Variants of the block GMRES method for solving multi-shifted linear systems with multiple right-hand sides simultaneously
Sun, Donglin

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
1 The mathematical problem

Numerical computations have become an important tool for studying many problems in science and engineering. Simulations of lattice quantum chromodynamics \[41\], hydraulic tomography \[80\], seismic problems \[6\], electromagnetic waves \[82\] and PageRank computations \[50\], to name only a few applications areas, require to solve linear systems with multiple shifts and multiple right-hand sides. The base form of these systems is given by

\[(A - \sigma_i I)X_i = B, \quad i = 1, 2, \ldots, L,\]

where the \(L\) scalars \(\{\sigma_i\}_{i=1}^{L} \subset \mathbb{C}\) are the shifts, \(A - \sigma_i I \in \mathbb{C}^{n \times n}\) are large square nonsingular sparse matrices of dimension \(n\) (here we denote by \(I\) the identity matrix of order \(n\)), \(B = [b^{(1)}, b^{(2)}, \ldots, b^{(p)}]\) is the \(n \times p\) matrix of the \(p\) right-hand sides \(b^{(j)}\) for \(j = 1, \ldots, p\) given at once, and \(X_i \in \mathbb{C}^{n \times p}\) is the unknown solution matrix corresponding to the given shift \(\sigma_i\).

The PageRank model uses graph theory to rank data relevance \[44\]. Basically, it models the behaviour of a random Web surfer who keeps visiting Web pages by clicking on hyperlinks available on the currently visited page or by jumping to an external page chosen at random. The linking structure of the related Web pages is represented by a directed graph named the Web link graph. Denote its adjacency matrix by \(G \in \mathbb{N}^{n \times n}\) where \(n\) is the number of nodes (pages), and \(G(i,j)\) is nonzero (being 1) only when page \(j\) has a link pointing to page \(i\). Then the transition matrix \(P \in \mathbb{R}^{n \times n}\) with respect to the Web link graph is defined as

\[
P(i,j) = \begin{cases} \frac{1}{\sum_{k=1}^{n} G(k,j)}, & \text{if } G(i,j) = 1, \\ 0, & \text{otherwise}. \end{cases}
\]

Finally, the PageRank model can be mathematically formulated as:

\[Ax = v, \quad \text{with } A = (I - \alpha P),\]
where $0 < \alpha < 1$ is the damping factor that determines the weight assigned in the model to the Web linking structure, $v \in \mathbb{R}^{n \times 1}$ ($v \geq 0$ and $\|v\|_1 = 1$) is called the personalization vector, and the solution $x$ is the unknown Web ranking vector. In some instances one needs to solve a finite number of deterministic PageRank problems, each for a different combination of damping factors $\alpha_i$ and personalization vectors $v^{(j)}$ (see e.g. [22]), leading to the solution of sequences of shifted linear systems with multiple right-hand sides. In quenched quantum chromodynamics using Wilson and Clover fermions, with application to quark propagator calculations, the right-hand sides $b^{(j)}$ represent different noise vectors and the shifts $\sigma_i$ correspond to different quark masses that are used in an extrapolation process [41, 53]. In the development of microwave and millimeter-wave circuits and modules in electromagnetics, the coefficient matrix is shifted by different multiples of the identity. The algebraic systems of equations not only have multiple right-hand sides corresponding to multiple sources, but also have multiple shifts for each right-hand side [82].

Direct methods based on variants of the Gaussian Elimination algorithm are robust and predictable in terms of both accuracy and costs [29, 30]. However, the Gaussian Elimination method can solve only one system at a time, requiring $O(n^2)$ storage and $O(n^3)$ floating-point arithmetic operations for a dense matrix with dimension $n$. Sparse direct methods require $O(n) + O(nnz)$ memory and algorithmic cost for a sparse matrix of order $n$ with $nnz$ nonzeros [26, p.108]. Additionally, direct methods need to access all the entries of the coefficient matrix, and these are not always available in practical applications. Therefore, we study iterative solution strategies that are matrix-free and thus can solve the memory bottlenecks of direct methods for this problem class [32, 55, 79, 85, 96].

The most basic iterative method for solving a general linear system

$$Ax = b,$$  \hspace{1cm} (1.2)

where $A$ is a large and sparse matrix of size $n$ and $b$ is the given right-hand side vector, is called Richardson iteration. It splits the coefficient matrix $A$ as $A = I - (I - A)$, and approximates the solution $x$ to (1.2) as

$$x_i = b + (I - A)x_{i-1} = x_{i-1} + r_{i-1},$$  \hspace{1cm} (1.3)

where $r_{i-1} = b - Ax_{i-1}$ is the residual at the $(i - 1)$ step of the iteration. Upon multiplying both sides of (1.3) with $-A$, and adding $b$, the following
relations are obtained:

\[ b - Ax_i = b - Ax_{i-1} - Ar_{i-1} = (I - A)r_{i-1} = (I - A)^i r_0 = P_i(A) r_0, \]

where \( P_i(A) = (I - A)^i \) and \( r_0 = b - Ax_0 \) is the initial residual vector corresponding to the initial guess \( x_0 \). The above relations show that the convergence can be fast when \( \|I - A\|_2 \ll 1 \) \cite{26}.

Without loss of generality, we assume that \( x_0 = 0 \) and thus \( r_0 = b \). By repeating the Richardson iteration (1.3), it follows that

\[ x_{i+1} = r_0 + r_1 + \ldots + r_i = \sum_{j=0}^{i} (I - A)^j r_0. \]

Hence,

\[ x_{i+1} \in \text{span} \left\{ r_0, Ar_0, A^2 r_0, \ldots, A^i r_0 \right\}. \]

The space

\[ K_i(A, r_0) = \text{span} \left\{ r_0, Ar_0, A^2 r_0, \ldots, A^{i-1} r_0 \right\} \tag{1.4} \]

is referred to as the Krylov subspace of dimension \( i \) generated by matrix \( A \) and vector \( r_0 \). Richardson iterations search for the approximate solution of a linear system into Krylov subspaces of increasing dimension. Numerical methods that compute an approximate solution to Eq. (1.2) belonging to the space (1.4) are called Krylov subspace methods. Ipsen and Meyer \cite{52} presented the following theorem to explain why using Krylov subspaces to construct an approximate solution of linear systems.

