handle the case of almost linear dependence of the right-hand sides. Numerical experiments illustrate the potential of the proposed method to solve efficiently the sequence of linear systems with multiple shifts and multiple right-hand sides, with and without preconditioner, also against other state-of-the-art solvers.

KEYWORDS

block Krylov subspace methods, deflated restarting, seed strategy, shifted systems

1 | INTRODUCTION

We consider solutions of linear systems with multiple right-hand sides, in which the coefficient matrices differ from each other by a scalar multiple of the identity, as follows:

$$(A - \sigma_i I)X_i = B, \quad \text{with } i = 1, 2, \ldots, L, \tag{1}$$

where $A - \sigma_i I \in \mathbb{C}^{n\times n}$ are square nonsingular matrices of large dimension n, $B = [b^{(1)}, b^{(2)}, \ldots, b^{(p)}]$ is a full rank matrix consisting of $p \ll n$, given right-hand sides $b^{(i)}$, and $\{\sigma_i\}_{i=1}^{L} \subset \mathbb{C}$ are called shifts. Many large-scale scientific

applications, such as lattice quantum chromodynamics,[1] Tikhonov–Phillips regularization,[2] Newton trust region methods,[3] and PageRank problems,[4] to name a few, require the solution of the sequence of linear systems with both multiple shifts and multiple right-hand sides given simultaneously. In this setting, matrix-free iterative methods such as Krylov subspace methods are of particular interest.[5–8]

Three main approaches can be distinguished for solving systems (1). The first obvious approach is to solve each of the p multishifted linear systems independently. In this case, shifted methods that take advantage of the shift-invariance property of the Krylov subspace may be of good choice.[9–17] Recently, Darnell et al.[18] have proposed an efficient method based on deflated restarting that corrects the computed solution for an extra right-hand side at moderate cost. The second approach is to use block Krylov subspace methods for solving the L linear systems with multiple right-hand sides in sequence. Block Krylov methods are computationally attractive because they use much larger search spaces, and additionally, they preserve the shift-invariance property of the Krylov subspace generated by the coefficient matrix.[19–22] Often, block Krylov methods are restarted after each cycle of, say, m iterations to limit the increasing memory and algorithmic costs for the iterative solution. The restarting procedure, however, may increase the number of iterations significantly, especially for matrices having small eigenvalues. It is known that the convergence of iterative solvers depends, to a large extent, on the distribution of the eigenvalues of the coefficient matrix.[23] In many cases, it is observed that removing the effect of the smallest eigenvalues can greatly improve convergence.[24–26] Morgan has extended the generalized minimum residual method (GMRES) with deflated restarting (shortly, GMRES-DR)[27] to a block variant (BGMRES-DR) that is suitable for solving linear systems with multiple right-hand sides.[28] Starting from the pioneering work by Morgan, many block deflated methods have been proposed in the past few years.[29–33]

Instead of applying Krylov subspace methods to linear systems with multiple shifts or block Krylov subspace methods to each single shifted linear system with multiple right-hand sides, it is more efficient to use a shifted block method for solving all of the shifted linear systems simultaneously. Recently, a modified shifted block GMRES (BGMRES-Sh) method for multiple shifts and multiple right-hand sides was proposed by Wu et al.[34] Inspired by the elegant methods mentioned above,[28,34] our main interest in this paper is to develop a new shifted BGMRES method with deflated restarting (hereafter named as BGMRES-DR-Sh) combined with the seed selection strategy[35] that is particularly effective for solving linear systems whose right-hand sides are nearly linearly dependent. Numerical experiments show the effectiveness of combining the virtues of the BGMRES-DR and the BGMRES-Sh methods.

Although Krylov subspace methods can solve the memory bottlenecks of sparse direct solvers, in practice, they may converge slowly, or sometimes, they even fail to converge especially for solving large problems. Thus, they need the assistance of efficient convergence acceleration techniques, or preconditioners, to reduce the overall solution time. For shifted linear systems, the choice of effective preconditioners is limited because conventional approaches often destroy the shift-invariance property of the Krylov subspace. As pointed out by Simoncini et al.,[36] the development of effective preconditioning strategies for this problem class remains an open research area. Shift-and-invert[37] and polynomial preconditioners[34] are two viable options. Recently, other preconditioning strategies have been proposed for solving shifted linear systems, see, for example, other works.[9,38] In this paper, we combine our BGMRES-DR-Sh method with the shift-and-invert preconditioner.

The remainder of this paper is organized as follows. A brief description of the BGMRES method with deflated restarting developed by Morgan[28] and of the BGMRES method for solving multishifted linear systems[34] is presented in Section 2. In Section 3, we describe a new shifted BGMRES method that deflates the small eigenvalues and is combined with the seed strategy. We also discuss how to use the right preconditioning technique that maintains the block shift-invariance property with our method. Numerical experiments are reported in Section 4. Finally, some conclusions and perspectives for future work are presented in Section 5.

Throughout this paper, the symbol $\|\cdot\|_q$ denotes the Euclidean norm when $q = 2$ and the Frobenius norm when $q = F$. The superscript H denotes the transpose conjugate of a vector or a matrix. $I_k \in \mathbb{C}^{k\times k}$ is the identity matrix of dimension k, and $0_{i\times j} \in \mathbb{C}^{i\times j}$ is the zero rectangular matrix with i rows and j columns. In addition, MATLAB notation is used; for example, $R(i,j)$ denotes the (i,j)th entry of a matrix R, and $R(1:m, 1:j)$ refers to the submatrix consisting of the first m rows and the first j columns of R.

2 | TWO VARIANTS OF BGMRES METHOD

In this section we briefly describe two existing variants of the BGMRES method that will be used in the development of the new BGMRES-DR-Sh method presented in this paper. Let X_0 be the initial block guess and $R_0 = B - AX_0$ be the

corresponding initial block residual. The block Krylov subspace generated by A from R_0 is defined as follows:

$$\mathbb{K}_m(A, R_0) = \text{span}\{R_0, AR_0, \ldots, A^{m-1}R_0\}.$$

Compute the reduced QR-decomposition of $R_0 = V_1R$, where $V_1 \in \mathbb{C}^{n\times p}$ is a matrix with orthonormal columns, and $R \in \mathbb{C}^{p\times p}$ is an upper triangular full-rank matrix. An orthonormal basis $\mathcal{V}_m = [V_1, V_2, \ldots, V_m] \in \mathbb{C}^{n\times n_m}$ of $\mathbb{K}_m(A, R_0)$ can be constructed by the standard block Arnoldi process described in Algorithm 1. Define $p_1 = p$, and p_{m+1} the column rank of the block orthonormal basis vector V_{m+1}. Then, $V_{m+1} \in \mathbb{C}^{n\times p_{m+1}}$, $W_m \in \mathbb{C}^{n\times p_m}$, and $H_{m+1,m} \in \mathbb{C}^{p_{m+1}\times p_m}$. It turns out that the product $m \times p$ is no longer the dimension of the search space but is an upper bound of the dimension of the search space. Let $n_m = \sum_{i=1}^{m} p_i$ be the sum of the column ranks of V_i ($i = 1, \ldots, L$). When no inexact breakdown has occurred, that is, $p_{m+1} = p_m = \ldots = p_1 = p$, the matrix $\mathcal{V}_m$ computed by the block Arnoldi process forms an orthonormal basis of the block Krylov subspace. As a consequence, the block Arnoldi relation[21] is written as follows:

$$A\mathcal{V}_m = \mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m. \tag{2}$$

We note that $\widetilde{\mathcal{H}}_m \in \mathbb{C}^{n_{m+1}\times n_m}$ has the following form:

$$\widetilde{\mathcal{H}}_m = \begin{bmatrix} \mathcal{H}_m \\ H_{m+1,m}E_m^H \end{bmatrix}, \tag{3}$$

where $\mathcal{H}_m \in \mathbb{C}^{n_m\times n_m}$ is supposed to be nonsingular with the $p \times p$ matrix $H_{i,j}$, and E_m is the matrix of the last p columns of I_{n_m}.

Algorithm 1 Block Arnoldi with blockwise modified Gram-Schmidt orthogonalization.

1: Choose a unitary matrix V_1 of size $n \times p$.
2: **for** j = 1, 2 ,..., m **do**
3: Compute $W_j = AV_j$
4: **for** i = 1, 2, ..., j **do**
5: $H_{i,j} = V_i^H W_j$
6: $W_j = W_j - V_i H_{i,j}$
7: **end for**
8: Compute the reduced QR-decomposition of W_j as $W_j = V_{j+1}H_{j+1,j}$
9: **end for**

The block minimum residual norm approach implemented in the BGMRES method[21] builds approximations to the solutions X at iteration m of the following form:

$$X_m = X_0 + \mathcal{V}_m Y_m,$$

where Y_m solves the least-squares problem, as follows:

$$Y_m = \underset{Y}{\text{argmin}}\|R_0 - A\mathcal{V}_m Y\|_F = \underset{Y}{\text{argmin}}\left\|ER - \widetilde{\mathcal{H}}_m Y\right\|_F. \tag{4}$$

In (4), we denote $R_0 = V_1R = \mathcal{V}_{m+1}ER$, and E is the $n_{m+1} \times p$ matrix whose upper $p \times p$ principal block is an identity matrix.

2.1 | BGMRES method with deflated restarting

In order to improve the convergence of restarted BGMRES, Morgan[28] proposed to recycle spectral information at each restart. In his BGMRES-DR method, the approximate eigenvectors corresponding to the small eigenvalues in magnitude are computed at the end of each cycle and are added to the next subspace. Let R_0 be the block residual at the end of the previous cycle. The subspace used by BGMRES-DR at the current cycle is as follows:

$$\text{span}\left(y_1, y_2, \ldots, y_k, R_0, AR_0, \ldots, A^{\frac{m-k}{p}-1}R_0\right),$$

where $y_1, y_2, \ldots, y_k$ are the harmonic Ritz vectors. The search space used by BGMRES-DR is also a block Krylov subspace, and this means that a block Arnoldi-type recurrence formula such as (2) exists, where $\widetilde{\mathcal{H}}_m$ is block or band upper-Hessenberg, except for a full leading $(k + p)$-by-$(k + p)$ portion. This approach has shown better performance than the standard BGMRES on many academic examples. We refer the reader to the work of Morgan[28] for further comments and computational details on the BGMRES-DR algorithm.

2.2 | Shifted BGMRES method

Recently, block Krylov subspace methods have been extended for solving linear systems of the form (1), see, the work of Wu et al.[34] We know that the block Krylov subspace is invariant under scalar shifts of the coefficient matrices,[31] that is,

$$\mathbb{K}_m(A, R_0) = \mathbb{K}_m(A - \sigma_i I, R_0),\ i = 1, \ldots, L. \tag{5}$$

From (2), we have the following:

$$(A - \sigma_i I)\mathcal{V}_m = \mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^i, \tag{6}$$

where

$$\widetilde{\mathcal{H}}_m^i = \widetilde{\mathcal{H}}_m - \sigma_i \begin{bmatrix} I_{n_m \times n_m} \\ 0_{p \times n_m} \end{bmatrix} = \begin{bmatrix} \mathcal{H}_m^i \\ H_{m+1,m} E_m^H \end{bmatrix}. \tag{7}$$

Let $(A - \sigma_1 I)X_1 = B$ be the base shifted block system and the other shifted block systems $(A - \sigma_i I)X_i = B, (i = 2, \ldots, k)$ be the additional shifted block systems as default. The approximate solutions of the base shifted block system can be formulated as $X_m^1 = X_0^1 + \mathcal{V}_m Y_m^1$, where X_0^1 is the initial approximation to the block solution with $R_0^1 = B - (A - \sigma_1 I)X_0^1$ being the initial block residual. Exploiting the seminal idea of the GMRES-Sh method,[12] Wu et al.[34] developed the BGMRES-Sh method for solving multishifted linear systems. The key idea of BGMRES-Sh is to apply the BGMRES algorithm to the base shifted block system and then generate the approximate solutions to the additional shifted block systems by imposing collinearity with the computed base block residual, that is, the block residuals are such that

$$R_m^i = R_m^1 W_m^i,\ i = 2, \ldots, L, \tag{8}$$

where $R_m^i, (i = 2, \ldots, L)$ are the residuals of the additional shifted block systems and W_m^i are some $p \times p$ nonsingular matrices. The approximate solutions of the additional shifted block systems can be expressed as $X_m^i = X_0^1 + \mathcal{V}_m Y_m^i, \quad (i = 2, \ldots, L)$. We set $R_{LS} = ER - \widetilde{\mathcal{H}}_m^1 Y_m^1$ as the least-squares residuals, where E is the $n_{m+1} \times p$ matrix whose upper $p \times p$ principal block is an identity matrix. Compute the reduced QR-decomposition of $R_{LS} = \widehat{V}\widehat{R}$. Then, Y_m^i and W_m^i can be simultaneously solved by solving the following modified linear systems:

$$\begin{bmatrix} \widetilde{H}_m^i & \widehat{V} \end{bmatrix} \begin{bmatrix} Y_m^i \\ \widehat{W_m^i} \end{bmatrix} = ERW_0^i,\ i = 2, 3, \ldots, L, \tag{9}$$

where $\widehat{W}_m^i = \widehat{R}W_m^i$. The pseudocode for the BGMRES-Sh method is shown as in Algorithm 2, which is not explicitly displayed in the work of Wu et al.[34]

Algorithm 2 BGMRES-Sh method.

