Numerical considerations of block GMRES methods when applied to linear discrete ill-posed problems

Lucas Onisk[a], Lothar Reichel[a], Hassane Sadok[b]

[a] *Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA.*
[b] *Laboratoire de Mathématiques Pures et Appliquées, Université du Littoral, Centre Universtaire de la Mi-Voix, Batiment H. Poincarré, 50 Rue F. Buisson, BP 699, 62228 Calais cedex, France.*

Abstract

Linear systems of equations with a matrix whose singular values decay to zero with increasing index number without a significant gap are commonly referred to as linear discrete ill-posed problems. We are interested in solving large systems of this kind when the right-hand side has $k > 1$ column vectors. The systems may be regarded as one system of equations with a block vector (with k columns) as the right-hand side and then solved by a block iterative method, or as k linear systems of equations (one for each right-hand side vector) that can be solved independently. Thus, the solution is a block vector with k columns. In many applications, including the restoration of color images, the right-hand side represents measurements that are contaminated by errors. Block iterative methods compute all columns of the solution block vector simultaneously. We will illustrate the performance of standard block GMRES methods and global GMRES methods, which also are block methods, and show that they may determine computed solutions of lower quality than when each column of the solution block vector is computed independently by a "standard" iterative method. We introduce a new local block GMRES method that can overcome the problems associated with block GMRES methods applied to linear discrete ill-posed problems.

Keywords: ill-posed problems, iterative method, block Arnoldi process, global Arnoldi process, GMRES, block GMRES
2010 MSC: 65F10, 65F22

1. Introduction

We are concerned with the approximate solution of linear systems of equations

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}^{\delta}, \tag{1}$$

Email addresses: lonisk@kent.edu (Lucas Onisk), reichel@math.kent.edu (Lothar Reichel), Hassane.Sadok@lmpa.univ-littoral.fr (Hassane Sadok)

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is a large matrix whose singular values decay to zero with increasing index number without a significant gap. This makes the matrix $\boldsymbol{A}$ severely ill-conditioned and possibly rank-deficient. Linear systems of equations with a matrix of this kind are commonly referred to as linear discrete ill-posed problems. They arise, for instance, from the discretization of a Fredholm integral equations of the first kind; see, e.g., Engl et al. [1] and Hansen [2] for discussions on ill-posed and linear discrete ill-posed problems, respectively. The right-hand side $\boldsymbol{B}^\delta = \left[\boldsymbol{b}_1^\delta, \ldots, \boldsymbol{b}_k^\delta\right] \in \mathbb{R}^{n\times k}$ is a block vector, i.e., a matrix with $k > 1$ columns. In many applications $k \ll n$, but this is not required by the algorithms considered in this paper.

In linear discrete ill-posed problems that arise in applications in science and engineering, the right-hand side typically represents measured data that is contaminated by an error, which is denoted by $\boldsymbol{E} \in \mathbb{R}^{n\times k}$. Throughout, we will assume that the entries of $\boldsymbol{E}$ are i.i.d. following a Gaussian distribution with mean zero. Let $\boldsymbol{B} \in \mathbb{R}^{n\times k}$ denote the unknown error-free block vector associated with $\boldsymbol{B}^\delta$. Then

$$\boldsymbol{B}^\delta = \boldsymbol{B} + \boldsymbol{E}.$$

We assume that $\boldsymbol{B}$ is in the range of $\boldsymbol{A}$, and that we are interested in computing an accurate approximation of the solution $\boldsymbol{X}^\dagger$ of minimal Frobenius norm of the unavailable linear systems of equations

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}. \tag{2}$$

The discrepancy principle (see below) can be used to determine how many steps of the chosen iterative method to carry out. Note that due to the ill-conditioning of $\boldsymbol{A}$ and the error $\boldsymbol{E}$ in $\boldsymbol{B}^\delta$, the least-squares solution of minimal Frobenius norm of (1) typically does not provide a useful approximation of $\boldsymbol{X}^\dagger$.

Let $\boldsymbol{X}_0 = \boldsymbol{0}$ be the initial iterate used by an iterative method. The successive iterates, $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$, generated by the methods discussed in this paper approach $\boldsymbol{X}^\dagger$ with increasing index number when the index number is small enough. However, due to the error in $\boldsymbol{B}^\delta$, iterates $\boldsymbol{X}_p$ with large index will diverge from $\boldsymbol{X}^\dagger$. Thus, one is interested in finding the optimal iterate $\boldsymbol{X}_s$ such that

$$\left\|\boldsymbol{X}_s - \boldsymbol{X}^\dagger\right\| \leq \left\|\boldsymbol{X}_p - \boldsymbol{X}^\dagger\right\|, \qquad p = 0, 1, \ldots \ .$$

Throughout this paper $\|\cdot\|$ denotes the Frobenius matrix norm unless otherwise noted.

Let a bound

$$\|\boldsymbol{E}\| \leq \delta$$

be known. The *discrepancy principle* prescribes that the iterations of an iterative method for the approximate solution of (1) be terminated as soon as an iterate $\boldsymbol{X}_p$ that satisfies

$$\left\|\boldsymbol{A}\boldsymbol{X}_p - \boldsymbol{B}^\delta\right\| \leq \tau\delta,$$

is found. Here $\tau > 1$ is a user-specified constant that is independent of δ; see [1] for a discussion on the discrepancy principle.