**Theorem 1.** A square linear system \( Ax = b \) has a solution lying in the Krylov subspace \( K_i(A, r_0) \) if and only if

\[ b \in \mathcal{R}(A^j), \tag{1.5} \]

with

\[ i = \begin{cases} m & \text{if } A \text{ is nonsingular}, \\ m - j & \text{if } A \text{ is singular}, \end{cases} \]

where \( \mathcal{R}(A) \) denotes the range of the matrix \( A \), \( m \) is the degree of the minimal polynomial of \( A \), and \( j \) is the index of the zero eigenvalue of \( A \).
From the mid-1970s on, the development of more sophisticated iterative methods than the basic and slowly convergent Richardson method began to flourish. Dongarra and Sullivan in [27] have listed Krylov subspace methods among the “Top Ten Algorithms of the 20th Century” for their “greatest influence on the development and practice of science and engineering in the 20th century”.

The main goal of this thesis is to develop efficient Krylov subspace methods for solving sequences of linear systems with multiple shifts and multiple right-hand sides simultaneously.

This introductory chapter is organised as follows. In Section 1.1, a short background on Krylov subspace methods is presented. Iterative solution strategies for solving multi-shifted linear systems with multiple right-hand sides are discussed in Section 1.2. We state the main research questions in Section 1.3. Finally, the chapter terminates with the structure of the thesis in Section 1.4.

1.1 Iterative solution of linear systems

Over the past three decades, many Krylov-type iterative methods have been developed for solving large linear systems fast and efficiently. The Krylov subspace method for solving linear system (1.2) that constructs an approximation at the $m$th iteration

$$x_m = x_0 + t_m,$$  \hspace{1cm} (1.6)

where $t_m \in \mathcal{K}_m(A, r_0)$ such that the $m$th residual $r_m = b - Ax_m$ satisfies some constraint. From the definition of Krylov subspaces, (1.6) can be written as

$$x_m = x_0 + q_{m-1}(A)r_0,$$

where $q_{m-1}$ is a polynomial of degree at most $m - 1$. The corresponding residuals are

$$r_m = b - Ax_m = r_0 - Aq_{m-1}(A)r_0 = p_m(A)r_0.$$

with a polynomial $p_m(t) = 1 - tq_{m-1}(t)$ of degree at most $m$ with $p_m(0) = 1$. The approximation $x_m \in x_0 + \mathcal{K}_m$ (or equivalently, the corresponding polynomial) is often found by requiring $x_m$ to be the minimizer of some functional. Roughly speaking, they differ in the way the subspace basis
1.1. Iterative solution of linear systems

is generated and in the criteria used to extract the approximate solution from the search space in construction. Four different approaches can be distinguished.

1. The minimum residual approach computes the approximate solution
   $x_m$ that minimizes the residual norm
   \[ \|r_m\|_2 = \min_{x \in x_0 + \mathcal{K}_m(A, r_0)} \|b - Ax\|_2. \]  
   (1.7)
   The Generalized Minimum Residual Method (GMRES) \[79\] by Saad and Schultz is the most popular algorithm in this class. Note that throughout this thesis, the symbol $\| \cdot \|_q$ is used to denote the Euclidean norm when $q = 2$ and the Frobenius norm when $q = F$.

2. The Ritz-Galerkin approach computes $x_m$ such that
   \[ r_m \perp \mathcal{K}_m(A, r_0). \]  
   (1.8)
   This approach leads to popular iterative algorithms such as the Conjugate Gradient (CG) method \[51\] proposed by Hestenes and Stiefel for symmetric, positive and definite linear systems and the Full Orthogonalization Method (FOM) \[78\] for general nonsymmetric problems.

3. The Petrov-Galerkin condition imposes
   \[ r_m \perp \mathcal{L}_m, \]  
   (1.9)
   where $\mathcal{L}_m$ is another subspace of dimension $m$ somehow related to the $m$-dimensional Krylov subspace. When $\mathcal{L}_m$ generated with $A^H$ and some vector $s_0$, as $\mathcal{L}_m = \mathcal{K}_m(A^H, s_0)$, we obtain the Biconjugate Gradient \[32\] and Quasi-Minimal Residual (QMR) \[37\] methods. The superscript $H$ denotes the transpose conjugate operation. More recently, the Conjugate Gradient Squared (CGS) \[88\] and Biconjugate Gradient Stabilized (Bi-CGSTAB) \[101\] methods are prominent examples in this class.

4. Finally, the minimum error approach, solving
   \[ \min_{x_m \in x_0 + A^H \mathcal{K}_m(A, r_0)} \|x - x_m\|_2 \]  
   (1.10)
   is pursued in methods such as the Symmetric LQ (SYMMLQ) and the Generalized Minimal Error (GMERR) methods \[72\].
In this work we focus mainly on the Generalized Minimum Residual Method (GMRES) \[79\] that is broadly used in applications owing to its robustness and smoothly convergence behaviour. In the GMRES method, an orthogonal basis \( \{ v_1, \ldots, v_m \} \) of \( \mathcal{K}_m(A,r_0) \) is obtained by the so-called Arnoldi procedure sketched in Algorithm 1.1. At each step, the algorithm multiplies the previous Arnoldi vector \( v_j \) by \( A \) and then orthonormalizes the resulting vector \( w_j \) against all previous \( v_i \)'s, by applying the standard Gram-Schmidt procedure from linear algebra. After \( m \) steps, Algorithm 1.1 produces the following matrix decompositions

\[
AV_m = V_{m+1} \overline{H}_m, \quad (1.11)
\]
\[
V_m^T AV_m = H_m, \quad (1.12)
\]

where we denote by \( V_m \) the \( n \times m \) matrix with column vectors \( v_1, \ldots, v_m \), \( \overline{H}_m \) is the \((m+1) \times m\) Hessenberg matrix with entries \( h_{i,j} \), and \( H_m \) is the matrix obtained from \( \overline{H}_m \) by deleting its last row.