1: *Start*. Let the p linearly independent right-hand sides be $B = [b^{(1)}, b^{(2)}, \ldots, b^{(p)}]$. Choose n_m: the maximum size of the underlying block approximation Krylov subspace in each cycle, ϵ: the tolerance, *maxiter*: the maximum number of iterations, $\{\sigma_i\}_{i=1}^L \subset \mathbb{C}$: shifts.
2: Choose initial guesses $X_0^i \in \mathbb{C}^{n \times p}$ and set $W_0^i = I (i = 1, \ldots, L)$.
3: Compute the initial shifted block residuals $R_0^i \in \mathbb{C}^{n \times p}$ as $R_0^i = B - (A - \sigma_i I)X_0^i (i = 1, \ldots, L)$.
4: Compute the reduced QR-decomposition of $R_0^1 = V_1 R$.
5: *Solve the base shifted block linear system*. Apply standard block GMRES(m). It generates $\mathcal{V}_{m+1}$, $\widetilde{\mathcal{H}}_m^1$ and forms the new approximate solutions $X_m^1 = X_0^1 + \mathcal{V}_m Y_m^1$ where $Y_m^1 = \underset{Y}{\operatorname{argmin}} \left\| ER - \widetilde{\mathcal{H}}_m^1 Y \right\|_F$. Compute the residual vectors $R_m^1 = R_0^1 - \mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^1 Y_m^1$. Check convergence, and proceed if not satisfied.
6: *Solve the additional shifted block linear systems*. Solve $\begin{bmatrix} \widetilde{\mathcal{H}}_m^i & \widehat{V} \end{bmatrix} \begin{bmatrix} Y_m^i \\ \widehat{W_m^i} \end{bmatrix} = ERW_0^i, i = 2, 3, \ldots, L$, for Y_m^i and $\widehat{W_m^i}$; Then form the approximate solutions $X_m^i = X_0^i + \mathcal{V}_m Y_m^i$ and compute the residual vectors R_m^i as $R_m^i = R_m^1 \widehat{R}^{-1}\widehat{W_m^i}$.
7: Check convergence, and restart if convergence is not achieved, i.e., go to 4.

3 | SHIFTED BGMRES METHOD WITH DEFLATED RESTARTING

In this section, we present the new shifted BGMRES method that allows deflated restarting and solves the multiple shifts simultaneously. The algorithm combines the favorable convergence property of BGMRES-Sh[34] and the augmentation strategy of BGMRES-DR.[28] First, we consider a base shifted block system in a cycle. Note that the first cycle of the BGMRES-DR method applied to the base shifted block system is the standard BGMRES method.

3.1 | Analysis of a cycle

Similar to the variants of GMRES-DR-type methods,[29,39] we need to take two main problems into consideration: what is the harmonic Ritz information that is recovered at restart, and how to use these harmonic Ritz vectors recycled at restart to maintain the shifted block Arnoldi-like relation at low computational cost. In the next section, we will answer both questions.

3.1.1 | Harmonic Ritz vectors

Definition 1. (Harmonic Ritz pair [40])
Consider a subspace $\mathcal{U}$ of $\mathbb{C}^n$. Given a matrix $B \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}$, and $y \in \mathcal{U}$, (λ, y) is called a harmonic Ritz pair of B with respect to $\mathcal{U}$ if and only if

$$By - \lambda y \perp B\mathcal{U} \tag{10}$$

or equivalently, using the canonical scalar product, if and only if

$$\forall \omega \in \text{Range}(B\mathcal{U}) \qquad \omega^H(By - \lambda y) = 0. \tag{11}$$

The vector y is called the harmonic Ritz vector associated with the harmonic Ritz value λ.
The following lemma presents the harmonic Ritz formulation used in the BGMRES-DR for shifted linear systems.

Lemma 1. *Let $\mathcal{U} = span\{\mathcal{V}_m\}$, where $\mathcal{V}_m$ is an orthonormal matrix built by BGMRES at the end of a cycle. The harmonic Ritz pairs $(\widetilde{\theta}_i, \widetilde{g}_i)$ associated with $\mathcal{V}_m$ satisfy the following property:*

$$\left(\widetilde{\mathcal{H}}_m^1\right)^H \left(\widetilde{\mathcal{H}}_m^1 \widetilde{g}_i - \widetilde{\theta}_i \begin{bmatrix} \widetilde{g}_i \\ 0_p \end{bmatrix}\right) = 0, \tag{12}$$

where $\mathcal{V}_m \widetilde{g}_i$ are the harmonic Ritz vectors associated with the corresponding harmonic Ritz values $\widetilde{\theta}_i$.

Proof. The proposed property, which extends the GMRES-DR to the block shift version, can be written as a Petrov–Galerkin condition, as follows:

$$\begin{aligned} ((A - \sigma_1 I)\mathcal{V}_m)^H \left((A - \sigma_1 I)\mathcal{V}_m \widetilde{g}_i - \widetilde{\theta}_i \mathcal{V}_m \widetilde{g}_i\right) = 0, \\ \left(\mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^1\right)^H \left(\mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^1 \widetilde{g}_i - \widetilde{\theta}_i \mathcal{V}_m \widetilde{g}_i\right) = 0. \end{aligned} \tag{13}$$

Because $\mathcal{V}_{m+1}$ has orthonormal columns, (13) can be written as follows:

$$\begin{aligned} \left(\widetilde{\mathcal{H}}_m^1\right)^H (\mathcal{V}_{m+1})^H \mathcal{V}_{m+1} \widetilde{\mathcal{H}}_m^1 \widetilde{g}_i - \widetilde{\theta}_i \left(\widetilde{\mathcal{H}}_m^1\right)^H (\mathcal{V}_{m+1})^H \mathcal{V}_m \widetilde{g}_i = 0, \\ \left(\widetilde{\mathcal{H}}_m^1\right)^H \left(\widetilde{\mathcal{H}}_m^1 \widetilde{g}_i - \widetilde{\theta}_i \begin{bmatrix} \widetilde{g}_i \\ 0_p \end{bmatrix}\right) = 0. \end{aligned}$$

□

Note that by using (7), the above property can be rewritten as

$$\left(\left(\mathcal{H}_m^1\right)^H \mathcal{H}_m^1 + E_m H_{m+1,m}^H H_{m+1,m} E_m^H\right) \widetilde{g}_i = \widetilde{\theta}_i \left(\mathcal{H}_m^1\right)^H \widetilde{g}_i. \tag{14}$$

Consequently, the k targeted harmonic Ritz pairs can be computed during the cycle of the shifted BGMRES method. Similar to the block variants,[29] the harmonic residual vectors $\widetilde{\mathcal{H}}_m^1 \widetilde{g}_i - \widetilde{\theta}_i \begin{bmatrix} \widetilde{g}_i \\ 0_p \end{bmatrix}$ are all contained in the subspace spanned by the least-squares residuals $R_{LS} = ER - \widetilde{\mathcal{H}}_m^1 Y_m^1$, that is, $\exists \alpha_i \in \mathbb{C}^p$ such that

$$\widetilde{\mathcal{H}}_m^1 \widetilde{g}_i - \widetilde{\theta}_i \begin{bmatrix} \widetilde{g}_i \\ 0_p \end{bmatrix} = R_{LS}\alpha_i. \tag{15}$$

3.1.2 | Shifted block Arnoldi-like relation

In this section, we will show that the shifted block Arnoldi-like relation still holds at low computational cost when the deflated restarting procedure is used. The k targeted harmonic Ritz vectors are given by $\mathcal{V}_m\widetilde{G}_k$, where $\widetilde{G}_k = [\widetilde{g}_1, \ldots, \widetilde{g}_k] \in \mathbb{C}^{n_m \times k}$. First, add zero rows of size p to $\widetilde{G}_k$ and then append the quasiresidual R_{LS} to $\widetilde{G}_k$. In the two cases, the new matrix is denoted as $\widetilde{G}_{k+1} = \begin{bmatrix} \begin{bmatrix} \widetilde{G}_k \\ 0_{p\times k} \end{bmatrix}, R_{LS} \end{bmatrix}$, and its dimension is $(n_{m+1}) \times (k+p)$. The reduced QR-decomposition of $\widetilde{G}_{k+1} = Q_{k+1}R_{k+1}$ can be written as follows:

$$\begin{bmatrix} \begin{bmatrix} \widetilde{G}_k \\ 0_{p\times k} \end{bmatrix}, R_{LS} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} Q_k \\ 0_{p\times k} \end{bmatrix}, Q_2 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} R_k \\ 0_{p\times k} \end{bmatrix}, R_2 \end{bmatrix}, \tag{16}$$

with

$$\widetilde{G}_k = Q_k R_k. \tag{17}$$

Proposition 1. *At each restart of shifted BGMRES method with deflated restarting, the shifted block Arnoldi-like relation, as follows:*

$$(A - \sigma_1 I)\mathcal{V}_k^{new} = \mathcal{V}_{k+1}^{new}\left(\widetilde{\mathcal{H}}_k^1\right)^{new}, \tag{18}$$

holds with $\mathcal{V}_{k+1}^{new} = \mathcal{V}_{m+1}Q_{k+1}$, $(\widetilde{\mathcal{H}}_k^1))^{new} = Q_{k+1}^H \widetilde{\mathcal{H}}_m^1 Q_k$.

Proof. Multiplying both sides of (15) with $\mathcal{V}_{m+1}$ gives the following:

$$\mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^1\widetilde{g}_i - \widetilde{\theta}_i\mathcal{V}_{m+1}\begin{bmatrix} \widetilde{g}_i \\ 0_p \end{bmatrix} = \mathcal{V}_{m+1}R_{LS}\alpha_i.$$

Using (6) and the above equation, we have the following:

$$(A - \sigma_1 I)\mathcal{V}_m\widetilde{G}_k = \mathcal{V}_{m+1}\widetilde{G}_{k+1}\begin{pmatrix} diag(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k) \\ \alpha_1, \ldots, \alpha_k \end{pmatrix}.$$

From (16) and (17), we deduce that

$$(A - \sigma_1 I)\mathcal{V}_m Q_k = \mathcal{V}_{m+1}Q_{k+1}R_{k+1}\begin{pmatrix} diag(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k) \\ \alpha_1, \ldots, \alpha_k \end{pmatrix} R_k^{-1}. \tag{19}$$

Using (6) and $Q_{k+1}^H Q_{k+1} = I$, we have the following:

$$R_{k+1}\begin{pmatrix} diag(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k) \\ \alpha_1, \ldots, \alpha_k \end{pmatrix} R_k^{-1} = Q_{k+1}^H \widetilde{\mathcal{H}}_m^1 Q_k. \tag{20}$$

If we denote

$$\mathcal{V}_{k+1}^{new} = \mathcal{V}_{m+1}Q_{k+1},\ (\widetilde{\mathcal{H}}_k^1)^{new} = Q_{k+1}^H\widetilde{\mathcal{H}}_m^1 Q_k,$$

and substitute (20) into (19), the shifted block Arnoldi-like relation can be obtained. □

Setting $\mathcal{V}_{k+1} = \mathcal{V}_{k+1}^{new}$, $\widetilde{\mathcal{H}}_k^1 = (\widetilde{\mathcal{H}}_k^1)^{new}$, BGMRES-DR-Sh then carries out $\frac{n_m - k}{p}$ block Arnoldi steps with starting vector V_{k+1} to maintain orthogonality against the $\mathcal{V}_{k+1}$. At the end of a cycle, the shifted block Arnoldi-like relation similar to (6) can be obtained. Then, we seek an approximate solution $X_m^1 = X_0^1 + \mathcal{V}_m Y_m^1$ over the subspace $X_0^1 + \text{Range}(\mathcal{V}_m)$, where Y_m^1 is the solution of the following least-squares problem:

$$Y_m^1 = \underset{Y}{\text{argmin}}\left\|R_0^1 - (A - \sigma_1 I)\mathcal{V}_m Y\right\|_F = \underset{Y}{\text{argmin}}\left\|ER - \widetilde{\mathcal{H}}_m^1 Y\right\|_F, \tag{21}$$

with $R_0^1 = \mathcal{V}_{m+1}ER$. Thus, the corresponding base shifted block residual is written as follows:

$$\begin{aligned} R_m^1 &= R_0^1 - (A - \sigma_1 I)\mathcal{V}_m Y_m^1 = \mathcal{V}_{m+1}(ER - \widetilde{\mathcal{H}}_m^1 Y_m^1), \\ &= \mathcal{V}_{m+1}R_{LS}; \end{aligned} \tag{22}$$

here, we set $R_{LS} = ER - \widetilde{\mathcal{H}}_m^1 Y_m^1$.