We note that when an upper bound for $\|\boldsymbol{E}\|$ is not available or the system (2) is not consistent, other approaches to identify a suitable iterate can be used, including the L-curve criterion and generalized cross validation; see, e.g., [3, 4, 5, 6, 7, 8, 9] for discussions. Some of these references are concerned with Tikhonov regularization, which is an alternative to the truncated iteration scheme outlined above. Many methods developed for Tikhonov regularization can be modified so that they can be used with truncated iteration.

The block generalized minimum residual (BGMRES) method is a popular iterative method for the solution of linear systems of algebraic equations with a large nonsymmetric matrix that stems from the discretization of a linear well-posed problem, such as an elliptic partial differential equations with Dirichlet boundary conditions and with a right-hand side block vector; see, e.g., [10]. The application of BGMRES to the solution of linear discrete ill-posed problems is described in, e.g., [11]. Tensor GMRES methods described in, e.g., [12, 13] also are block GMRES methods and the findings of the present paper apply.

Let the initial iterate be $\boldsymbol{X}_0 = \boldsymbol{0} \in \mathbb{R}^{n\times k}$. Then the p^{th} iterate, $\boldsymbol{X}_p$, computed by the standard BGMRES method applied to (1) satisfies

$$\left\|\boldsymbol{A}\boldsymbol{X}_p - \boldsymbol{B}^\delta\right\| = \min_{\boldsymbol{X}\in\mathbb{K}_p(\boldsymbol{A},\boldsymbol{B}^\delta)} \left\|\boldsymbol{A}\boldsymbol{X} - \boldsymbol{B}^\delta\right\|, \tag{3}$$

where

$$\begin{aligned}\mathbb{K}_p(\boldsymbol{A},\boldsymbol{B}^{\boldsymbol{\delta}}) &= \text{block span}\left\{\boldsymbol{B}^{\boldsymbol{\delta}}, \boldsymbol{A}\boldsymbol{B}^\delta, \ldots, \boldsymbol{A}^{p-1}\boldsymbol{B}^{\boldsymbol{\delta}}\right\}\\ &= \left\{\boldsymbol{X}\in\mathbb{R}^{n\times k} : \boldsymbol{X} = \sum_{i=0}^{p-1} \boldsymbol{A}^i\boldsymbol{B}^\delta\Omega_i,\ \ \Omega_i\in\mathbb{R}^{k\times k}\ \text{ for } i = 0,1,\ldots,p-1\right\}\end{aligned}$$

is a block Krylov subspace of order p; see, e.g., [14]. We tacitly assume that p is sufficiently small so that all required computations can be carried out without breakdown.

It is the purpose of the present paper to illustrate that block GMRES methods may yield inferior accuracy when compared to the standard GMRES methods applied to the k linear systems of equations with the matrix $\boldsymbol{A}$ and right-hand side vectors $\boldsymbol{b}_j^\delta$, $j = 1, 2, \ldots, k$. The ℓ-shifted BGMRES method introduced in [11] was found to often yield solutions of higher accuracy than the unshifted BGMRES method (0-shifted BGMRES). It is a block generalization of the GMRES-type method discussed in [15]. Our comparison also includes the ℓ-shifted global GMRES (gl-GMRES) method. This block method differs from the BGMRES methods in the choice of inner product in the block orthogonalization process. The gl-GMRES method without shift (0-shifted gl-GMRES) was introduced by Jbilou et al. [16, 17] and the shifted variant is discussed in [11]. Finally, we introduce local BGMRES (lo-BGMRES) methods (both shifted and unshifted) that can overcome the inferior results achieved by the aforementioned block GMRES methods.

This paper is organized as follows. Section 2 reviews the available block GMRES methods used in our comparison and introduces the new local block

GMRES methods, specifically, the lo-BGMRES and the ℓ-shifted lo-BGMRES methods. Computed examples are presented in Section 3, and concluding remarks can be found in Section 4.

2. Block GMRES methods

This section discusses in order the ℓ-shifted BGMRES method, the ℓ-shifted gl-GMRES method, and the new local BGMRES methods: lo-BGMRES and ℓ-shifted lo-BGMRES.

2.1. ℓ-shifted BGMRES

The block Arnoldi process, defined by Algorithm 1 below, is initialized by computing the QR factorization $\boldsymbol{B}^{\delta} = \hat{\boldsymbol{Q}}\hat{\boldsymbol{R}}$ with $\hat{\boldsymbol{Q}} \in \mathbb{R}^{n\times k}$ having orthonormal columns and $\hat{\boldsymbol{R}} \in \mathbb{R}^{k\times k}$ being upper triangular. If the columns of $\boldsymbol{B}^{\delta}$ are linearly dependent, then we reduce k so that the range of the so obtained matrix $\hat{\boldsymbol{Q}} \in \mathbb{R}^{n\times k'}$ with $k' < k$ agrees with the range of $\boldsymbol{B}^{\delta}$. We will for notational simplicity assume that $k' = k$. The p^{th} step of the block Arnoldi iteration applied to the matrix $\boldsymbol{A}$ with initial matrix $\boldsymbol{V}_1 = \hat{\boldsymbol{Q}}$ gives the block Arnoldi decomposition