**Algorithm 1.1 Arnoldi algorithm.**

1. Choose a vector \( v_1 \) such that \( \|v_1\|_2 = 1 \)
2. for \( j = 1, 2, \ldots, m \) do
3. Compute \( w_j = Av_j \)
4. for \( i = 1, 2, \ldots, j \) do
5. \( h_{i,j} = (w_j, v_j) \)
6. \( w_j = w_j - h_{ij} v_i \)
7. \( h_{j+1,j} = \|w_j\|_2 \)
8. end for
9. if \( h_{j+1,j} = 0 \) then
10. stop
11. end if
12. \( v_{j+1} = w_j / h_{j+1,j} \)
13. end for

The GMRES solution \( x_m \) minimizes the residual norm \( r_m = b - Ax_m \) over all vectors in \( x_0 + \mathcal{K}_m \). It is computed such that

\[
\|r_m\|_2 = \|b - Ax_m\|_2 = \min_{x \in x_0 + \mathcal{K}_m(A,r_0)} \|b - Ax\|_2. \quad (1.13)
\]

Since any vector \( x_m \) in \( x_0 + \mathcal{K}_m \) can be written as

\[
x_m = x_0 + V_m y_m, \quad (1.14)
\]
from the Arnoldi relation (1.11) we obtain

\[ b - Ax_m = b - A(x_0 + V_m y_m), \]
\[ = r_0 - AV_m y_m, \]
\[ = \beta v_1 - V_{m+1} H_m y_m, \]
\[ = V_{m+1}(\beta e_1 - H_m y_m), \]

with \( \beta = \|r_0\|_2 \) and \( v_1 = r_0/\beta \). Finally, by applying the orthonormality of the column vectors of \( V_{m+1} \), the least-squares problem (1.13) can be rewritten in the form

\[ J(y_m) = \|b - Ax_m\|_2 = \|b - A(x_0 + V_m y_m)\|_2 = \|\beta e_1 - H_m y_m\|_2. \] (1.15)

The minimizer \( y_m \) is cheap to compute, since it requires to solve a small \((m + 1) \times m\) least-squares problem. However, as \( m \) increases, the memory and computational requirements of the Gram-Schmidt orthogonalization implemented in the Arnoldi procedure tend to become large, sometimes making the GMRES method impractical to solve high-dimensional linear systems arising in applications. Therefore, the algorithm is restarted periodically using the last computed approximation. For the sake of completeness, the restarted version of the GMRES method is sketched in Algorithm 1.2.

### 1.2 Krylov methods for sequences of multi-shifted linear systems with multiple right-hand sides

For the iterative solution of linear systems with multiple shifts and multiple right-hand sides given simultaneously by

\[ (A - \sigma_i I)X_i = B, \quad i = 1, 2, \ldots, L, \]

where \( B = [b^{(1)}, b^{(2)}, \ldots, b^{(p)}] \), three general routes can be followed. The first route is described below; the two other routes can be found in Section 1.2.2 and Section 1.2.3, respectively.

The first obvious approach is to solve each of the \( p \) multi-shifted linear systems independently, that is, to solve separately

\[ (A - \sigma_i I)x_i = b_j, \quad \text{with } i = 1, 2, \ldots, L, \] (1.16)
Algorithm 1.2  Restarted GMRES algorithm (GMRES(\(m\))).

1: Choose \(x_0\), compute \(r_0 = b - Ax_0\) and \(v_1 = r_0 / \|r_0\|_2\), \(\beta = \|r_0\|_2\)
2: \textbf{for} \(j = 1, 2, \ldots, m\) \textbf{do}
3:  \hspace{1em} Compute \(w_j = Av_j\)
4:  \hspace{1em} \textbf{for} \(i = 1, 2, \ldots, j\) \textbf{do}
5:  \hspace{2em} \(h_{i,j} = (w_j, v_j)\)
6:  \hspace{2em} \(w_j = w_j - h_{ij} v_i\)
7:  \hspace{2em} \(h_{j+1,j} = \|w_j\|_2\)
8:  \hspace{1em} \textbf{end for}
9:  \hspace{1em} \textbf{if} \(h_{j+1,j} = 0\) \textbf{then}
10:     \hspace{2em} goto line 13
11: \hspace{1em} \textbf{end if}
12: \hspace{1em} \(v_{j+1} = w_j / h_{j+1,j}\)
13: Define \(V_{m+1} = [v_1, v_2, \ldots, v_{m+1}]\), \(\overline{H}_m = (h_{k,l})_{1 \leq k \leq m+1, 1 \leq l \leq m}\)
14: Compute \(y_m = \arg \min_y \|\beta e_1 - \overline{H}_m y \|_2\)
15: \hspace{1em} \textbf{if} \(\|\beta e_1 - \overline{H}_m y \|_2 / \|b\|_2 \leq \text{tol}\) \textbf{then}
16:     \hspace{2em} \(x_m = x_0 + V_m y_m\), stop
17: \hspace{1em} \textbf{end if}
18: Compute \(x_m = x_0 + V_m y_m\)
19: Set \(x_0 = x_m\), \(r_0 = b - Ax_0\) and goto line 1
20: \textbf{end for}
for each right-hand side $b_j$. The shift $\sigma_1$ is generally called the base shift and the other shifts are called the additional shifts. Accordingly, throughout the thesis we will refer to $(A - \sigma_1 I)X_1 = B$ as the base block system and to $(A - \sigma_i I)X_i = B$, for $i = 2, \ldots, L$, as the shifted block systems. These so-called multi-shifted linear systems arise in a wide range of applications, like higher-order implicit methods for solving time-dependent partial differential equations [97], control theory [24], and quantum chromodynamics [80]. Here we consider the shifted Krylov subspace methods. The idea of shifted Krylov subspace methods is to save computational work by using every matrix-vector multiplication with the matrix $A$ for all $L$ systems instead of solving the systems individually. In this computational setting, it is advantageous to use shifted variants of Krylov subspace methods (see e.g. [6, 28, 38, 39, 54, 58, 83, 91]) that exploit the shift-invariance property of the Krylov subspace $K_m(A, \tilde{b})$ (1.4), as discussed in the next section.