The additional shifted block residual vectors associated with approximate solutions of the following form:

$$X_m^i = X_0^i + \mathcal{V}_m Y_m^i, \tag{23}$$

is given by as follows:

$$R_m^i = B - (A - \sigma_i I)X_m^i = R_0^i - \mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^i Y_m^i,\ i = 2, 3, \ \dots, L. \tag{24}$$

In order to get the additional shifted block residuals and take advantage of the block Krylov subspace shift-invariance property, we force the block residuals as follows:

$$R_m^i = R_m^1 W_m^i,\ i = 2, \ \dots, L, \tag{25}$$

where W_m^i are some $p \times p$ nonsingular matrices. We assume that the initial shifted block residuals are collinear, $R_0^i = R_0^1 W_0^i$, $i = 2, \ \dots, L$ (note that this relation is satisfied by taking $X_0^i = 0$).

On the other hand, (25) can be rewritten as follows:

$$R_0^i - \mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^i Y_m^i = \mathcal{V}_{m+1} R_{LS} W_m^i,\ i = 2, \ \dots, L. \tag{26}$$

From (24) and (26), the Y_m^i and W_m^i can be simultaneously solved by solving the following linear system:

$$\left[\ \widetilde{\mathcal{H}}_m^i \ \ R_{LS}\ \right] \begin{bmatrix} Y_m^i \\ W_m^i \end{bmatrix} = ERW_0^i,\ i = 2, 3, \ \dots, L. \tag{27}$$

Finally, the additional shifted block systems can be expressed as in (23), where $Y_m^i, \quad (i = 2, \ \dots, L)$ is the solution of (27).

One practical difficulty with the implementation and use of block Krylov methods for the simultaneous solution of multishifted linear systems of the form (1) is the possible linear dependence of the various residuals, which may slow down or even prevent the convergence. Various methods have been proposed to overcome this problem.[29,41–43] The issue is not trivial, both theoretically and computationally, and we will analyze it in a separate study. Another difficulty is that the coefficient matrix in (27) will generally become more and more ill-conditioned when the base shifted block system is close to convergence, as the base shifted block residual $\|R_m^1\|_F = \|\mathcal{V}_{m+1}R_{LS}\|_F = \|R_{LS}\|_F$ tends to zero. In order to deal with this problem, we use a modified process to solve the additional shifted block systems, similar to the work of Wu et al.[34] We set $R_{LS} = \widehat{V}\widehat{R}$ as the reduced QR-decomposition, where $\widehat{V}$ is an $n \times p$ orthonormal matrix and $\widehat{R}$ is a $p \times p$ upper-triangular matrix. Then, (27) can be rewritten as follows:

$$\left[\ \widetilde{\mathcal{H}}_m^i \ \ \widehat{V}\ \right] \begin{bmatrix} Y_m^i \\ \widehat{W_m^i} \end{bmatrix} = ERW_0^i,\ i = 2, 3, \ \dots, L, \tag{28}$$

where $\widehat{W}_m^i = \widehat{R}W_m^i$. Thus, we can use the quantity Y_m^i, $i = 2, \ \dots, L$ to form the approximate solutions X_m^i, $i = 2, \ \dots, L$ of the additional shifted block systems in (23). Observe from (22) and (25) that computing the Frobenius norms of the additional shifted residuals only needs to evaluate $\widehat{W}_m^i$ as follows:

$$\begin{aligned} \left\|R_m^i\right\|_F &= \left\|R_m^1 W_m^i\right\|_F = \left\|\mathcal{V}_{m+1}R_{LS}W_m^i\right\|_F \\ &= \left\|\mathcal{V}_{m+1}\widehat{V}\widehat{R}(\widehat{R})^{-1}\widehat{W}_m^i\right\|_F \\ &= \left\|\widehat{W}_m^i\right\|_F,\ i = 2, \ \dots, L, \end{aligned}$$

and therefore, it is cheaper than using (25). The pseudocode for the BGMRES-DR-Sh method is shown as in Algorithm 3 and Algorithm 4.

Algorithm 3 BGMRES-DR-Sh method.

1: *Start*. Let the p linearly independent right-hand sides be $B = [b^{(1)}, b^{(2)}, \dots, b^{(p)}]$. Choose n_m: the maximum size of the underlying block approximation Krylov subspace in each cycle, k: the desired number of approximate eigenvectors, ϵ: the tolerance, *maxiter*: the maximum number of iterations, $\{\sigma_i\}_{i=1}^L \subset \mathbb{C}$: shifts.
2: Choose an initial guess $X_0^i \in \mathbb{C}^{n\times p}$ and set $W_0^i = I$.
3: Compute the initial block residual $R_0^i \in \mathbb{C}^{n\times p}$ as $R_0^i = B - (A - \sigma_i I)X_0^i$.
4: *Solve the base shifted block linear system and computation of* $\widehat{V}$: see Algorithm 4.
5: *Solve the additional shifted block linear systems*. Solve $\left[\ \widetilde{\mathcal{H}}_m^i \ \ \widehat{V}\ \right] \begin{bmatrix} Y_m^i \\ \widehat{W_m^i} \end{bmatrix} = ERW_0^i,\ i = 2, 3, \dots, L$, for Y_m^i and $\widehat{W_m^i}$; Then form the approximate solutions $X_m^i = X_0^i + \mathcal{V}_m Y_m^i$, and compute the residual vectors R_m^i as $R_m^i = R_m^1 \hat{R}^{-1}\widehat{W_m^i}$.
6: Check convergence, and restart if convergence is not achieved, i.e., go to 4.

Algorithm 4 BGMRES-DR-Sh method: computation of $\hat{V}$.

1: Compute the reduced QR-decomposition of $R_0^1 = V_1 R$.
2: Find approximate solutions for the first cycle. Apply standard Block GMRES. It generates $\mathcal{V}_{m+1}$, $\widetilde{\mathcal{H}}_m^1$ and the new approximate solution $X_m^1 = X_0^1 + \mathcal{V}_m Y_m^1$ where $Y_m^1 = \underset{Y}{\operatorname{argmin}} \left\| ER - \widetilde{\mathcal{H}}_m^1 Y \right\|_F$. Compute the reduced QR-decomposition of $Z_{m+1} = \hat{V}\hat{R}$. Compute $R_m^1 = R_0^1 - \mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^1 Y_m^1$. Check convergence, and proceed if not satisfied.
3: Begin restart. Let $X_0^1 = X_m^1$ and $R_0^1 = R_m^1$. Compute the k smallest (or others, if desired) eigenpairs of $\widetilde{\theta}_i$, $\widetilde{g}_i$ of the matrix $(\mathcal{H}_m^1)^H \mathcal{H}_m^1 + E_m H_{m+1,m}^H H_{m+1,m} E_m^H$. Then store the $\widetilde{g}_i$ into the matrix $\tilde{G}_k$. For real matrices, it is necessary to separate $\widetilde{g}_i$ into real and imaginary parts if complex, in order to form an matrix $\tilde{G}_k$.
4: Append a $p \times k$ zero matrix at the bottom and the column $ER - \widetilde{\mathcal{H}}_m^1 Y_m^1$. In the two cases, the new matrix is denoted as $\tilde{G}_{k+1}$ and its dimension is $n_{m+1} \times (k+p)$.
5: Perform the reduced QR-decomposition of $\tilde{G}_{k+1}$ as $\widetilde{G}_{k+1} = Q_{k+1} R_{k+1}$. Define $Q_k \in \mathbb{C}^{n_m \times k}$ as $Q_k = Q_{k+1}(1 : n_m, 1 : k)$.
6: Set $\mathcal{V}_{k+1}^{new} = \mathcal{V}_{m+1} Q_{k+1}$, $\widetilde{\mathcal{H}}_k^{new} = Q_{k+1}^H \widetilde{\mathcal{H}}_m^1 Q_k$. And then set $\mathcal{V}_{k+1} = \mathcal{V}_{k+1}^{new}$, $\widetilde{\mathcal{H}}_k^1 = \widetilde{\mathcal{H}}_k^{new}$.
7: Orthogonalize $V_{k+1}, \ldots, V_{k+p}$ against the earlier columns of the new $\mathcal{V}_{k+1}$.
8: Block Arnoldi iteration: Form the rest of $\mathcal{V}_{m+1}$ and $\widetilde{\mathcal{H}}_m^1$ applying the block Arnoldi iteration starting with V_{k+1}, go to 2.

We do not present the computational costs of BGMRES-DR-Sh because it is similar to the work of Giraud et al.[39] Compared with BGMRES-Sh, the main computational difference is in the calculation of $\mathcal{V}_m$, which is the same as FGMRES versus FGMRES-DR mentioned in the work of Giraud et al.[39]

3.2 | Seed selection strategy

As shown in the theoretical error bound analysis reported in other works[35,44] for the approximate solution computed from the projection process, the seed strategy can enable us to reduce the norms of the additional residuals especially when the right-hand sides are almost linearly dependent. With regard to the multiple-shift and multiple right-hand side linear systems, it is also crucial to select a suitable seed shifted block system. A good seed shifted block system should give a good initial guess for the nonseed systems. In order to get more information of each shifted block systems, the seed block system can be chosen as a linear combination of the current residuals, such as in the work of Smith.[45] However, as pointed out by Simoncini et al.,[35] this will cause extra work to solve an artificial system. In this work, we use the seed selection strategy similar to other works[35,46] to improve the performance of the proposed BGMRES-DR-Sh method for solving the multishifted systems. During each cycle, we choose the seed block system that has maximum residual norm among the nonconverged shifts. That is, we choose the seed block system such that $\|R_m^1(:, seed)\|_2 \geq \|R_m^i(:, j)\|_2, j = 1, \ldots, p$, that is,

$$\left\| R_m^1(:, seed) \right\|_2 = \max_{j=1,\ldots,p} \left\| R_m^i(:, j) \right\|_2, \; i = 1, \ldots, s. \tag{29}$$

Here, s denotes the number of nonconverged shifted systems. Once the seed system is selected, we reorder the columns of the shifted block residuals R_m^i and set $R_0^1(new) = R_m^1(seed)$, $R_0^i(new) = R_m^i$ obtained from the previous cycle, solving the seed block system by the BGMRES-DR-Sh method. Then, we force the block residuals similar to (25) as follows:

$$R_m^i(new) = R_m^1(new) W_m^i(new), \; i = 2, \ldots, s, \tag{30}$$

where the initial W can be computed with $W_0^i(new) = (R_0^1(new))^{-1} R_0^i(new)$. A Galerkin projection of the block residual onto the shifted block Krylov subspace generated by the seed block system is used to obtain approximate solutions for the additional block systems. We remove the seed shift system from a series of shift systems if the seed block system converges to the required accuracy. The whole process is repeated until all the shifted block systems converge.