$$\boldsymbol{A}\boldsymbol{V}_{pk} = \boldsymbol{V}_{(p+1)k}\boldsymbol{H}_{(p+1)k,pk}, \tag{4}$$

where the matrix $\boldsymbol{V}_{(p+1)k} = [\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_{p+1}] \in \mathbb{R}^{n\times(p+1)k}$ has orthonormal columns and $\boldsymbol{H}_{(p+1)k,pk} \in \mathbb{R}^{(p+1)k\times pk}$ is of upper block Hessenberg form with k subdiagonal blocks of size $p\times p$. The blocks $\boldsymbol{V}_i$ of $\boldsymbol{V}_{(p+1)k}$ for $i = 1, 2, \ldots, p+1$ are of size $n \times k$ and the matrix $\boldsymbol{V}_{pk} \in \mathbb{R}^{n\times pk}$ contains the first p blocks of $\boldsymbol{V}_{(p+1)k}$. Moreover,

$$\boldsymbol{V}_i^T\boldsymbol{V}_j = \begin{cases} \boldsymbol{I}_k & \text{for} \quad i = j, \\ \boldsymbol{O}_k & \text{for} \quad i \neq j, \end{cases}$$

where $\boldsymbol{I}_k$ and $\boldsymbol{O}_k$ denote the identity and zero matrices of order k, respectively. The range of the matrix $\boldsymbol{V}_{(p+1)k}$ is the block Krylov subspace $\mathbb{K}_{p+1}(\boldsymbol{A}, \boldsymbol{B}^{\delta})$ under the assumption that all upper Hessenberg matrices generated by Algorithm 1 are nonsingular. This algorithm is the foundation of the BGMRES method (see [10, 18]) and its ℓ-shifted variant.

The ℓ-shifted BGMRES method solves the minimization problem

$$\left\|\boldsymbol{A}\boldsymbol{X}_p^{(\ell)} - \boldsymbol{B}^{\delta}\right\| = \min_{\boldsymbol{X}\in\mathbb{K}_p(\boldsymbol{A},\boldsymbol{A}^{\ell}\boldsymbol{B}^{\delta})} \left\|\boldsymbol{A}\boldsymbol{X} - \boldsymbol{B}^{\delta}\right\|, \tag{5}$$

where the solution of minimal norm is sought in the range restricted block Krylov subspace

$$\mathbb{K}_p(\boldsymbol{A}, \boldsymbol{A}^{\ell}\boldsymbol{B}^{\boldsymbol{\delta}}) = \text{block span}\left\{\boldsymbol{A}^{\ell}\boldsymbol{B}^{\boldsymbol{\delta}}, \boldsymbol{A}^{\ell+1}\boldsymbol{B}^{\delta}, \ldots, \boldsymbol{A}^{\ell+p-1}\boldsymbol{B}^{\boldsymbol{\delta}}\right\} \tag{6}$$

for $\ell \in \mathbb{N}$ small (usually 1 - 4). When $\ell = 0$, we define $\boldsymbol{A}^{\ell} = \boldsymbol{I}_n$. In this case the above minimization problem reduces to (3). As part of the ℓ-shifted BGMRES

Algorithm 1: Block Arnoldi

Input: $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B}^{\delta} \in \mathbb{R}^{n\times k}$
Output: $\boldsymbol{V}_{(p+1)k} \in \mathbb{R}^{n\times(p+1)k}$ and $\boldsymbol{H}_{(p+1)k,pk} \in \mathbb{R}^{(p+1)k\times pk}$

Compute QR factorization $\left[\hat{\boldsymbol{Q}}, \hat{\boldsymbol{R}}\right] = \boldsymbol{B}^{\delta}$;
Set $\boldsymbol{V}_1 = \hat{\boldsymbol{Q}}$;
for $p = 1, 2, \ldots, n$ **do**
    Compute $\boldsymbol{W}_p = \boldsymbol{A}\boldsymbol{V}_p$;
    for $i = 1, 2, \ldots, p$ **do**
        $\boldsymbol{H}_{i,p} = \boldsymbol{V}_i^T \boldsymbol{W}_p$;
        $\boldsymbol{W}_p = \boldsymbol{W}_p - \boldsymbol{V}_i \boldsymbol{H}_{i,p}$;
    end
    Compute QR factorization $[\boldsymbol{V}_{p+1}, \boldsymbol{H}_{p+1,p}] = \boldsymbol{W}_p$;
end

method, the minimization problem (5) may be rewritten as

$$\left\|\boldsymbol{A}\boldsymbol{X}_p^{(\ell)} - \boldsymbol{B}^{\delta}\right\| = \min_{\boldsymbol{Y}\in\mathbb{R}^{pk\times k}} \left\|\boldsymbol{R}^{(\ell+1)}_{(\ell+p+1)k,pk}\boldsymbol{Y} - \left(\boldsymbol{Q}^{(\ell+1)}_{(\ell+p+1)k}\right)^T \boldsymbol{E}_1\hat{\boldsymbol{R}}\right\|, \quad (7)$$