1.2.1 The shifted GMRES method (GMRES-Sh)

Krylov subspaces satisfy the important shift-invariance property [39]

$$K_m(A, b) = \text{span}\{b, Ab, A^2b, \ldots, A^{m-1}b\} = K_m(A - \sigma_i I, \tilde{b}), \quad (1.17)$$

as long as the starting vectors are collinear, i.e., $\tilde{b} = \gamma b$ for some $\gamma \in \mathbb{C}\{0\}$. This invariance property has an important consequence: the Krylov basis generated for the base system can be used to approximate simultaneously the solution of the additional shifted systems by performing matrix-free operations [6, 28, 38, 39, 54, 58, 83, 91]. It is straightforward to establish the following result.

**Theorem 2.** For shifted systems, the standard Arnoldi relation (1.11) can be rewritten as

$$(A - \sigma_i I)V_m = V_{m+1}\overline{H}_m^i, \quad (1.18)$$

where $\overline{H}_m^i = \overline{H}_m - \sigma_i \begin{bmatrix} I_{m \times m} & 0_{1 \times m} \end{bmatrix}$. The matrix $V_m$ and $\overline{H}_m$ are the same as in (1.11) and are independent of the shift $\sigma_i$.

**Proof.** By subtracting $\sigma_i V_m$ from both sides of Eq. (1.11), it is straightforward to obtain the analogous shifted Arnoldi-like relation (1.18). □
Thus, the approximate solutions of the shifted systems can be written as

\[ x^i_m = x^i_0 + V_m y^i_m, \quad i = 1, \ldots, L, \quad (1.19) \]

where \( x^i_0 \) is the initial approximations. The corresponding shifted residuals are given by

\[
\begin{align*}
  r^i_m &= b - (A - \sigma_i I)(x^i_0 + V_m y^i_m), \\
        &= r^i_0 - (A - \sigma_i I)V_m y^i_m, \\
        &= c v_1 - V_{m+1} \overline{H}_m^i y^i_m, \\
        &= V_{m+1}(c e_1 - \overline{H}_m^i y^i_m). \\
\end{align*}
\quad (1.20)
\]

Note that the superscript \( i \) ranges from 1 to \( L \); it denotes quantities related to the \( i \)-th shift such as \( x_m, r_m, y_m \) and \( \overline{H}_m \). Then \( y^i_m \) can be computed by solving the following GMRES minimization problem

\[
y^i_m = \arg\min_{y^i} \left\| c e_1 - \overline{H}_m^i y^i \right\|_2,
\]

where we used the orthonormality of \( V_{m+1} \). Similar approaches can be established to derive unrestarted Krylov subspace methods for solving multi-shifted linear systems such as, e.g., the CG-Sh \([35]\), the BiCGStab-Sh \([38]\) and the QMR-Sh \([34]\) methods.

In many circumstances the search space built to approximate all of the shifted solutions satisfactorily within the required accuracy can grow very large, and the method needs to be restarted by taking the current system residuals as the new generating vectors. In other words, all systems need to have collinear right-hand sides for the restarted Krylov subspace methods. While for some methods, like the restarted FOM method \([83]\), the residuals are naturally collinear, making them straightforward choices for the solution of shifted systems, the GMRES shifted residuals are not collinear at restart in general. In \([39]\), Frommer and Glässner propose to modify the restarted GMRES method by forcing the residuals \( r^i_m \) to be collinear to \( r^1_m \), that is,

\[
r^i_m = \beta^i_m r^1_m, \quad i = 2, 3, \ldots, L, \quad (1.21)
\]

where

\[
r^i_m = b - (A - \sigma_i I)x^i_m = r^i_0 - V_{m+1} \overline{H}_m^i y^i_m. \quad (1.22)
\]
1.2. Krylov methods for sequences of multi-shifted linear systems with multiple right-hand sides

In order to exploit relation (1.17), we assume that the initial shifted residuals are collinear, \( r_0^i = \beta_0^i r_0 \), \( i = 2, \ldots, L \) (note that this relation is satisfied by taking \( x_0^i = 0 \)). Then, the collinearity condition (1.21) and Eqns. (1.20) and (1.22) give

\[
\begin{align*}
    r_m^i &= \beta_m^i r_0^1 \\
    \Leftrightarrow & \quad r_0^i - V_{m+1} \overline{H}_m^i y_m^i = \beta_m^i V_{m+1} z_{m+1}^i \\
    \Leftrightarrow & \quad \beta_0^i r_0^1 - V_{m+1} \overline{H}_m^i y_m^i = \beta_m^i V_{m+1} z_{m+1}^i \\
    \Leftrightarrow & \quad V_{m+1} (\overline{H}_m^i y_m^i + \beta_m^i z_{m+1}^i) = \beta_0^i r_0^1 \\
    \Leftrightarrow & \quad V_{m+1} (\overline{H}_m^i y_m^i + \beta_m^i z_{m+1}^i) = \beta_0^i V_{m+1} c e_1, \\
\end{align*}
\]

where we denote by \( z_{m+1} \) the quasi-residual vector \( z_{m+1} = ce_1 - \overline{H}_m y_m^i \). The unknowns \( y_m^i \) and \( \beta_m^i \) can be calculated easily by solving the following linear systems

\[
\begin{bmatrix}
    \overline{H}_m^i & z_{m+1}^i
\end{bmatrix}
\begin{bmatrix}
    y_m^i \\
    \beta_m^i
\end{bmatrix} = \beta_0^i c e_1, \quad i = 2, 3, \ldots, L. \quad (1.23)
\]

Finally, the approximate solutions of the additional shifted systems can be expressed as \( x_m^i = x_0^i + V_m y_m^i \), \( i = 2, \ldots, L \).