The main difference between a seed shifted BGMRES with deflated restarting version and Algorithm 3 is that we solve the shifted block system that has a maximum residual instead of the base shifted block system in Step 4.

3.3 | Combining the BGMRES-DR-Sh method with preconditioning

Although the method seen in the previous subsection is theoretically well defined, in practice, it is likely to suffer from slow convergence or even stagnation. Clearly, one way to accelerate the convergence is to use preconditioning technology by transforming the original sequence of linear systems into an equivalent system with more favorable spectral properties. For the solution of multishifted linear systems (1), however, the block shift-invariance property (5) of the block Krylov subspace may not be ensured if general preconditioning is used. Special preconditioners that preserve property (5) are shift-and-invert,[47] multishift,[38] and polynomial preconditioners.[34] We refer the reader to other works[9,47–49] for some discussions on preconditioning techniques for solving shifted linear systems.

In this paper, we solve the system of equations using the right preconditioner of the form $A_\tau = A - \tau I$, as follows[38]:

$$(A - \sigma_i I)(A - \tau I)^{-1} \widetilde{X}_i = B, \quad i = 1, \ldots, L, \tag{31}$$

where $X_i = (A - \tau I)^{-1}\widetilde{X}_i$. Because

$$(A - \sigma_i I)(A - \tau I)^{-1} = I + (\tau - \sigma_i)(A - \tau I)^{-1}, \tag{32}$$

the shift-invariance property $\mathbb{K}_m((A - \sigma_i I)(A - \tau I)^{-1}, B) = \mathbb{K}_m((A - \tau I)^{-1}, B)$ is satisfied. Note that $\mathbb{K}_m((A - \tau I)^{-1}, B)$ is independent of σ_i. This means that the construction of the preconditioned Krylov subspace does not require matrix-vector products (mvps) with $(A - \sigma_i I)$. Although the application of the preconditioner $A - \tau I$ is not cheap in general, the basis of the preconditioned Krylov subspace needs to be computed only once. Owing to the shift-invariance property of the corresponding Krylov subspace, the solution of the other shifted systems can be carried out at a relatively low cost. We run the block Arnoldi algorithm on the matrix $(A - \tau I)^{-1}$ starting from the vector B to get a basis for the Krylov subspace $\mathbb{K}_m((A - \tau I)^{-1}, B)$. As a consequence, we get the following relations:

$$Z_m = \mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m, \tag{33}$$

$$(A - \tau I)Z_m = \mathcal{V}_m, \tag{34}$$

where $\mathcal{V}_m^H \mathcal{V}_m = I$. Multiplying both sides of (33) by $(\tau - \sigma_i)$ and adding it to the (34), we obtain an Arnoldi-like relation, as follows:

$$(A - \sigma_i I)Z_m = \mathcal{V}_{m+1}\widetilde{\mathcal{H}}_m^i, \tag{35}$$

where

$$\widetilde{\mathcal{H}}_m^i = \begin{bmatrix} I_{n_m \times n_m} \\ 0_{p \times n_m} \end{bmatrix} + (\tau - \sigma_i)\widetilde{\mathcal{H}}_m. \tag{36}$$

Here, the columns of matrix $\mathcal{V}_{m+1}$ span the Krylov subspace $\mathbb{K}_m((A - \tau I)^{-1}, B)$. Similar to the analysis presented in Section 3.1, first, the preconditioned BGMRES-DR algorithm is run on the base shifted block system, and then, the approximate solutions to the additional shifted block systems are generated by imposing collinearity with the computed base block residual. The preconditioned BGMRES-DR-Sh method (denoted as PBGMRES-DR-Sh) seeks approximate solutions of the form $X_m^i = X_0^i + Z_m Y_m^i$ $(i = 1, 2, \ldots, L)$. The spectral analysis of the preconditioned matrix $(A - \sigma_i I)(A - \tau I)^{-1}$ proposed by Saibaba et al.[38] has shown that the preconditioner A_τ is well suited only for values of σ_i near τ. Note that the left preconditioner of the form $A_\tau = A - \tau I$ can also be combined with the BGMRES-DR-Sh method because the shift-invariance property is satisfied. Although not considered in this paper, the right preconditioner can be easily extended to flexible and inexact Krylov methods.[37,38]

3.4 | Detecting approximate linear dependence of the residuals

Similar to other block Krylov subspace algorithms, the BGMRES-DR-Sh method may suffer from potentially very slow convergence due to the presence of (approximately) linearly dependent residuals. One natural strategy to alleviate this problem and try to restore fast convergence is to discard redundant information from the Krylov subspace. This strategy is known as right-hand side deflation. In the initial deflation approach, the SVD of the block right-hand sides matrix B (or of the initial block residual matrix R_0 if the initial guess is not zero) is computed, and the smallest singular values are filtered to remove the linearly dependent directions from the Krylov basis.[41,42] The BGMRES-DR-Sh method cannot be combined with the initial deflation strategy, as the shifted block Arnoldi-like relation (18) is not maintained after filtering the linearly dependent columns in the initial block residual at the first cycle. Another approach, known as Arnoldi deflation, monitors the possible column rank deficiency of the matrix W_j at Step 8 of the block Arnoldi procedure (Algorithm 1).[42] Recently, Robbé et al.[43] introduced the inexact breakdown technique similar to the Arnoldi deflation to reduce the block size of the BGMRES method. Identifying the best strategy to detect a possible approximate linear dependence of the residuals in the BGMRES-DR-Sh method is not a trivial task. We will look at this problem in a separate study.

4 | NUMERICAL EXPERIMENTS

In this section, we investigate the effectiveness of the proposed BGMRES-DR-Sh method for solving multishifted linear systems (1) arising in different fields. For all of the experiments, the right-hand sides $B = [b^{(1)}, \ldots, b^{(p)}] \in \mathbb{C}^{n \times p}$ are p linearly independent columns randomly generated with normal distribution. The initial block guess is equal to $0 \in \mathbb{C}^{n \times p}$. The k smallest harmonic Ritz values correspond to the harmonic vectors that are recycled in BGMRES-DR-Sh. These are computed using the MATLAB function *eigs* that calls ARPACK.[50] The following stopping criterion is used for the iterative solution:

$$\frac{\left\|R_m^i(:, j)\right\|_2}{\|B(:, j)\|_2} \leq \varepsilon, \; i = 1, \ldots, L.$$

No seed selection strategy is used except for the experiments presented in Section 4.1.3. The shift-and-invert preconditioner is only implemented in Section 4.2.1 to ensure the convergence within a reasonable amount of iterations. All the numerical experiments are performed using Windows 7 and MATLAB (version R2014a) running on a PC with an Inter(R) Core(TM) i5-4590 CPU @3.30 GHz and 8 GB of memory.

4.1 | Performance analysis of the BGMRES-DR-Sh method

First, we compare the performance of the BGMRES-DR-Sh method introduced in Section 3 with the BGMRES-Sh method by Wu et al.[34] in terms of number of mvps, maximum true relative residuals, and CPU solution time in seconds (these quantities are referred to *mvps*, *trr*, and *cpu*, respectively), to study the effect on convergence of the deflated restarting procedure. In our experiments, the number of right-hand sides is $p = 6$, the maximum dimension of the search space is $m = 90$, the number of deflated harmonic Ritz vectors is $k = 10$, the tolerance is $\varepsilon = 10^{-6}$, and $maxiter = 5{,}000$. We choose the shifts $\sigma = 0, -0.4, -2$, which are the same as in the work of Darnell et al.[18] For the sake of comparison with related works in the literature, we choose two different sets of matrices. The first four representative problems[28] have dimension 1,000 and are bidiagonal with superdiagonal entries all equal to one: matrix 1 has diagonal entries 0.1, 1, 2, 3, … , 999; matrix 2 has diagonal entries 1, 2, 3, … , 1,000; and matrix 3 has diagonal entries 11, 12, 13, … , 1,010, whereas matrix 4 has diagonal entries 10.1, 10.2, 10.3, … , 19.9, 20, 21, … , 919, 920. The second set of matrices used in Examples 5, 6, 7, 8 are extracted from the Sparse Matrix Collection available at the University of Florida,[51] described in Table 1.

As shown in Table 2, the BGMRES-DR-Sh method outperforms the BGMRES-Sh method in all cases in terms of both *Mvps* and *Cpu* time, converging to the targeted accuracy where the BGMRES-Sh method cannot converge. The symbol "-" is used to indicate no convergence, and three values in the *Trr* column means that maximum true relative residuals correspond to three different shifts. The corresponding convergence histories are shown in Figure 1, which report on the maximum relative residual norms against the number of matrix-vector products. We clearly see from Figure 1 that the BGMRES-DR-Sh method outperforms the BGMRES-Sh method especially on Examples 1, 2, 5, and 6, due to the presence of small eigenvalues on the spectra of these problems. This makes the BGMRES-DR-Sh method more effective because

TABLE 1 Characteristics of examples 5, 6, 7, 8

Example	Name	Size	Nonzeros	Problem kind
5	bfwa398	398	3,678	electromagnetics
6	orsreg1	2,205	14,133	computational fluid dynamics
7	pde2961	2,961	14,585	two-dimensional/three-dimensional problem
8	sherman4	1,104	3,786	computational fluid dynamics

TABLE 2 Numerical behavior of the BGMRES-Sh and BGMRES-DR-Sh methods with $\varepsilon = 10^{-6}$

Example	Method	*Mvps*	*Trr*(max)	*Cpu*
1	BGMRES-Sh	–	–	–
	BGMRES-DR-Sh	**648**	9.9933e−07 9.9513e−07 9.7075e−07	**0.7241**
2	BGMRES-Sh	2,158	9.9690e−07 9.9140e−07 9.9242e−07	1.8655
	BGMRES-DR-Sh	**520**	9.2339e−07 9.8931e−07 9.4725e−07	**0.5233**
3	BGMRES-Sh	403	8.9853e−07 9.9129e−07 9.8459e−07	0.4317
	BGMRES-DR-Sh	**332**	9.6970e−07 9.6649e−07 9.3198e−07	**0.3511**
4	BGMRES-Sh	456	9.6458e−07 9.9220e−07 8.8725e−07	0.4881
	BGMRES-DR-Sh	**443**	9.7678e−07 9.9071e−07 9.9901e−07	**0.4775**
5	BGMRES-Sh	–	–	–
	BGMRES-DR-Sh	**595**	8.3238e−07 5.7918e−07 9.2956e−07	**0.6456**
6	BGMRES-Sh	–	–	–
	BGMRES-DR-Sh	**4,655**	9.8521e−07 9.9245e−07 9.9870e−07	**6.5856**
7	BGMRES-Sh	1,793	9.9434e−07 9.9038e−07 6.7145e−07	1.7885
	BGMRES-DR-Sh	**1,204**	9.6114e−07 9.3280e−07 6.3293e−07	**1.3051**
8	BGMRES-Sh	1,975	9.8136e−07 9.4067e−07 9.9710e−07	1.3432
	BGMRES-DR-Sh	**424**	9.4382e−07 9.3570e−07 9.2801e−07	**0.4553**

Note. BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting; *Mvps* = matrix-vector products.