where the matrix $\boldsymbol{E}_1 \in \mathbb{R}^{(\ell+p+1)k\times k}$ is made up of the first k columns of $\boldsymbol{I}_{(\ell+p+1)k}$ and zeros elsewhere. The matrices $\boldsymbol{Q}^{(\ell+1)}_{(\ell+p+1)k} \in \mathbb{R}^{(\ell+p+1)k\times(\ell+p+1)k}$ and $\boldsymbol{R}^{(\ell+1)}_{(\ell+p+1)k,pk} \in \mathbb{R}^{(\ell+p+1)k\times pk}$ come from the $(\ell+1)^{st}$ QR factorization of the method. Denote the solution of (7) by $\boldsymbol{Y}_p^{(\ell)}$. Then the approximate solution of (1) determined by the ℓ-shifted BGMRES method is given by $\boldsymbol{X}_p^{(\ell)} = \boldsymbol{V}_{(\ell+p)k}\boldsymbol{Q}^{(\ell)}_{(\ell+p)k,pk}\boldsymbol{Y}_p^{(\ell)}$. Algorithm 2 below computes $\boldsymbol{X}_p^{(\ell)}$ and is terminated according to the discrepancy principle. Further details and discussion can be found in [11].

2.2. ℓ-shifted gl-GMRES

We begin by introducing notation necessary to describe the global Arnoldi iteration and the gl-GMRES method. These techniques were first discussed by Jbilou et al. [16, 17]. We then outline the ℓ-shifted gl-GMRES algorithm.

The Kronecker product of two matrices $\boldsymbol{G} = [g_{i,j}] \in \mathbb{R}^{n\times n}$ and $\boldsymbol{H} \in \mathbb{R}^{p\times p}$ is defined by

$$\boldsymbol{G}\otimes\boldsymbol{H} := \begin{bmatrix} g_{1,1}\boldsymbol{H} & g_{1,2}\boldsymbol{H} & \cdots & g_{1,n}\boldsymbol{H} \\ g_{2,1}\boldsymbol{H} & g_{2,2}\boldsymbol{H} & \cdots & g_{2,n}\boldsymbol{H} \\ \vdots & \vdots & & \vdots \\ g_{n,1}\boldsymbol{H} & g_{n,2}\boldsymbol{H} & \cdots & g_{n,n}\boldsymbol{H} \end{bmatrix} \in \mathbb{R}^{np\times np}.$$

For two general matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we have

$$\begin{aligned} \boldsymbol{A}\otimes\boldsymbol{B} &= (\boldsymbol{I}\otimes\boldsymbol{B})(\boldsymbol{A}\otimes\boldsymbol{I}) \\ &= (\boldsymbol{A}\otimes\boldsymbol{I})(\boldsymbol{I}\otimes\boldsymbol{B}), \end{aligned} \quad (8)$$

Algorithm 2: ℓ-shifted BGMRES ($\ell \geq 1$) with discrepancy principle

Input: $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B}^{\delta} \in \mathbb{R}^{n\times k}$, and $\ell \in \{1, 2, 3, \dots\}$
Output: $\boldsymbol{X}_p^{(\ell)} \in \mathbb{R}^{n\times k}$

Compute QR factorization: $\left[\hat{\boldsymbol{Q}}, \hat{\boldsymbol{R}}\right] = \boldsymbol{B}^{\delta}$;
Set $\boldsymbol{V}_1 = \hat{\boldsymbol{Q}}$ & $\boldsymbol{X}_0^{(\ell)} = \boldsymbol{0}$;
Compute ℓ steps of block Arnoldi: $\boldsymbol{A}\boldsymbol{V}_{\ell k} = \boldsymbol{V}_{(\ell+1)k}\boldsymbol{H}_{(\ell+1)k,\ell k}$;
for $p = 1, 2, \dots$ **do**
    Compute next block Arnoldi step: $\boldsymbol{A}\boldsymbol{V}_{(\ell+p)k} = \boldsymbol{V}_{(\ell+p+1)k}\boldsymbol{H}_{(\ell+p+1)k,(\ell+p)k}$;
    Compute QR factorization: $\left[\boldsymbol{Q}^{(1)}_{(p+1)k}, \boldsymbol{R}^{(1)}_{(p+1)k,pk}\right] = \boldsymbol{H}_{(p+1)k,pk}$;
    for $j = 1, 2, \dots, \ell$ **do**
        Compute QR factorization: $\left[\boldsymbol{Q}^{(j+1)}_{(j+p+1)k}, \boldsymbol{R}^{(j+1)}_{(j+p+1)k,pk}\right] = \boldsymbol{H}_{(j+p+1)k,(j+p)k}\boldsymbol{Q}^{(j)}_{(j+p)k,pk}$;
    end
    Compute $\boldsymbol{Y}_p^{(\ell)}$ as the minimizer of $\left\| \boldsymbol{R}^{(\ell+1)}_{(\ell+p+1)k,pk}\boldsymbol{Y} - \left(\boldsymbol{Q}^{(\ell+1)}_{(\ell+p+1)k}\right)^T \boldsymbol{E}_1\hat{\boldsymbol{R}} \right\|$;
    Compute $\boldsymbol{X}_p^{(\ell)} = \boldsymbol{V}_{(\ell+p)k}\boldsymbol{Q}^{(\ell)}_{(\ell+p)k,pk}\boldsymbol{Y}_p^{(\ell)}$;
    Compute $\|\boldsymbol{r}_p\| = \left\| \boldsymbol{A}\boldsymbol{X}_p^{(\ell)} - \boldsymbol{B}^{\delta} \right\|$;
    if $\|\boldsymbol{r}_p\| \leq \tau\delta$ **then**
        Stop;
    end
end