1.2.2 The block GMRES method (BGMRES)

The second approach for solving systems (1.1) is to apply block Krylov subspace methods to the \( L \) shifted linear systems with multiple right-hand sides in sequence. That is, to solve separately for each \( \sigma_i \) the linear system

\[
(A - \sigma_i I)X = B. \quad (1.24)
\]

By an abuse of notation, we denote the coefficient matrix of (1.24) still as \( A \). Block variants of Krylov methods have shown important advantages for the solution of multiple right-hand sides systems over methods designed for solving single right-hand side systems. They approximate \( X \) over the block Krylov subspace defined as

\[
\mathbb{K}_m(A, R_0) = \text{blockspan} \left( R_0, A R_0, \ldots, A^{m-1} R_0 \right),
\]

\[
= \left\{ \sum_{k=0}^{m-1} A^k R_0 \gamma_k \forall \gamma_k \in \mathbb{C}^{p \times p} \text{ with } k \mid 0 \leq k \leq m - 1 \right\} \in \mathbb{C}^{n \times p},
\]
Chapter 1. The mathematical problem

where $R_0 = B - AX_0$ is the initial block residual corresponding to the initial block guess $X_0$ of $X$. The block Krylov subspace $\mathbb{K}_m(A, R_0)$ contains all Krylov subspaces generated by each initial residual $\mathbb{K}_m(A, R_0(:, i))$, for every $1 \leq i \leq p$, plus all possible linear combinations of the vectors contained in these subspaces. For convenience of presentation, MATLAB notation is used; for example, $R_0(i,j)$ denotes the $(i,j)$-th entry of matrix $R_0$, and $R_0(:, i)$ corresponds to its $i$-th column. In block methods, each column $X_m(:, l)$ of the approximate solution $X_m$ is searched in the space $\sum_{j=1}^p \mathbb{K}_m(A, R_0(:, j))$, that is

\[
X_m(:, l) - X_0(:, l) \in \left\{ \sum_{k=0}^{m-1} \sum_{j=1}^p A^k R_0(:, j) \gamma_k(j, l) \gamma_k(j, l) \in \mathbb{C} \forall 1 \leq l \leq p \right\},
\]
\[
\in \sum_{j=1}^p \mathbb{K}_m(A, R_0(:, j)),
\]

(1.26)

whereas the Krylov solution for a single right-hand side is searched in $\mathbb{K}_m(A, R_0(:, j))$. Therefore, block Krylov subspace methods use much larger search spaces to approximate the solution of each linear system than conventional single right-hand side methods \cite{48,78}. In turn, they may have much higher memory and computational demands. However, they can solve the whole sequence at once and they perform block matrix-vector products, maximizing efficiency on modern cache-based computer architectures.

We review the block GMRES method which is an effective Krylov subspace solver for linear systems with multiple right-hand sides. The block GMRES method forms the starting point of our research. Let $R_0 = V_1 R$ be the $QR$-decomposition of $R_0$, 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. Block GMRES computes an orthonormal basis for $\mathbb{K}_m(A, R_0)$ by performing $m$ steps of the block Arnoldi procedure based on a Modified Block Gram-Schmidt process started with matrix $V_1$. The procedure is described in Algorithm 1.3. Let $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$ be the block orthonormal sequence spanning the block Krylov subspace $\mathbb{K}_m(A, R_0)$. Then the approximate solution to system \cite{1.24} can be written as $X_m = X_0 + \mathcal{V}_m Y_m$, where $Y_m$ minimizes the residual $R_m = R_0 - A \mathcal{V}_m Y_m$. An expression for $Y_m$ can be formed by using the block Arnoldi relation

\[
A \mathcal{V}_m = \mathcal{V}_{m+1} \mathcal{H}_m,
\]

(1.27)
1.2. Krylov methods for sequences of multi-shifted linear systems with multiple right-hand sides

where \( V_m \in \mathbb{C}^{n \times mp} \) and \( V_{m+1} \in \mathbb{C}^{n \times (m+1)p} \) are such that \( V_{m+1}^H V_{m+1} = I_{m+1} \), and \( \tilde{H}_m \in \mathbb{C}^{(m+1)p \times mp} \) has the following form:

\[
\tilde{H}_m = \begin{bmatrix}
    H_{1,1} & H_{1,2} & \cdots & H_{1,m} \\
    H_{2,1} & H_{2,2} & \cdots & H_{2,m} \\
    0_{p \times p} & H_{3,2} & \cdots & \vdots \\
    \vdots & \ddots & \ddots & \vdots \\
    0_{p \times p} & 0_{p \times p} & 0_{p \times p} & H_{m+1,m}
\end{bmatrix} = \begin{bmatrix} H_m & H_{m+1,m} E_m^H \end{bmatrix}.
\]

(1.28)

Note that \( H_m \in \mathbb{C}^{mp \times mp} \) in (1.28) is a block Hessenberg matrix with \( p \times p \) blocks \( H_{i,j} \), and \( E_m \) is the matrix consisting of the last \( p \) columns of \( I_{mp} \), where \( I_{mp} \in \mathbb{C}^{mp \times mp} \) is the identity matrix of dimension \( mp \). The minimization problem is relatively small, it requires to solve the least-squares problem

\[
Y_m = \arg\min_Y \| ER - \tilde{H}_m Y \|_F,
\]

(1.29)

with \( R_0 = V_1 R = V_{m+1} E R \), where \( E \) is an \((m+1)p \times p\) matrix whose upper \( p \times p \) principal block is the identity matrix. Throughout this thesis the block GMRES method will be shortly referred to as BGMRES.

Algorithm 1.3 Block Arnoldi algorithm.

1: Choose a unitary matrix \( V_1 \) of size \( n \times p \)
2: for \( j = 1, 2, \ldots, m \) do
3: \hspace{1em} Compute \( W_j = AV_j \)
4: \hspace{1em} for \( i = 1, 2, \ldots, j \) do
5: \hspace{2em} \( H_{i,j} = V_i^H W_j \)
6: \hspace{2em} \( W_j = W_j - V_i H_{i,j} \)
7: \hspace{1em} end for
8: \hspace{1em} \( W_j = V_{j+1} H_{j+1,j} \) (reduced QR-decomposition)
9: end for
10: Define \( V_{m+1} = [V_1, V_2, \ldots, V_{m+1}] \), \( \tilde{H}_m = (H_{k,l})_{1 \leq k \leq m+1, 1 \leq l \leq m} \)