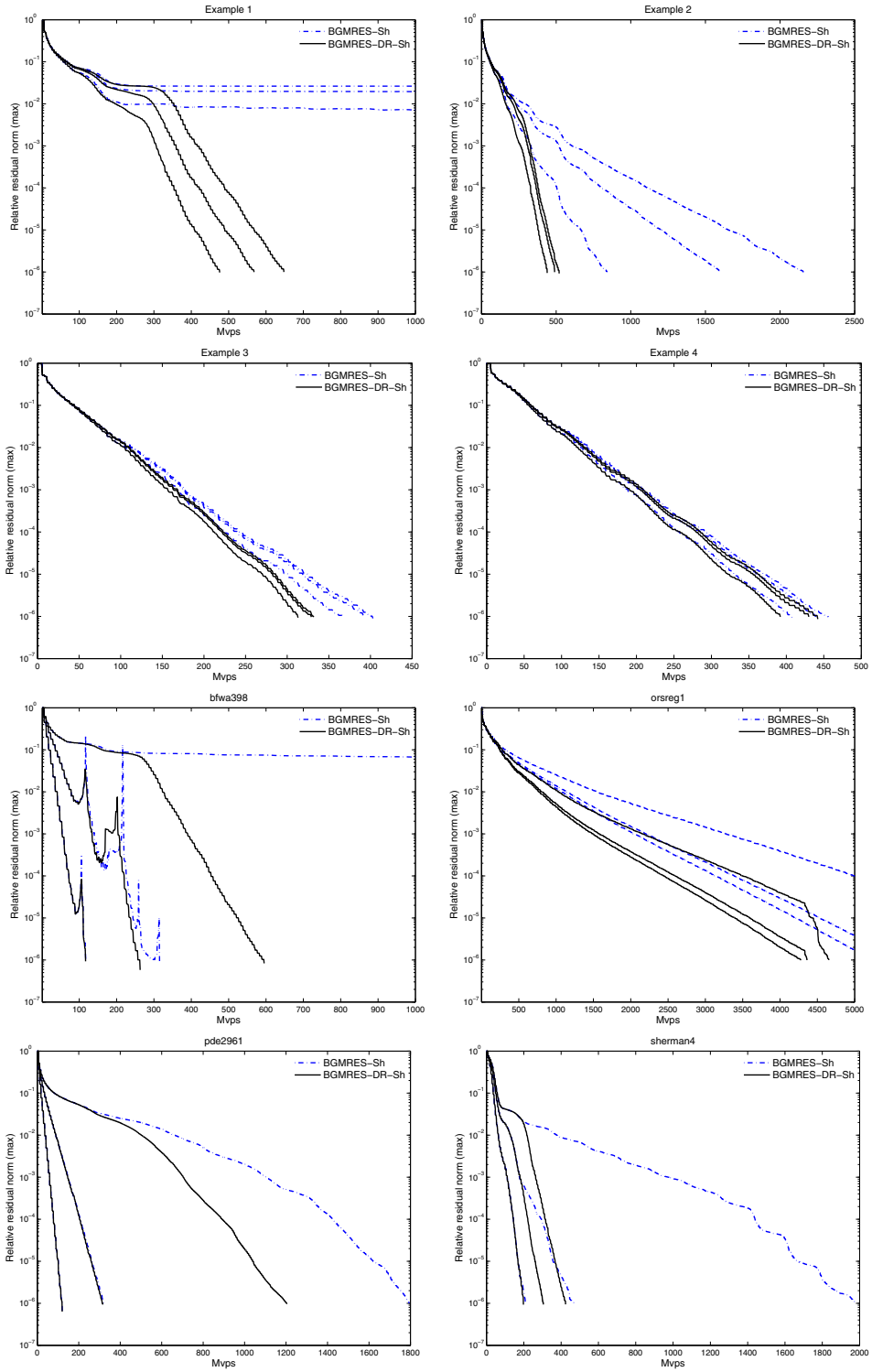

FIGURE 1 Convergence histories for Examples 1–8. The different convergence curves of the shifted block Krylov methods correspond to different choices of the shifts. BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting

deflating small eigenvalues is very beneficial in terms of convergence. For problems that do not exhibit such a cluster of eigenvalues close to zero, such as Examples 3 and 4, the performances of the BGMRES-Sh and BGMRES-DR-Sh methods are similar. In Figure 2, we depict the distribution of the smallest harmonic Ritz values over the spectrum of the coefficient matrices for Examples 5–8. The presence of eigenvalues close to zero makes these problems difficult to solve. We can notice that the BGMRES-DR-Sh method does a good job in approximating the eigenvalues to be deflated. This explains the overall good performance improvement.

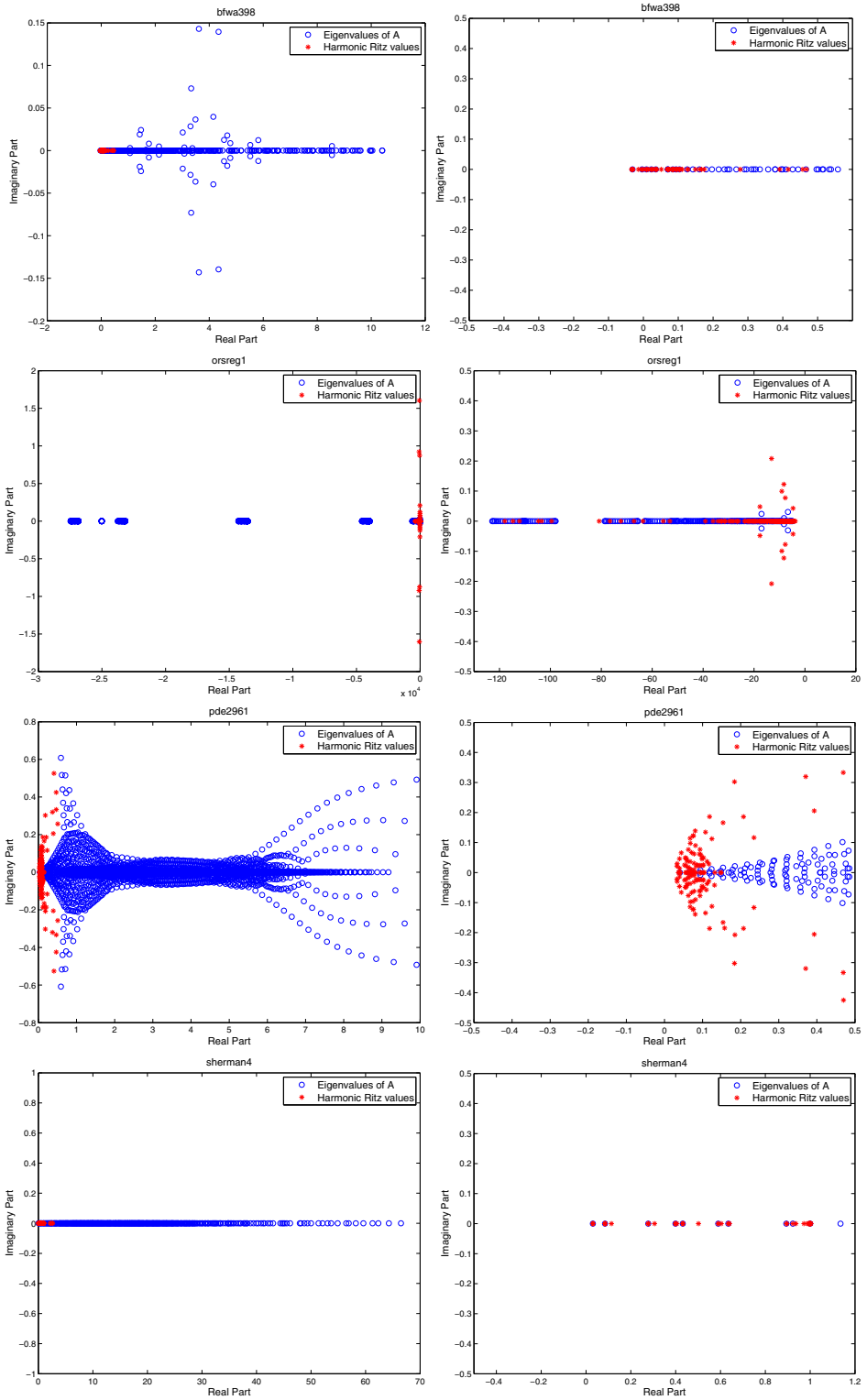

FIGURE 2 Distribution of the smallest harmonic Ritz values and plot of the spectrum for Examples 5, 6, 7, 8

4.1.1 | Performance analysis for different choices of shifts

We further analyze the effectiveness of the BGMRES-DR-Sh method with different choices of shifts for solving the same sequence of linear systems on Examples 1 and 2. In the experiments reported on in Table 3, we use the set of shifts $\Sigma_1 = [0.99, 0.95, 0.90]$ and $\Sigma_2 = [-1, -1.5, -2]$. For the shifts Σ_1, the BGMRES-Sh method failed to converge, whereas the BGMRES-DR-Sh method performed well. As Σ_1 shifts the spectrum of the coefficient matrices close to origin, potentially

TABLE 3 Number of *Mvps* and solution time in seconds (*Cpu*) for the BGMRES-Sh and BGMRES-DR-Sh methods using different shifts

Example	Method	Σ_1		Σ_2	
		Mvps	***Cpu***	***Mvps***	***Cpu***
1	BGMRES-Sh	–	–	1999	2.0308
	BGMRES-DR-Sh	**784**	**0.9411**	**514**	**0.5965**
2	BGMRES-Sh	–	–	1129	1.3462
	BGMRES-DR-Sh	**728**	**0.8334**	**463**	**0.5414**

Note. BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting; *Mvps* = matrix-vector products.

slowing down the convergence, this proves that the deflated restarting strategy is useful to speed up the convergence. Shifting the matrix $A-\sigma_i I$ by Σ_2 improves the convergence of the BGMRES-Sh method. Moreover, the convergence curves of the BGMRES-DR-Sh method plotted in Figure 1 for different shifts are similar, showing that the shift does not have an important effect when deflation is used.[18]

4.1.2 | Performance analysis for different number of right-hand sides

In the experiments illustrated in Table 4, we increased the number of right-hand sides to $p = 9, 15, 18$, setting the restart parameter m equal to $m = 90, 120$. We see that in most cases, the BGMRES-DR-Sh method performed noticeably better than the BGMRES-Sh method in terms of *Mvps* and *Cpu* time. The larger the length of the restart cycle, the better the performance of the BGMRES-DR-Sh method. For instance, with $m = 90$ and $p = 18$, the BGMRES-DR-Sh method failed to converge on Example 5, whereas it performed well for $m = 120$. When the restart value is increased, the dimension of the shifted block Krylov subspace also increases, and the performance of the shifted block Krylov subspace methods improves. Note that in Table 4, the experiments on Example 6 are not reported because in this problem, it was not possible to achieve convergence, increasing the number of right-hand sides. Obviously, preconditioning is an essential component of the iterative solution, an issue that will be shortly discussed in Section 4.2.1.

4.1.3 | The effect of the seed selection strategy

In this section, we test the performance of the BGMRES-DR-Sh method using the seed selection strategy (denoted as S-BGMRES-DR-Sh). We consider the same matrix on Example 1. The right-hand sides are set to be $B_1 = \text{rand}(n, p)$, $B_2 = [b^{(1)}, b^{(2)}, \ldots, b^{(p)}]$ with $b^{(j)} = b^{(1)} + 10^{-3} * u^{(j)}, j = 2, 3, \ldots, p$, where $b^{(1)}, u^{(j)}$ are random vectors. The shifts are $\sigma = [-2, -1, -0.8, -0.4, 0, 0.95, 0.99]$, and we set $p = 10$ on the right-hand sides. Convergence histories for Example 1 and different sets are plotted in Figure 3. We see that the S-BGMRES-DR-Sh method seems to have about the same convergence as the BGMRES-DR-Sh method for all right-hand sides, and it requires a reduced number of *mvps* with respect to the BGMRES-DR-Sh method. In the cases the BGMRES-DR-Sh method failed to converge when it is deflated small number harmonic Ritz vectors.