for suitably sized identity matrices $\boldsymbol{I}$. We define vec$(\cdot)$ as the operation which transforms a general matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ to a vector $\boldsymbol{a} \in \mathbb{R}^{mn}$ by stacking the columns of $\boldsymbol{A}$ from left to right. Additionally, we define the inner product

$$\begin{aligned} \langle \boldsymbol{A}, \boldsymbol{B} \rangle_F &:= \operatorname{tr}(\boldsymbol{A}^T \boldsymbol{B}) \\ &= (\operatorname{vec}(\boldsymbol{A}))^T \operatorname{vec}(\boldsymbol{B}), \end{aligned} \tag{9}$$

where tr$(\cdot)$ denotes the trace.

Let $\boldsymbol{a} = [a_1\, a_2\, \cdots\, a_p]^T \in \mathbb{R}^p$ and introduce the product

$$\boldsymbol{V}_{pk} * \boldsymbol{a} = \sum_{i=1}^{p} a_i \boldsymbol{V}_i,$$

where $\boldsymbol{V}_{pk} = [\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_p] \in \mathbb{R}^{n\times pk}$ and $\boldsymbol{V}_i \in \mathbb{R}^{n\times k}$ for $i = 1, 2, \ldots, p$. Further, we have that

$$\begin{aligned} \boldsymbol{V}_{pk} * \boldsymbol{H} &= [\boldsymbol{V}_{pk} * \boldsymbol{H}_{:,1}, \boldsymbol{V}_{pk} * \boldsymbol{H}_{:,2}, \ldots, \boldsymbol{V}_{pk} * \boldsymbol{H}_{:,p}] \\ &= \boldsymbol{V}_{pk}\big(\boldsymbol{H} \otimes \boldsymbol{I}_k\big), \end{aligned} \tag{10}$$

where $\boldsymbol{H}_{:,i}$, for $i = 1, 2, \ldots, p$, denotes the i^{th} column of the matrix $\boldsymbol{H} \in \mathbb{R}^{p\times p}$. Additionally, it can be shown that

$$(\boldsymbol{V}_{pk} * \boldsymbol{H}) * \boldsymbol{a} = \boldsymbol{V}_{pk} * \boldsymbol{H}\boldsymbol{a}.$$

The global Arnoldi iteration, implemented by Algorithm 3, uses the inner product (9). This algorithm executes p steps with the matrix $\boldsymbol{A}$ and initial block vector $\boldsymbol{V}_1 = \boldsymbol{B}^\delta / \|\boldsymbol{B}^\delta\|$.

Algorithm 3: Global Arnoldi

Input: $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B}^\delta \in \mathbb{R}^{n\times k}$
Output: $\boldsymbol{V}_{(p+1)k} \in \mathbb{R}^{n\times(p+1)k}$ and $\boldsymbol{H}_{p+1,p} \in \mathbb{R}^{(p+1)\times p}$

```
Set V_1 = B^δ/‖B^δ‖;
for p = 1, 2, ..., n do
    Compute W_p := A V_p;
    for i = 1, 2, ..., p do
        h_{i,p} = ⟨W_p, V_i⟩_F;
        W_p := W_p − h_{i,p} V_i;
    end
    h_{p+1,p} = ‖W_p‖ ;
    V_{p+1} = W_p / h_{p+1,p};
end
```

At step p, the global Arnoldi iteration gives the relation

$$\boldsymbol{A}\boldsymbol{V}_{pk} = \boldsymbol{V}_{(p+1)k} * \boldsymbol{H}_{p+1,p}, \tag{11}$$

where $\boldsymbol{V}_{(p+1)k} = [\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_{p+1}] \in \mathbb{R}^{n\times(p+1)k}$ is the same as defined in the previous section and the block column vectors $\boldsymbol{V}_i \in \mathbb{R}^{n\times k}$ are F-orthonormal, i.e.,

$$\langle \boldsymbol{V}_i, \boldsymbol{V}_j \rangle_F = \begin{cases} 1 & \text{for} \quad i = j, \\ 0 & \text{for} \quad i \neq j. \end{cases} \tag{12}$$

The matrix $\boldsymbol{V}_{pk} \in \mathbb{R}^{n\times pk}$ is composed of the first p blocks of $\boldsymbol{V}_{(p+1)k}$. We also note that the Hessenberg matrix is of dimension $(p+1)\times p$. We assume that all subdiagonal entries of $\boldsymbol{H}_{p+1,p}$ are positive. This is the generic case. For linear discrete ill-posed problems breakdown occurs exceedingly rarely. We therefore will not discuss this issue.