1.2.3 The shifted block GMRES method (BGMRES-Sh)

The approaches described in the two previous sections can be merged in one single Krylov subspace formulation. Instead of applying shifted Krylov
subspace methods to the solution of each of the \( p \) multi-shifted linear systems independently \([28, 38, 39, 54, 56, 83, 91]\), or block Krylov subspace methods to the solution of the \( L \) linear systems with multiple right-hand sides in sequence \([1, 10, 48, 65]\), it would be much more efficient to use a shifted block method that solves the whole sequence of multi-shifted linear systems with multiple right-hand sides \((1.1)\) simultaneously \([95, 103]\). This can be achieved by taking advantage of the block shift-invariance property \([89]\) of the block Krylov subspace \( \mathbb{K}_m(A, R_0) \), which can be written as

\[
\mathbb{K}_m(A, R_0) = \mathbb{K}_m(A - \sigma_i I, R_0), \quad i = 1, \ldots, L,
\]  

and computing a unique basis that can be employed for solving the whole sequence of shifted block systems. Indeed, Eq. \((1.27)\) implies the Arnoldi-like relation

\[
(A - \sigma_i I) \mathbb{V}_m = \mathbb{V}_{m+1} \tilde{\mathbb{H}}_m^i,
\]

where

\[
\tilde{\mathbb{H}}_m^i = \tilde{\mathbb{H}}_m - \sigma_i \begin{bmatrix}
I_{mp \times mp} \\
0_{p \times mp}
\end{bmatrix} = \begin{bmatrix}
\mathbb{H}_m^i \\
H_{m+1, m} E_m^H
\end{bmatrix}.
\]

Due to the very large block orthogonalization costs (see, e.g., \([73]\), in particular Table 2.4 for some figures of complexity), the block Arnoldi procedure is often restarted periodically using the last computed approximation as starting point. As we mentioned in Section 1.2.1, in order to ensure the shift-invariant property, all systems need to have collinear residuals for the restarted Krylov subspace methods, like for the shifted GMRES \([39]\). The natural analog to this property for shifted block Krylov subspace methods is that the block residuals need to have columns spanning the same subspace. In this thesis, we adopt the terminology used in \([40]\) and denote blocks having this property as “cospatial”. Upon restarting the Arnoldi procedure after \( m \) steps, the block shift-invariance property does not hold anymore because the block residuals \( R_m^1 \) and \( R_m^i \), \( i = 2, \ldots, L \) of the base block system and of the additional shifted block systems are in general no longer cospatial. This means that \( \mathbb{K}_m(A - \sigma_1 I, R_m^1) \neq \mathbb{K}_m(A - \sigma_i I, R_m^i), \quad i = 2, \ldots, L \), where we denote by \( R_m^i = B - (A - \sigma_i I) X_m^i, \quad i = 1, 2, \ldots, L \) the block residuals at iteration \( m \) of the \( i \)-th shifted linear system.

In order to establish Eq. \((1.17)\) in the block case, Wu, Wang and Jin \([103]\) forced the shifted block residuals to be cospatial at each restart, i.e., they impose that

\[
R_m^i = R_m^i W_m^i, \quad i = 2, \ldots, L.
\]
The corresponding shifted block Krylov subspace method is called BGMRES-Sh. Here $W_m^i$ are $p \times p$ nonsingular matrices. Assuming that the initial shifted block residuals are cospatial, $R^i_0 = R^i_0 W^i_0$, $i = 2, \ldots, L$. Note that this relation is satisfied by taking $X^i_0 = 0$. The approximate solution of the base block system can be formulated as $X^1_m = X^1_0 + Y^1_m$, where $X^1_0$ is the initial approximation and $Y^1_m$ minimizes the base shifted block residual

\[
R^1_m = R^1_0 - (A - \sigma_1 I)Y^1_m
= Y^1_m(ER - \tilde{H}^1_m Y^1_m)
= Y^1_m R_{LS}.
\] (1.34)

In Eq. (1.35), we denote the block quasi-residual by $R_{LS} = ER - \tilde{H}^1_m Y^1_m$ and $R^i_0 = B - (A - \sigma_i I)X^i_0 = Y^i_{m+1} ER$, where $E$ is an $(m + 1)p \times p$ matrix whose upper $p \times p$ principal block is the identity matrix. Since $Y^i_{m+1}$ is orthonormal, it follows from Eq. (1.34) that $Y^i_m$ can be computed by solving the following minimization problem

\[
Y^i_m = \arg \min_{Y^i} \| ER - \tilde{H}^i_m Y^i_m \|_F.
\]

The additional shifted block residuals $R^i_m$ are given by

\[
R^i_m = B - (A - \sigma_i I)X^i_m = R^i_0 - Y^i_{m+1} \tilde{H}^i_m Y^i_m.
\] (1.36)

Combining the cospatial condition (1.33) with Eqns. (1.35) and (1.36) gives

\[
\begin{align*}
R^i_m &= R^i_0 W^i_m \\
\Leftrightarrow R^i_0 - Y^i_{m+1} \tilde{H}^i_m Y^i_m &= Y^i_{m+1} R_{LS} W^i_m \\
\Leftrightarrow R^1_0 W^i_0 - Y^i_{m+1} \tilde{H}^i_m Y^i_m &= Y^i_{m+1} R_{LS} W^i_m \\
\Leftrightarrow Y^i_{m+1}(\tilde{H}^i_m Y^i_m + R_{LS} W^i_m) &= R^1_0 W^i_0 \\
\Leftrightarrow Y^i_{m+1}(\tilde{H}^i_m Y^i_m + R_{LS} W^i_m) &= Y^i_{m+1} ER W^i_0.
\end{align*}
\]

The unknowns $Y^i_m$ and $W^i_m$ can be calculated easily by solving the following linear systems

\[
\begin{bmatrix}
\tilde{H}^i_m & \hat{V}
\end{bmatrix}
\begin{bmatrix}
Y^i_m \\
W^i_m
\end{bmatrix}
= ER W^i_0, \; i = 2, 3, \ldots, L.
\] (1.37)
Chapter 1. The mathematical problem

In Eq. (1.37) we have set \( \hat{W}_m^i = \hat{R}W_m^i \). Furthermore, \( \hat{V}, \hat{R} \) are the QR factors of the reduced QR-decomposition of \( R_{LS} \). The approximate solutions of the additional shifted block systems can be expressed as \( X_m^i = X_0^i + V_m Y_m^i, \) where \( Y_m^i, \) \( (i = 2, \ldots, L) \) is the solution of Eq. (1.37). The complete pseudocode of the BGMRES-Sh method is given as Algorithm 1.4.