4.2 | Comparison between BGMRES, GMRES-DR-Sh, and BGMRES-DR-Sh

The BGMRES-DR-Sh method solves all of the multishifted linear systems simultaneously and has the ability to deflate the small eigenvalues of the coefficient matrix. In this section, we compare its convergence performance against the standard BGMRES method[21] for solving $L = 4$ linear systems with multiple right-hand sides in sequence and against the shifted GMRES method with deflated restarting procedure (GMRES-DR-Sh[34]) for solving $p = 6, 15$ multishifted linear systems independently. A set of general sparse matrices was selected from the University of Florida Sparse Matrix Collection,[51] described in Table 5. As for quantum chromodynamics problems denoted as Π_3 and Π_4, each matrix $\Pi_i, i = 3, 4$ exists with a critical value κ_c with $\frac{1}{\kappa_c} < \frac{1}{\kappa} < \infty$ such that $A = I - \kappa\Pi_i$ is real-positive. Here, we defined $A = (\frac{1}{\kappa_c} + 10^{-3})I - \Pi_i$ as the base matrix. We set the shifts $\sigma_i = [-0.0001, -0.001, -0.01, -0.1]$, the number of harmonic Ritz vectors $k = 5$, and *maxiter* = 10,000 for the stopping criterion parameters. In the experiments shown in Table 6, we use $p = 6$ and 15 right-hand sides, setting the restart parameter m equal to $m = 8 \times p$ and $10 \times p$, respectively (for instance, $m = 48, 60$ for $p = 6$). Note that if the maximum dimension of the standard Krylov subspace method is equal to jp, then the maximum dimension for the corresponding block variant with p right-hand sides is j. Then, for a fair comparison, we set the maximum dimension of

TABLE 4 Number of *mvps* and solution time in seconds (*Cpu*) for the BGMRES-Sh and BGMRES-DR-Sh methods with different pairs (m, p)

m = 90		p = 9		p = 15		p = 18	
Example	**Method**	***Mvps***	***Cpu***	***Mvps***	***Cpu***	***Mvps***	***Cpu***
1	BGMRES-Sh	–	–	–	–	–	–
	BGMRES-DR-Sh	**1,060**	**1.2022**	**2310**	**3.0844**	**2953**	**4.3655**
2	BGMRES-Sh	4,096	4.0117	–	–	–	
	BGMRES-DR-Sh	**874**	**0.9805**	**1,770**	**2.4391**	**2,287**	**3.4517**
3	BGMRES-Sh	676	0.7166	1,531	1.9311	2,107	2.8840
	BGMRES-DR-Sh	**513**	**0.5888**	**1,075**	**1.4726**	**1,330**	**1.9596**
4	BGMRES-Sh	**760**	**0.8034**	**1,618**	**2.0419**	**2,155**	**2.9525**
	BGMRES-DR-Sh	766	0.8618	1,682	2.3256	2,475	3.4565
5	BGMRES-Sh	–	–	–	–	–	–
	BGMRES-DR-Sh	**856**	**0.7046**	–	–	–	–
7	BGMRES-Sh	3,286	3.7931	–	–	–	–
	BGMRES-DR-Sh	**1,807**	**2.5074**	**2,857**	**4.4156**	**3,396**	**5.5450**
8	BGMRES-Sh	4,141	3.2765	–	–	–	–
	BGMRES-DR-Sh	**672**	**0.8311**	**1,277**	**1.7176**	**1,836**	**2.6171**
m = 120		**p = 9**		**p = 15**		**p = 18**	
Example	**Method**	***Mvps***	***Cpu***	***Mvps***	***Cpu***	***Mvps***	***Cpu***
1	BGMRES-Sh	–	–	–	–	–	–
	BGMRES-DR-Sh	**862**	**1.1924**	**1,721**	**2.7151**	**2,099**	**3.6527**
2	BGMRES-Sh	3,065	3.5271	–	–	–	–
	BGMRES-DR-Sh	**736**	**1.0317**	**1,350**	**2.2377**	**1,828**	**3.2175**
3	BGMRES-Sh	598	0.7931	1,261	1.8944	1,633	2.6289
	BGMRES-DR-Sh	**496**	**0.6860**	**874**	**1.4261**	**1,131**	**1.9881**
4	BGMRES-Sh	703	0.9213	1,405	**2.1228**	2,005	3.2899
	BGMRES-DR-Sh	**674**	**0.9130**	**1,371**	2.2702	**1,819**	**3.2447**
5	BGMRES-Sh	–	–	–	–	–	–
	BGMRES-DR-Sh	**700**	**0.7003**	**1,281**	**1.4606**	**1,605**	**1.8817**
7	BGMRES-Sh	2,993	3.7289	–	–	–	–
	BGMRES-DR-Sh	**1,739**	**2.5624**	**2,752**	**4.5695**	**3,383**	**5.9732**
8	BGMRES-Sh	2,740	2.5472	–	–	–	–
	BGMRES-DR-Sh	**573**	**0.7814**	**1,057**	**1.5844**	**1,515**	**2.3469**

Note. BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting; *Mvps* = matrix-vector products.

TABLE 5 Characteristics of the test matrix problems used in our experiments

Example	Name	Size	Nonzeros	Problem kind
Π_1	poisson3Db	85,623	2,374,949	computational fluid dynamics problem
Π_2	light_in_tissue	29,282	406,084	electromagnetics problem
Π_3	conf6_0-8×8-20	49,152	1,916,928	quantum chromodynamics problem
Π_4	conf6_0-8×8-30	49,152	1,916,928	quantum chromodynamics problem

the search space $m = 8, 10$ in the GMRES-DR-Sh method. As shown in Table 6, the BGMRES-DR-Sh method outperforms the tested methods in terms of *Mvps* and *Cpu* time in all cases, and the shifted block Krylov subspace methods always perform better than the block nonshift variant. Especially for large right-hand sides number and small restart parameters, the BGMRES-DR-Sh method can converge to the targeted accuracy, whereas the other methods failed to converge.

4.2.1 | Experiments with preconditioning

We studied the effect of preconditioning by comparing the unpreconditioned BGMRES, BGMRES method with an *ILU*(0) preconditioner,[21] preconditioned BGMRES-Sh, and preconditioned BGMRES-DR-Sh methods (denoted as BGMRES, BGMRES (*ILU*), PBGMRES-Sh, and PBGMRES-DR-Sh, respectively) on Examples 5 and 6. We choose a right preconditioner in shifted form $A - \tau I$, for different values of τ. In our experiments, we used a sparse direct solver (the built-in

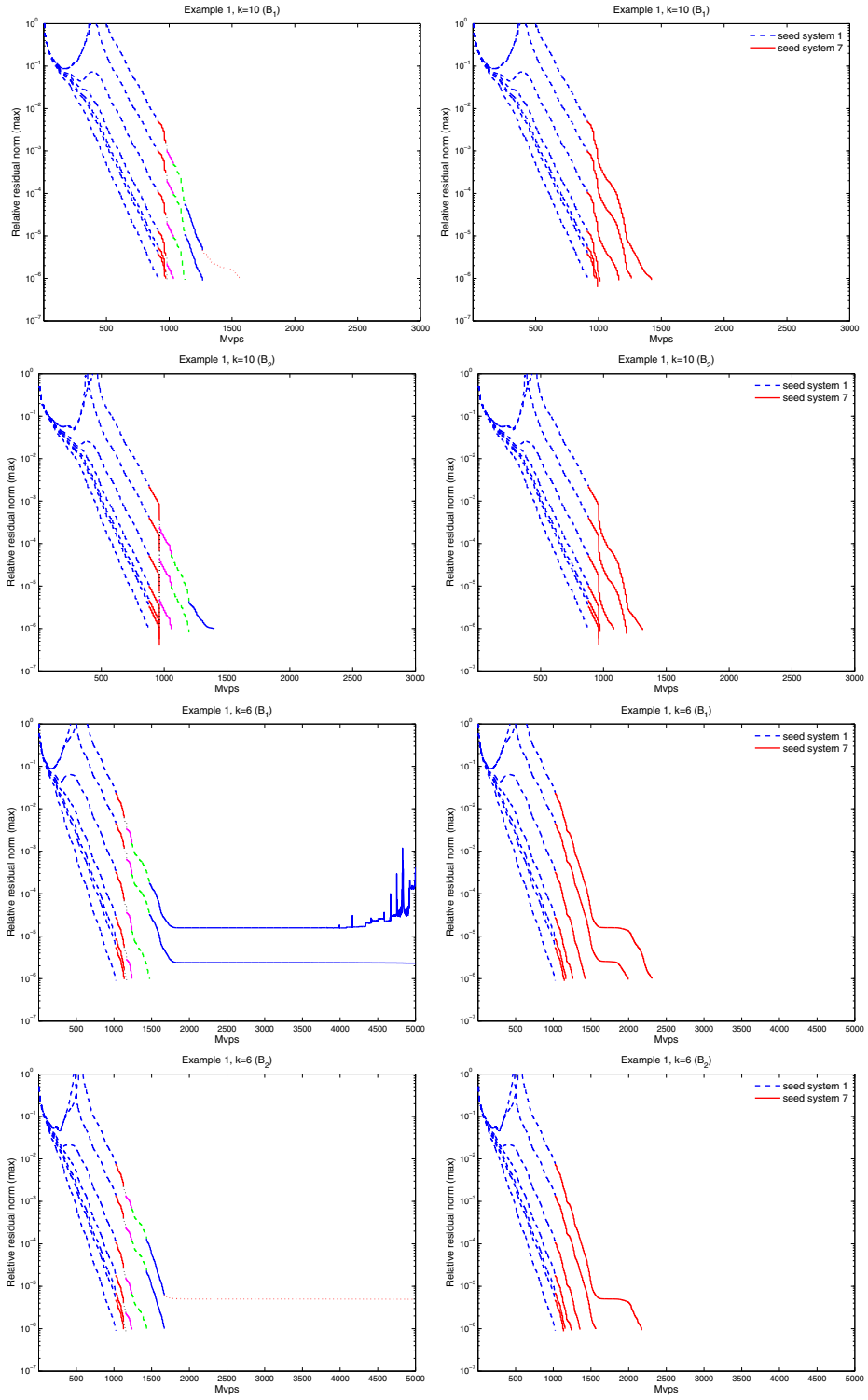

FIGURE 3 Comparative convergence histories for solving Example 1. Left: Convergence histories for the BGMRES-DR-Sh method. Right: Convergence histories for the S-BGMRES-DR-Sh method

MATLAB's "\" operator). As we can see in Table 7, the PBGMRES-DR-Sh method works well also in the cases when the BGMRES-DR-Sh method failed to converge. On Example 6, PBGMRES-Sh and PBGMRES-DR-Sh exhibit the same numerical behavior. This is due to the fact that the methods converge with no restart or only one restart. The BGMRES method with an $ILU(0)$ preconditioner still failed to converge on Example 6. On Example 5 with $\tau = -1.8, -3$, the PBGMRES-Sh method does not work, whereas the PBGMRES-DR-Sh method performed quite well. For different values of τ, systems with shifts close to $|\tau|$ converge rapidly.[31]

TABLE 6 Numerical behavior of the BGMRES, GMRES-DR-Sh, BGMRES-Sh, and BGMRES-DR-Sh methods for solving general test problems

		Example							
$m(p=6)$	Method	Π_1		Π_2		Π_3		Π_4	
		Mvps	*Cpu*	*Mvps*	*Cpu*	*Mvps*	*Cpu*	*Mvps*	*Cpu*
48	BGMRES	6,798	115.2054	–	–	–	–	–	–
	GMRES-DR-Sh	–	–	–	–	–	–	–	–
	BGMRES-Sh	4,478	73.3457	–	–	6,439	121.4615	6,565	125.6879
	BGMRES-DR-Sh	**2,303**	**45.5708**	**5,240**	**60.8445**	**3,963**	**86.4901**	**2,826**	**60.6758**
60	BGMRES	5,748	108.3229	–	–	–	–	–	–
	GMRES-DR-Sh	3,439	48.7659	–	–	7,634	117.8173	6,154	100.8173
	BGMRES-Sh	3,619	62.3493	–	–	5,533	118.8265	5,791	121.1183
	BGMRES-DR-Sh	**2,060**	**43.1373**	**4,382**	**54.4453**	**3,319**	**79.5106**	**2,432**	**56.6361**
		Example							
$m(p=15)$	Method	Π_1		Π_2		Π_3		Π_4	
		Mvps	*Cpu*	*Mvps*	*Cpu*	*Mvps*	*Cpu*	*Mvps*	*Cpu*
120	BGMRES	–	–	–	–	–	–	–	–
	GMRES-DR-Sh	–	–	–	–	–	–	–	–
	BGMRES-Sh	–	–	–	–	–	–	–	–
	BGMRES-DR-Sh	**5110**	**157.7496**	–	–	**8,561**	**325.3643**	**6,191**	**245.8153**
150	BGMRES	–	–	–	–	–	–	–	–
	GMRES-DR-Sh	8,521	164.7189	–	–	–	–	–	–
	BGMRES-Sh	9,682	285.2324	–	–	9,041	391.2042	9,886	423.4163
	BGMRES-DR-Sh	**4,416**	**141.5174**	–	–	**6,667**	**290.1264**	**5,026**	**221.3367**

Note. GMRES = generalized minimum residual method; BGMRES = block GMRES; BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting; *Mvps* = matrix-vector products.