The ℓ-shifted gl-GMRES method solves the minimization problem (5) which may be rewritten as

$$\min_{\boldsymbol{y}\in\mathbb{R}^p} \left\| \boldsymbol{R}^{(\ell+1)}_{\ell+p+1,p}\boldsymbol{y} - \|\boldsymbol{B}^\delta\| \left(\boldsymbol{Q}^{(\ell+1)}_{\ell+p+1}\right)^T \boldsymbol{e}_1 \right\|, \tag{13}$$

where $\boldsymbol{e}_1 \in \mathbb{R}^{\ell+p+1}$ denotes the first axis vector. The matrices $\boldsymbol{R}^{(\ell+1)}_{\ell+p+1,p}$ and $\boldsymbol{Q}^{(\ell+1)}_{\ell+p+1}$ come from the $(\ell+1)^{st}$ QR factorization during the execution of the method. Further details may be found in [11]. Denoting the solution of (13) by $\boldsymbol{y}^{(\ell)}_p$, the approximate solution of (1) using the ℓ-shifted gl-GMRES method is given by

$$\boldsymbol{X}^{(\ell)}_p = \boldsymbol{V}_{(\ell+p)k}\left(\boldsymbol{Q}^{(\ell)}_{\ell+p,p}\boldsymbol{y}^{(\ell)}_p \otimes \boldsymbol{I}_k\right).$$

Algorithm 4 describes the ℓ-shifted gl-GMRES method. Termination of the algorithm is achieved with the discrepancy principle.

2.3. Local BGMRES methods

We now turn our attention to the new local variants of BGMRES: lo-BGMRES and ℓ-shifted lo-BGMRES. As we demonstrate in Section 3, these methods can achieve higher accuracy than BGMRES, gl-GMRES and their ℓ-shifted variants when applied to the solution of linear discrete ill-posed problems.

The lo-BGMRES method determines an approximate solution to (1) by computing approximate solutions to the systems of equations

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}^\delta_j \quad \text{for} \quad j = 1, 2, \ldots, k. \tag{14}$$

For $j = 1, 2, \ldots, k$, let ${}^j\boldsymbol{x}_0 = \boldsymbol{0} \in \mathbb{R}^n$. Then the p^{th} iterate of the j^{th} subproblem, which is denoted by ${}^j\boldsymbol{x}_p$ and computed by lo-BGMRES, satisfies

$$\left\| \boldsymbol{A}\,{}^j\boldsymbol{x}_p - \boldsymbol{b}^\delta_j \right\| = \min_{\boldsymbol{x}\in{}^j\mathbb{K}_p(\boldsymbol{A},\boldsymbol{b}^\delta_j)} \left\| \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^\delta_j \right\|, \tag{15}$$

where ${}^j\mathbb{K}_p(\boldsymbol{A}, \boldsymbol{b}^\delta_j) = \text{span}\left\{\boldsymbol{b}^\delta_j, \boldsymbol{A}\boldsymbol{b}^\delta_j, \ldots, \boldsymbol{A}^{p-1}\boldsymbol{b}^\delta_j\right\}$ is the j^{th} Krylov subspace. Here and throughout this subsection $\|\cdot\|$ stands for the Euclidean vector norm. We assume that p is small enough so that $\dim({}^j\mathbb{K}_p(\boldsymbol{A}, \boldsymbol{b}^\delta_j)) = p$. For clarity,

Algorithm 4: ℓ-shifted gl-GMRES ($\ell \geq 1$) with discrepancy principle

Input: $\boldsymbol{A} \in \mathbb{R}^{n\times n}, \boldsymbol{B}^{\delta} \in \mathbb{R}^{n\times k}$, and $\ell \in \{1, 2, 3, \dots\}$
Output: $\boldsymbol{X}_p^{(\ell)} \in \mathbb{R}^{n\times k}$

Set $\boldsymbol{V}_1 = \boldsymbol{B}^{\delta}/\|\boldsymbol{B}^{\delta}\|$ & $\boldsymbol{X}_0^{(\ell)} = \boldsymbol{0}$;
Compute ℓ steps of global Arnoldi: $\boldsymbol{A}\boldsymbol{V}_{\ell k} = \boldsymbol{V}_{(\ell+1)k} * \boldsymbol{H}_{\ell+1,\ell}$;
for $p = 1, 2, \dots$ **do**
    Compute next global Arnoldi step:
    $\boldsymbol{A}\boldsymbol{V}_{(\ell+p)k} = \boldsymbol{V}_{(\ell+p+1)k} * \boldsymbol{H}_{\ell+p+1,\ell+p}$;
    Compute QR factorization: $\left[\boldsymbol{Q}_{p+1}^{(1)}, \boldsymbol{R}_{p+1,p}^{(1)}\right] = \boldsymbol{H}_{p+1,p}$;
    for $j = 1, 2, \dots, \ell$ **do**
        Compute QR factorization:
        $\left[\boldsymbol{Q}_{j+p+1}^{(j+1)}, \boldsymbol{R}_{j+p+1,p}^{(j+1)}\right] = \boldsymbol{H}_{j+p+1,j+p}\boldsymbol{Q}_{j+p,p}^{(j)}$;
    end
    Compute $\boldsymbol{y}_p^{(\ell)}$ as the minimizer of
    $\left\| \boldsymbol{R}_{\ell+p+1,p}^{(\ell+1)}\boldsymbol{y} - \|\boldsymbol{B}^{\delta}\| \left(\boldsymbol{Q}_{\ell+p+1}^{(\ell+1)}\right)^T \boldsymbol{e}_1 \right\|$;
    Compute $\boldsymbol{X}_p^{(\ell)} = \boldsymbol{V}_{(\ell+p)k}\left(\boldsymbol{Q}_{\ell+p,p}^{(\ell)}\boldsymbol{y}_p^{(\ell)} \otimes \boldsymbol{I}_k\right)$;
    Compute $\|\boldsymbol{r}_p\| = \left\|\boldsymbol{A}\boldsymbol{X}_p^{(\ell)} - \boldsymbol{B}^{\delta}\right\|$;
    if $\|\boldsymbol{r}_p\| \leq \tau\delta$ **then**
        Stop;
    end
end