Algorithm 1.4 BGMRES-Sh method.

1: \textit{Start}. Let the \( p \) linearly independent right-hand sides be \( B = [b^{(1)}, b^{(2)}, \ldots, b^{(p)}] \). Choose \( m \): the maximum size of the underlying block approximation Krylov subspace in each cycle, \( \epsilon \): the tolerance, \( \text{maxiter} \): 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 \( R_1^0 = V_1 R \).
\% Solve the base block system
5: Apply standard block GMRES(\( m \)). It generates \( V_{m+1}, \hat{H}_m^1 \) and forms the new approximate solutions \( X_1^m = X_0^1 + V_m Y_1^m \), where \( Y_1^m = \arg\min_{Y_1} \left\| ER - \hat{H}_m^1 Y_1 \right\|_F \).
Compute the residual vectors \( R_1^m = R_0^1 - V_{m+1} \hat{H}_m^1 Y_1^m \). Check convergence, and proceed if not satisfied.
\% Solve the additional shifted block systems
6: Solve \( \begin{bmatrix} \hat{H}_m^i & \hat{V} \end{bmatrix} \begin{bmatrix} Y_m^i \\ W_m^i \end{bmatrix} = ERW_m^i, \) \( i = 2, 3, \ldots, L, \) for \( Y_m^i \) and \( W_m^i \).
Then form the approximate solutions \( X_m^i = X_0^i + V_m Y_m^i \) and compute the residual vectors \( R_m^i \) as \( R_m^i = R_m^1 \hat{R}^{-1} \hat{W}_m^i \).
7: Check convergence, and restart if convergence is not achieved, i.e., go to 4.

1.3 Some research questions

Shifted block Krylov methods are computationally attractive to use as they compute a unique basis and use much larger spaces for solving the whole sequence of shifted block linear systems. Drawbacks are some
lack of general theory compared to the single right-hand side case and the large memory requirements due to the block orthogonalization procedure. Restarting periodically the algorithm may remedy memory problems, but it may also slow down the convergence significantly especially for matrices having small eigenvalues. It has been shown \cite{78} that at each iteration of GMRES, basically the next largest eigenvalue in magnitude is removed from the spectrum of the linear system. Therefore, the restarting procedure can destroy information on the very small eigenvalues of $A$ and the superlinear convergence may be lost \cite{102}. Although in theory the distribution of the eigenvalues alone does not determine the convergence of iterative Krylov solvers for non-Hermitian systems \cite{46}, for many problems and applications it has been observed that a tightly clustered spectrum around a single point away from the origin is generally very beneficial, whereas widely spread eigenvalues and the presence of clusters close to zero are often disadvantageous. Krylov methods build a polynomial expansion of the coefficient matrix that must be equal to one at zero and has to approximate zero on the set of the eigenvalues. Clearly, a low-degree polynomial with value 1 at the origin cannot be small at many points \cite{45}; therefore, “removing” the smallest eigenvalues can greatly improve the convergence \cite{4, 14, 16, 18, 66–69}.

There are essentially two different approaches for exploiting information about the smallest eigenvalues in the iterative solution procedure. The first idea is to compute a few, say $k$, approximate eigenvectors of $A$ corresponding to the $k$ smallest eigenvalues in magnitude, and augment the Krylov subspace with those directions. This approach is referred in the literature to as the augmented subspace approach (see \cite{13, 42, 66–69, 77, 95}). The second idea exploits spectral information gathered during the Arnoldi process to determine an approximation of an invariant subspace of $A$ associated with the eigenvalues nearest the origin, and uses this information to construct a preconditioner or to adapt an existing one \cite{4, 9, 16, 18, 31, 57}. In this thesis, we consider both approaches and use akin ideas to develop new robust variants of the shifted block BGMRES method for solving Eq. (1.1) more efficiently.

Another practical difficulty in the implementation of shifted block Krylov subspace methods is their lack of robustness due to the approximate linear dependence of the various residuals that may slow down or even prevent the convergence. A number of methods have been proposed to overcome this problem \cite{1, 12, 48, 60, 73, 76}. One technique is called explicit reduction.
or initial deflation \[48, 60, 76\]. Before starting the iterative solution, linear dependency in the block right-hand side \(B\) or in the initial block residual \(R_0\) can be detected by computing a singular value decomposition (SVD) of these matrices \[12, 48, 73\]. The smallest singular values can then be filtered out according to a certain deflation tolerance to remove the useless information from the block Krylov subspace. If the Krylov method is restarted, the strategy can be applied again at the beginning of each cycle. That is, after the computation of the block residual, to reduce effectively the Krylov space dimension. Recently, Calandra, Gratton, Langou, et.al. \[12\] have proposed a flexible variant of block GMRES method based on initial deflation (named as BFGMRESd) to detect the possible convergence of a linear combination of the systems. Another approach known as Arnoldi deflation \[48, Section 13\] detects a possible rank deficiency occurring during the block Arnoldi procedure and removes the linearly dependent columns of the matrix \(W_j\) at step 8 of Algorithm 1.3, ensuring that \(V_{j+1}\) has full rank. A rank-revealing QR algorithm can be used to determine the deficiency. As it often arises in applications, the problem of efficiently handling the approximate linear dependence of the right-hand sides is thoroughly addressed in this thesis.

### 1.3.1 Preconditioning techniques

Iterative methods can solve many of the memory bottlenecks of direct solution methods. To be computationally efficient, it is established that they need to be assisted by robust preconditioners that reduce notably the number of Krylov iterations required to converge. The main aim of preconditioning is to improve the spectral properties of the coefficient matrix in Eq. (1.2) by transforming the linear system into the equivalent form

\[
MAx = Mb, \tag{1.38}
\]

where \(M\) is a nonsingular matrix called preconditioner, close to \(A^{-1}\) in some sense, cheap to compute and to apply, and such that the eigenvalues of the preconditioned matrix \(MA\) are well clustered away from the origin. Consequently, solving \(1.38\) may be significantly faster than solving \(1.2\). Eq. \(1.38\) is preconditioned from the left, but one can also precondition from the right: \(AMy = b, x = M^{-1}y\). Hence the spectral properties of \(AM\) are to be considered.