TABLE 7 Numerical behavior of the Krylov solves for Examples 5 and 6 with $p = 18, m = 90$

Method		Example 5		Example 6	
		Mvps	*Cpu*	*Mvps*	*Cpu*
	BGMRES	–	–	–	–
	BGMRES (*ILU*)	1,139	0.8559	–	–
	BGMRES-DR-Sh	–	–	–	–
$\tau = 0.1$	PBGMRES-Sh	287	0.8524	96	1.4196
	PBGMRES-DR-Sh	226	0.6355	96	1.4196
$\tau = -0.2$	PBGMRES-Sh	273	0.7864	**90**	**1.2973**
	PBGMRES-DR-Sh	**202**	**0.5380**	**90**	**1.2973**
$\tau = -1.8$	PBGMRES-Sh	–	–	99	1.5356
	PBGMRES-DR-Sh	491	1.2429	99	1.5356
$\tau = -3$	PBGMRES-Sh	–	–	126	1.9215
	PBGMRES-DR-Sh	669	1.6881	126	1.9215

Note. Mvps = matrix-vector products; BGMRES = block GMRES; BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting.

In Figure 4, we display the spectrum of the preconditioned matrix *bfwa*398 using the right preconditioning and the distribution of the smallest harmonic Ritz values. It can be seen that the preconditioner can cluster effectively most of the eigenvalues near one. However, it still leaves a few outliers close to zero. This problem can be observed for many preconditioners and applications and motivates the use of our method also in combination with preconditioning, due to its ability to deflate the smallest eigenvalues near zero.

Observe that on Example Π_2 in Table 6, none of the tested methods converged to the targeted accuracy. On Example Π_2, we choose a shift-and-invert preconditioner in form $A - \tau I$ with $\tau = -0.0005$ for the GMRES-DR-Sh, BGMRES-Sh, and BGMRES-DR-Sh methods, and an *ILU*(0) preconditioner is implemented in the BGMRES method. As shown in Table 8, the PBGMRES-DR-Sh and PBGMRES-Sh methods need less *Mvps* and *Cpu* time, whereas the PGMRES-DR-Sh method still failed to converge.

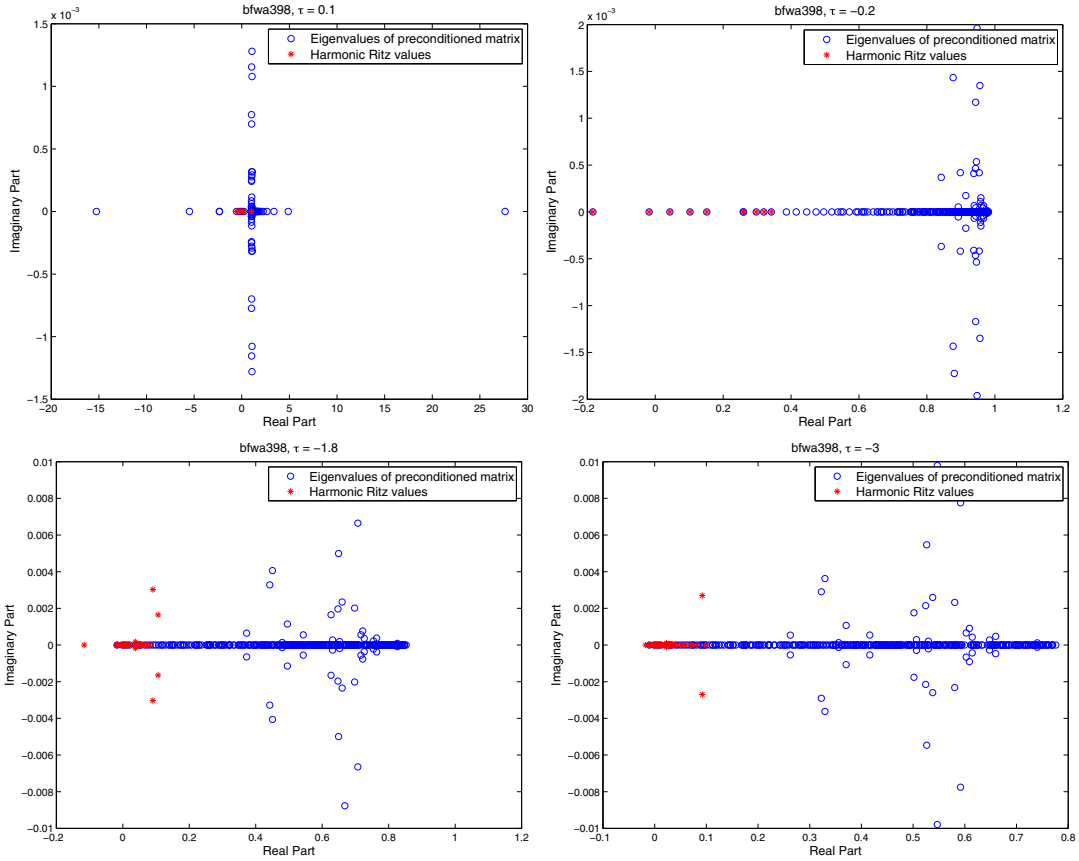

FIGURE 4 Distribution of the smallest harmonic Ritz values and the spectrum of the preconditioner matrix

TABLE 8 Numerical behavior of the preconditioned Krylov methods on Example $\Pi_2$with $p = 15, m = 120$

Method	*Mvps*	*Cpu*
BGMRES (*ILU*)	5501	115.7811
PGMRES-DR-Sh	–	–
PBGMRES-Sh	**126**	**66.1002**
PBGMRES-DR-Sh	**126**	**66.1002**

Note. *Mvps* = matrix-vector products; BGMRES = block GMRES.

TABLE 9 Characteristics of Web matrices

Name	Size	Nonzeros	Description
wb-cs-stanford	9,914	36,854	Stanford CS web
web-stanford	281,903	2,312,497	Web graph of Stanford.edu
cnr-2000	325,557	3,216,152	Small Web crawl of Italian CNR domain
wikipedia-20051105	1,634,989	19,753,078	Wikipedia pages

4.3 | A case study in PageRank application

In the recent years, with the booming development of the internet, Web search engines have become an extremely important tool for Web information retrieval. Due to the use of its efficient PageRank algorithm,[4] Google is one the most popular Web search engine. The PageRank model uses graph theory to rank data relevance, that is, the Web link structure can be represented by an $n \times n$ matrix P whose element $p_{i,j}$ is the probability of transitioning from page j to page i. In a practical

setting, it may be required to solve the PageRank problem with multiple damping factors and multiple personalization vectors. Mathematically, this problem can be rewritten as follows:

$$(I - \alpha_i P^T)X_i = B, \quad \text{with} \quad i = 1, 2, \ldots, L, \tag{37}$$

where $0 < \alpha_i < 1$ are the damping factors, $B = [b_1, \ldots, b_p]$ with $b_j (j = 1, \ldots, p)$ are the personalization vectors, and the coefficient matrices $I - \alpha_i P^T$ are huge. Note that when the damping factors are close to 1, the problems are more difficult to solve due to a clustering of eigenvalues close to zero in the spectrum of the coefficient matrix. This is an ideal setting to test the BGMRES-Sh and BGMRES-DR-Sh methods. In our experiments, we selected test matrices from the University of Florida Sparse Matrix Collection.[51] The characteristics of the Web matrices used in our experiments are listed in Table 9. We chose as damping vectors the values $\alpha = 0.99, 0.97, 0.95$. From Figure 5, it can be seen that BGMRES-DR-Sh

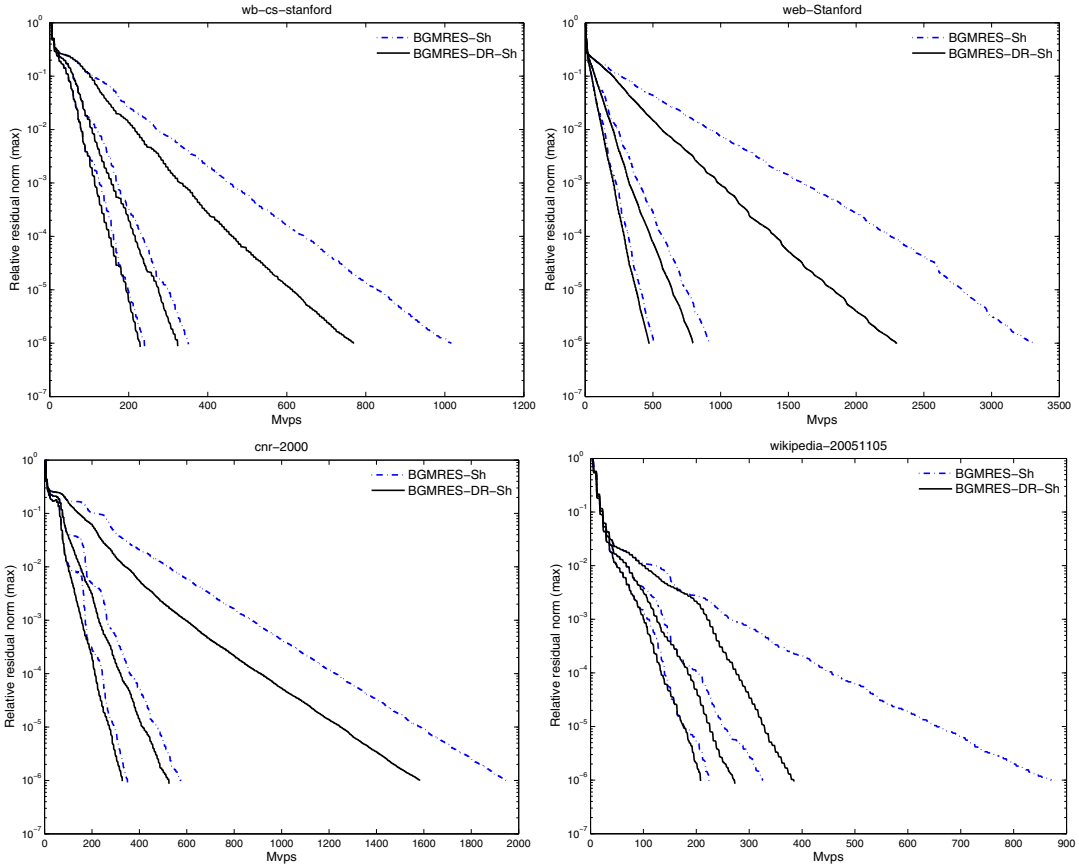

FIGURE 5 Convergence histories of Web matrices. BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting

TABLE 10 Numerical behavior of BGMRES-Sh and BGMRES-DR-Sh with $\varepsilon = 10^{-6}$

Example	Method	*Mvps*	*Trr*(max)	*Cpu*
wb-cs-stanford	BGMRES-Sh	1,016	9.5271e−07 9.7461e−07 9.8496e−07	3.2448
	BGMRES-DR-Sh	**769**	8.7789e−07 9.6572e−07 9.8117e−07	**2.3543**
web-stanford	BGMRES-Sh	3,313	9.9783e−07 9.8968e−07 9.9764e−07	361.1401
	BGMRES-DR-Sh	**2,299**	9.9723e−07 9.8244e−07 9.7780e−07	**280.3105**
cnr-2000	BGMRES-Sh	1,952	9.9987e−07 9.9210e−07 9.7871e−07	224.5758
	BGMRES-DR-Sh	**1,583**	8.6513e−07 9.2656e−07 9.8939e−07	**202.7926**
wikipedia-20051105	BGMRES-Sh	871	9.8664e−07 9.3966e−07 9.8496e−07	682.8839
	BGMRES-DR-Sh	**385**	8.5560e−07 9.4734e−07 9.3959e−07	**338.0457**

Note. BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting; *Mvps* = matrix-vector products.