we let the pre-superscript of a matrix in Algorithm 5 denote the index of the column being selected; elsewhere in this paper this superscript indicates that the j^{th} subproblem is being considered.

Differently from applying k independent instances of a modified Gram-Schmidt-based Arnoldi iteration, the computations of the local block Arnoldi process described by Algorithm 5 are initiated by

$$\boldsymbol{V}_1 = \left[\boldsymbol{b}_1^\delta/\|\boldsymbol{b}_1^\delta\|, \boldsymbol{b}_2^\delta/\|\boldsymbol{b}_2^\delta\|, \ldots, \boldsymbol{b}_k^\delta/\|\boldsymbol{b}_k^\delta\|\right].$$

The matrix $\boldsymbol{A}$ is then multiplied by $\boldsymbol{V}_i$ for $i = 1, 2, \ldots, p$ in each step of the iterative process. The inner j^{th} loop (lines $\mathbf{4} - \mathbf{11}$) of Algorithm 5 may be viewed as a modified Gram-Schmidt process for local orthogonalization and can be carried out in parallel.

Algorithm 5: Local block Arnoldi

Input: $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B}^\delta \in \mathbb{R}^{n\times k}$

Output: ${}^j\tilde{\boldsymbol{V}}_{p+1} \in \mathbb{R}^{n\times(p+1)}$ and ${}^j\boldsymbol{H}_{p+1,p} \in \mathbb{R}^{(p+1)\times p}$ for $j = 1, \ldots, k$

1 Set $\boldsymbol{V}_1 = \left[\boldsymbol{b}_1^\delta/\|\boldsymbol{b}_1^\delta\|, \boldsymbol{b}_2^\delta/\|\boldsymbol{b}_2^\delta\|, \ldots, \boldsymbol{b}_k^\delta/\|\boldsymbol{b}_k^\delta\|\right]$;
2 **for** $p = 1, 2, \ldots, n$ **do**
3 Compute $\boldsymbol{W}_p = \boldsymbol{A}\boldsymbol{V}_p$;
4 **for** $j = 1, 2, \ldots, k$ **do**
5 **for** $s = 1, 2, \ldots, p$ **do**
6 ${}^jh_{s,p} = \left({}^j\boldsymbol{V}_s\right)^T {}^j\boldsymbol{W}_p$;
7 ${}^j\boldsymbol{W}_p = {}^j\boldsymbol{W}_p - {}^jh_{s,p}\,{}^j\boldsymbol{V}_s$;
8 **end**
9 ${}^jh_{p+1,p} = \|{}^j\boldsymbol{W}_p\|$;
10 ${}^j\boldsymbol{V}_{p+1} = {}^j\boldsymbol{W}_p/{}^jh_{p+1,p}$;
11 **end**
12 **end**

The p^{th} step of the j^{th} local Arnoldi process computed by Algorithm 5 is given by

$$\boldsymbol{A}\,{}^j\tilde{\boldsymbol{V}}_p = {}^j\tilde{\boldsymbol{V}}_{p+1}\,{}^j\boldsymbol{H}_{p+1,p} \quad \text{for} \quad j = 1, 2, \ldots, k, \tag{16}$$

where the matrix ${}^j\tilde{\boldsymbol{V}}_{p+1} = [{}^j\tilde{\boldsymbol{v}}_1, {}^j\tilde{\boldsymbol{v}}_2, \ldots, {}^j\tilde{\boldsymbol{v}}_{p+1}] \in \mathbb{R}^{n\times(p+1)}$ has orthonormal columns with initial column ${}^j\tilde{\boldsymbol{v}}_1 = \boldsymbol{b}_j^\delta/\|\boldsymbol{b}_j^\delta\|$. The matrix ${}^j\tilde{\boldsymbol{V}}_p \in \mathbb{R}^{n\times p}$ is made up of the first p columns of ${}^j\boldsymbol{V}_{p+1}$ and ${}^j\boldsymbol{H}_{p+1,p} \in \mathbb{R}^{(p+1)\times p}$ is an upper Hessenberg matrix with positive subdiagonal entries. The columns of ${}^j\tilde{\boldsymbol{V}}_{p+1}$ span the j^{th} Krylov subspace ${}^j\mathbb{K}_{p+1}(\boldsymbol{A}, \boldsymbol{b}_j^\delta)$. The decomposition (16) is the foundation for the implementation of the lo-BGMRES and ℓ-shifted lo-BGMRES methods.