For the solution of multi-shifted linear systems with multiple right-hand sides \(1.1\), the block shift-invariance property \(1.30\) may not be maintained
1.3. Some research questions

if conventional preconditioning techniques are used. Special preconditioners need to be used to ensure that

$$\mathbb{K}_m(AM^{-1}, B) = \mathbb{K}_m((A - \sigma_i I)\hat{M}^{-1}, B),$$  \hspace{1cm} (1.39)

where $\hat{M}$ is the preconditioner for the shifted systems. Note that the above condition is satisfied when the preconditioned shifted matrix can be written as a shifted preconditioned matrix

$$(A - \sigma_i I)\hat{M}^{-1} = AM^{-1} - \eta I.$$  \hspace{1cm} (1.40)

Two important examples of preconditioners that satisfy Eq. (1.39) are the shift-and-invert \cite{64, 80} and polynomial preconditioners \cite{2, 35, 53, 103} reviewed below.

**Shift-and-invert preconditioner**

We consider a shift-and-invert preconditioner of the form $M = A - \tau I$ for some $\tau \in \mathbb{C}$ \cite{80}:

$$(A - \sigma_i I)\hat{M}^{-1} = A(A - \tau I)^{-1} - \eta I$$

$$= [A - \eta(A - \tau I)](A - \tau I)^{-1}$$

$$= [A + \frac{\eta \tau}{1 - \eta}I](1 - \eta)(A - \tau I)^{-1}.$$  

If we choose

$$\frac{\eta \tau}{1 - \eta} = -\sigma_i,$$

then the matrix $\hat{M}^{-1}$ becomes

$$\hat{M}^{-1} = (1 - \eta)(A - \tau I)^{-1}.$$  

The shift-invariance property (1.39) is satisfied. Indeed, $\mathbb{K}_m((A - \sigma_i I)(A - \tau I)^{-1}, B) = \mathbb{K}_m(A(A - \tau I)^{-1}, B)$. Note that $\mathbb{K}_m(A(A - \tau I)^{-1}, B)$ is independent of $\sigma_i$. This means that the construction of the preconditioned Krylov subspace does not require matrix–vector products 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 solutions of the other shifted systems can be carried out at relatively low cost.
Polynomial preconditioner

As observed in [35,53], polynomial preconditioners also satisfy Eq. (1.39). Let

$$\hat{M}^{-1} = \hat{p}_N(A),$$

and

$$M^{-1} = p_N(A),$$

where $\hat{p}_N(A) = \sum_{i=0}^{N} \hat{\eta}_i A^i$ is a polynomial preconditioner of degree $N$ that is applied to the shifted matrix $(A - \sigma_i I)$ and $p_N(A) = \sum_{i=0}^{N} \eta_i A^i$ is a polynomial preconditioner of degree $N$ for $A$. The shift-invariance property can be satisfied by taking $\gamma_i$ and $\hat{\eta}_i$ such that

$$(A - \sigma_i I)\hat{p}_N(A) = Ap_N(A) - \gamma_i I.$$ 

In practice, the parameters $\gamma_i$ and $\hat{\eta}_i$ can be computed by using recursive formulas. We refer to the literature, for instance [2,35,53,103].

For other shifted preconditioning methods, we point the reader to the rich literature available in this field, for instance [2,5,6,8,80,90].

1.4 Structure of this thesis

This thesis is based on the journal papers [92–95]. It is organized as follows.

In Chapter 2, we present the development of a new variant of the shifted block GMRES method that recycles spectral information at each restart. This method combines the attractive numerical properties of the BGMRES-DR [69] and BGMRES-Sh [103] methods for solving sequence of linear systems with multiple shifts and multiple right-hand sides simultaneously. The effectiveness of the newly proposed solver, referred to as shifted block GMRES with deflated restarting, or shortly BGMRES-DR-Sh, is mainly due to its ability to mitigate the negative effects due to the presence of the eigenvalues near zero from the convergence.

In Chapter 3, we illustrate some difficulties with the implementation and use of shifted block Krylov subspace methods in the presence of approximately linearly dependent right-hand sides. We propose a variant of
the block shifted GMRES method based on initial deflation (DBGMRES-Sh) that may overcome these difficulties. The numerical experiments show a significant enhancement of robustness compared to the original BGMRES-Sh method. As preconditioning is mostly required in applications, we present a preconditioned DBGMRES-Sh algorithm that can accommodate both fixed and variable preconditioning at each iterate, still being able to take advantage of the block shift-invariance property.

Special care needs to be taken if a linear combination of the systems has converged, or if some of the right-hand sides converge much faster than others. In Chapter 4, we present a shifted, augmented and deflated BGMRES (SAD-BGMRES) method that can handle numerical problems due to the occurrence of inexact breakdowns in the block Arnoldi procedure. In addition, it recycles spectral information at restart for faster convergence. Finally, it solves the whole sequence of multi-shifted linear systems and multiple right-hand sides simultaneously exploiting the shift-invariance property of Krylov subspaces. Numerical experiments confirm the robustness and the effectiveness of the SAD-BGMRES algorithm. It outperforms conventional block Krylov algorithms for solving general sparse linear systems arising from different fields.

In Chapter 5, we exploit approximate invariant subspaces recycled over the iterations to adapt an existing preconditioner for the simultaneous solution of sequences of linear systems with multiple right-hand sides. This approach differs from the augmented subspace approach considered in Chapters 2 and 4. The new class of block iterative solvers has the ability to handle the approximate linear dependence of the block of right-hand sides and exploits approximate invariant subspaces recycled over the iterations to mitigate the bad effects that small eigenvalues can have on the convergence, by adapting an existing preconditioner. We illustrate the numerical behavior of the spectrally updated and initially deflated block GMRES (SPID-BGMRES) method on a set of linear systems arising from the discretization of the Dirac equation and of boundary integral equations in electromagnetics scattering.

In Chapter 6, we summarize our findings by comparing the theoretical properties and the numerical performance of the BGMRES-DR-Sh, the DBGMRES-Sh and the SAD-BGMRES methods for solving sequences of multi-shifted linear systems with multiple right-hand sides, also against the
conventional shifted block GMRES method (BGMRES-Sh). We conclude with final remarks and some plans for future research.