TABLE 11 Sensitivity of BGMRES-DR-Sh to the accuracy of the harmonic Ritz vectors computation

cnr-2000 $m = 100$			
Backward error	***Mvps***	**Time(eigs)**	***Cpu***
10^{-6}	1,499	0.2085	208.2818
10^{-7}	1,499	0.2093	209.2815
10^{-8}	1,488	0.2096	207.9408
10^{-15}	1,488	0.4678	212.7469
web-stanford m=150			
Backward error	***Mvps***	**Time (eigs)**	***Cpu***
10^{-5}	1,883	0.1625	318.6143
10^{-6}	1,878	0.1733	312.9319
10^{-7}	1,878	0.1683	315.6753
10^{-15}	1,878	0.2481	318.8239
wikipedia-20051105 m=250			
Backward error	***Mvps***	**Time (eigs)**	***Cpu***
10^{-2}	354	0.6462	634.3478
10^{-3}	354	0.8059	638.7437
10^{-5}	354	0.8195	620.6223
10^{-15}	354	0.8196	630.9228
wb-cs-stanford m=270			
Backward error	***Mvps***	**Time (eigs)**	***Cpu***
10^{-2}	598	0.1858	3.7101
10^{-3}	603	0.1897	3.7301
10^{-5}	603	0.2001	3.7246
10^{-15}	603	0.2421	3.7699

Note. BGMRES-DR-Sh = Shifted BGMRES method with deflated restarting; *Mvps* = matrix-vector products.

converges faster than BGMRES-Sh, and from Table 10, it clearly outperforms BGMRES-Sh in terms of both *mvps* and *cpu* time.

4.4 | Sensitivity of BGMRES-DR-Sh to the accuracy of the eigencomputation

As mentioned earlier, the harmonic Ritz vectors are computed using ARPACK. A natural question is to study how the performance of the proposed solver depend on the accuracy of the harmonic Ritz vectors. In order to investigate this parameter, we monitor the backward error associated with the computation of each eigenvector. In Table 11, we give the number of *mvps* and the solution time for BGMRES-DR-Sh for different backward errors in the computation of the harmonic Ritz vectors. The *time*(*eigs*) reported in the table is the time spent for computing the harmonic Ritz vectors, whereas the quantity *cpu* is the total elapsed time for solving the sequence of multiple linear systems. We conclude that high accuracy is not necessary. For instance, on the wikipedia-20051105 example, harmonic Ritz vectors computed with a backward error of about 10^{-2} are as effective to reduce the number of iterations as those calculated with a backward error of about 10^{-15}. With the larger m, the harmonic Ritz vectors are more difficult to solve. In general, the larger m, the higher *mvps* we need to solve the problems. However, as illustrated in Table 11, the BGMRES-DR-Sh can perform well when a backward error is about 10^{-2} and $m = 270$.

5 | CONCLUSIONS

In this paper, we have presented the development of a new shifted block Krylov subspace method that combines the attractive numerical properties of the BGMRES-DR and BGMRES-Sh methods for solving the sequence of linear systems with multiple shifts and multiple right-hand sides simultaneously. The effectiveness of the proposed method is mainly due to its ability to deflate the smallest eigenvalues near zero. Our numerical experiments have shown that the BGMRES-DR-Sh method has a good potential to outperform the BGMRES, BGMRES-Sh method, and GMRES-DR-Sh method on several

realistic matrix problems, with and without using a preconditioner. The design of robust preconditioning techniques that maintain the shift-invariance property of the Krylov subspace and can be combined with BGMRES-DR-Sh will be the topic of a future study.

ACKNOWLEDGEMENTS

This research is supported by National Natural Science Foundation of China (NSFC) (61772003). The first author is also funded by the University of Groningen Ubbo Emmius scholarship.

ORCID

Dong-Lin Sun http://orcid.org/0000-0003-4750-6730
Yan-Fei Jing http://orcid.org/0000-0002-6120-9583

REFERENCES

1. Frommer A, Nöckel B, Güsken S, Lippert T, Schilling K. Many masses on one stroke economic computation of quark propagators. Int J Modern Phys C. 1995;6(5):627–638.
2. Frommer A, Maass P. Fast CG-based methods for Tikhonov-Phillips regularization. SIAM J Sci Comput. 1999;20(5):1831–1850.
3. Celis MR, Dennis JE, Tapia RA. A trust region strategy for nonlinear equality constrained optimization. In: Boggs PT, Byrd RH, Schnabel RB, editors. Proceedings SIAM Conference on Numerical Optimization. Philadelphia: SIAM; 1985. p. 71–82.
4. Haveliwala TH. Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search. IEEE Trans Knowl Data Eng. 2003;15(4):784–796.
5. Fletcher R. Conjugate gradient methods for indefinite systems. Lecture Notes in Mathematics, volume 506. New York: Springer, 1976.
6. Jing Y-F, Huang T-Z, Zhang Y, et al. Lanczos-type variants of the COCR method for complex nonsymmetric linear systems. J Comput Phys. 2009;228(17):6376–6394.
7. Saad Y, Schultz MH. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J Sci Stat Comput. 1986;7(3):856–869.
8. Sun D-L, Jing Y-F, Huang T-Z, Carpentieri B. A quasi-minimal residual variant of the BiCORSTAB method for nonsymmetric linear systems. Comput Math Appl. 2014;67(10):1743–1755.
9. Baumann M, Gijzen MB. Nested Krylov methods for shifted linear systems. SIAM J Sci Comput. 2015;37(5):S90–S112.
10. Du L, Sogabe T, Zhang S-L. IDR(s) for solving shifted nonsymmetric linear systems. J Comput Appl Math. 2015;274:35–43.
11. Frommer A. BiCGStab(l) for families of shifted linear systems. Computing. 2003;70(2):87–109.
12. Frommer A, Glässner U. Restarted GMRES for shifted linear systems. SIAM J Sci Comput. 1998;19(1):15–26.
13. Gu X-M, Huang T-Z, Meng J, Sogabe T, Li H-B, Li L. BiCR-type methods for families of shifted linear systems. Comput Math Appl. 2014;68(7):746–758.
14. Jing Y-F, Huang T-Z. Restarted weighted full orthogonalization method for shifted linear systems. Comput Math Appl. 2009;57(9):1583–1591.
15. Kilmer ME, de Sturler E. Recycling subspace information for diffuse optical tomography. SIAM J Sci Comput. 2006;27(6):2140–2166.
16. Simoncini V. Restarted full orthogonalization method for shifted linear systems. BIT Numer Math. 2003;43(2):459–466.
17. Soodhalter KM, Szyld DB, Xue F. Krylov subspace recycling for sequences of shifted linear systems. Appl Numer Math. 2014;81:105–118.
18. Darnell D, Morgan RB, Wilcox W. Deflated GMRES for systems with multiple shifts and multiple right-hand sides. Linear Algebra Appl. 2008;429(10):2415–2434.
19. Baker AH, Dennis JM, Jessup ER. On improving linear solver performance: A block variant of GMRES. SIAM J Sci Comput. 2006;27(5):1608–1626.
20. Boyse WE, Seidl AA. A block QMR method for computing multiple simultaneous solutions to complex symmetric systems. SIAM J Sci Comput. 1996;17(1):263–274.
21. Saad Y. Iterative methods for sparse linear systems. 2nd ed. Philadelphia: SIAM; 2003.
22. Simoncini V, Gallopoulos E. A hybrid block GMRES method for nonsymmetric systems with multiple right-hand sides. J Comput Appl Math. 1996;66(1):457–469.
23. Greenbaum A. Iterative methods for solving linear systems frontiers in applied mathematics, vol. 17. Philadelphia: SIAM; 1997.
24. Carpentieri B, Duff IS, Giraud L. A class of spectral two-level preconditioners. SIAM J Sci Comput. 2003;25(2):749–765.
25. Carpentieri B, Giraud L, Gratton S. Additive and multiplicative two-level spectral preconditioning for general linear systems. SIAM J Sci Comput. 2007;29(4):1593–1612.
26. Gutknecht MH. Deflated and augmented Krylov subspace methods: A framework for deflated BiCG and related solvers. SIAM J Matrix Anal Appl. 2014;35(4):1444–1466.
27. Morgan RB. GMRES with deflated restarting. SIAM J Sci Comput. 2002;24(1):20–37.

28. Morgan RB. Restarted block-GMRES with deflation of eigenvalues. Appl Numer Math. 2005;54(2):222–236.
29. Agullo E, Giraud L, Jing Y-F. Block GMRES method with inexact breakdowns and deflated restarting. SIAM J Matrix Anal Appl. 2014;35(4):1625–1651.
30. Calandra H, Gratton S, Lago R, Vasseur X, Carvalho LM. A modified block flexible GMRES method with deflation at each iteration for the solution of non-hermitian linear systems with multiple right-hand sides. SIAM J Sci Comput. 2013;35(5):S345–S367.
31. Meerbergen K, Bai Z. The lanczos method for parameterized symmetric linear systems with multiple right-hand sides. SIAM J Matrix Anal Appl. 2010;31(4):1642–1662.
32. Stathopoulos A, Orginos K. Computing and deflating eigenvalues while solving multiple right-hand side linear systems with an application to quantum chromodynamics. SIAM J Sci Comput. 2010;32(1):439–462.
33. Zhong H-X, Wu G, Chen G-L. A flexible and adaptive simpler block GMRES with deflated restarting for linear systems with multiple right-hand sides. J Comput Appl Math. 2015;282:139–156.
34. Wu G, Wang Y-C, Jin X-Q. A preconditioned and shifted GMRES algorithm for the pagerank problem with multiple damping factors. SIAM J Sci Comput. 2012;34(5):A2558–A2575.
35. Simoncini V, Gallopoulos E. An iterative method for nonsymmetric systems with multiple right-hand sides. SIAM J Sci Comput. 1995;16(4):917–933.
36. Simoncini V, Szyld DB. Recent computational developments in Krylov subspace methods for linear systems. Numer Linear Algebra Appl. 2007;14(1):1–59.
37. Gu G-D, Zhou X-L, Lin L. A flexible preconditioned Arnoldi method for shifted linear systems. J Comput Math. 2007;25:522–530.
38. Saibaba AK, Bakhos T, Kitanidis PK. A flexible Krylov solver for shifted systems with application to oscillatory hydraulic tomography. SIAM J Sci Comput. 2013;35(6):A3001–A3023.
39. Giraud L, Gratton S, Pinel X, Vasseur X. Flexible GMRES with deflated restarting. SIAM J Sci Comput. 2010;32(4):1858–1878.
40. Paige CC, Parlett BN, Van der Vorst HA. Approximate solutions and eigenvalue bounds from Krylov subspaces. Numer Linear Algebra Appl. 1995;2(2):115–133.
41. Calandra H, Gratton S, Langou J, Pinel X, Vasseur X. Flexible variants of block restarted GMRES methods with application to geophysics. SIAM J Sci Comput. 2012;34(2):A714–A736.
42. Gutknecht MH. Block Krylov space methods for linear systems with multiple right-hand sides: An introduction. In: Siddiqi AH, Duff IS, Christensen O, editors. Modern mathematical models, methods and algorithms for real world systems. New Delhi: Anamaya Publishers, 2007; p. 420–447.
43. Robbé M, Sadkane M. Exact and inexact breakdowns in the block GMRES method. Linear Algebra Appl. 2006;419(1):265–285.
44. Chan TF, Wan WL. Analysis of projection methods for solving linear systems with multiple right-hand sides. SIAM J Sci Comput. 1997;18(6):1698–1721.
45. Smith CF. The performance of preconditioned iterative methods in computational electromagnetics [PhD thesis]. Urbana-Champaign Illinois: University of Illinois; 1987.
46. Jing Y-F, Yuan P, Huang T-Z. A simpler GMRES and its adaptive variant for shifted linear systems. Numer Linear Algebra Appl. 2017; 24(1). https://doi.org/doi:10.1002/nla.2076.
47. Meerbergen K. The solution of parametrized symmetric linear systems. SIAM J Matrix Anal Appl. 2003;24(4):1038–1059.
48. Benzi M, Bertaccini D. Approximate inverse preconditioning for shifted linear systems. BIT Numer Math. 2003;43(2):231–244.
49. Luo W-H, Huang T-Z, Li L, Zhang Y, Gu X-M. Efficient preconditioner updates for unsymmetric shifted linear systems. Comput Math Appl. 2014;67(9):1643–1655.
50. Lehoucq RB, Sorensen DC, Yang C. ARPACK users' guide: Solution of large scale eigenvalue problems with implicitly restarted arnoldi methods. Philadelphia: SIAM; 1998.
51. Davis TA, Hu Y. The University of Florida sparse matrix collection. NA Digest. 1994:92.

How to cite this article: Sun D-L, Huang T-Z, Jing Y-F, Carpentieri B. A block GMRES method with deflated restarting for solving linear systems with multiple shifts and multiple right-hand sides. *Numer Linear Algebra Appl*. 2018;e2148. https://doi.org/10.1002/nla.2148