Post initialization, lo-BGMRES and ℓ-shifted lo-BGMRES may be viewed as block versions of GMRES and ℓ-shifted GMRES, respectively. Using the

notation defined so far, the minimization problem (15) may be rewritten as

$$\begin{aligned}\min_{\boldsymbol{x}\in {}^{j}\mathbb{K}_p(\boldsymbol{A},\boldsymbol{b}_j^\delta)} \left\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}_j^\delta\right\| &= \min_{\boldsymbol{y}\in\mathbb{R}^p} \left\|\boldsymbol{A}\,{}^{j}\tilde{\boldsymbol{V}}_p\boldsymbol{y}-\boldsymbol{b}_j^\delta\right\| \\ &= \min_{\boldsymbol{y}\in\mathbb{R}^p} \left\|{}^{j}\boldsymbol{R}_{p+1,p}\boldsymbol{y}-\|\boldsymbol{b}_j^\delta\|\left({}^{j}\boldsymbol{Q}_{p+1}\right)^T\boldsymbol{e}_1\right\|\end{aligned} \tag{17}$$

for $j=1,2,\ldots,k$, where $\boldsymbol{e_1}\in\mathbb{R}^{p+1}$ is the first unit axis vector. Here, we have used the relation (16) and the QR factorization ${}^{j}\boldsymbol{H}_{p+1,p} = {}^{j}\boldsymbol{Q}_{p+1}\,{}^{j}\boldsymbol{R}_{p+1,p}$. In equation (17), we also used the fact that the first column of ${}^{j}\tilde{\boldsymbol{V}}_{p+1}$ is $\boldsymbol{b}_j^\delta/\|\boldsymbol{b}_j^\delta\|$. Denote the minimizer of (17) by ${}^{j}\boldsymbol{y}_p$ for $j=1,2,\ldots,k$. Then the approximate solution of (14) may be written as ${}^{j}\boldsymbol{x}_p = {}^{j}\tilde{\boldsymbol{V}}_p\,{}^{j}\boldsymbol{y}_p$. Thus, the p^{th} approximate solution of (1) via lo-BGMRES may be written as $\boldsymbol{X}_p = [{}^{1}\boldsymbol{x}_p, {}^{2}\boldsymbol{x}_p, \ldots, {}^{k}\boldsymbol{x}_p]$. Algorithm 6 below summarizes the lo-BGMRES method and is terminated according to the discrepancy principle.

Algorithm 6: lo-BGMRES with discrepancy principle

Input: $\boldsymbol{A}\in\mathbb{R}^{n\times n}$ and $\boldsymbol{B}^\delta\in\mathbb{R}^{n\times k}$
Output: $\boldsymbol{X}_p\in\mathbb{R}^{n\times k}$

1 Set $\boldsymbol{V}_1 = \left[\boldsymbol{b}_1^\delta/\|\boldsymbol{b}_1^\delta\|, \boldsymbol{b}_2^\delta/\|\boldsymbol{b}_2^\delta\|, \ldots, \boldsymbol{b}_k^\delta/\|\boldsymbol{b}_k^\delta\|\right]$ & $\boldsymbol{X}_0 = \boldsymbol{0}$;
2 **for** $p=1,2,\ldots$ **do**
3 Compute next local block Arnoldi step:
  $\boldsymbol{A}\,{}^{j}\tilde{\boldsymbol{V}}_p = {}^{j}\tilde{\boldsymbol{V}}_{p+1}\,{}^{j}\boldsymbol{H}_{p+1,p}$ for $j=1,2,\ldots,k$;
4 **for** $j=1,2,\ldots,k$ **do**
5 Compute QR factorization: $\left[{}^{j}\boldsymbol{Q}_{p+1}, {}^{j}\boldsymbol{R}_{p+1,p}\right] = {}^{j}\boldsymbol{H}_{p+1,p}$;
6 Compute ${}^{j}\boldsymbol{y}_p$ as the minimizer of
  $\left\|{}^{j}\boldsymbol{R}_{p+1,p}\boldsymbol{y}-\|\boldsymbol{b}_j^\delta\|\left({}^{j}\boldsymbol{Q}_{p+1}\right)^T\boldsymbol{e}_1\right\|$;
7 Compute ${}^{j}\boldsymbol{x}_p = {}^{j}\tilde{\boldsymbol{V}}_p\,{}^{j}\boldsymbol{y}_p$ and update j^{th} column of $\boldsymbol{X}_p$ with ${}^{j}\boldsymbol{x}_p$;
8 Compute $\left\|{}^{j}\boldsymbol{r}_p\right\| = \left\|\boldsymbol{A}\,{}^{j}\boldsymbol{x}_p-\boldsymbol{b}_j^\delta\right\|$;
9 **if** $\left\|{}^{j}\boldsymbol{r}_p\right\| \le \tau\delta$ **then**
10 Stop;
11 **end**
12 **end**
13 **end**

We note that computation of the norm of the residual in line **8** of Algorithm 6 does not require an additional matrix-vector product evaluation with the matrix $\boldsymbol{A}$ (see Proposition 6.9 in [10]). This result was extended in [11] to ℓ-shifted GMRES, gl-GMRES, and BGMRES methods and therefore applies to this work. Additionally, we clarify that δ in line **9** corresponds to the norm of the error in

the j^{th} right-hand side vector $\boldsymbol{b}_j^\delta$ for $j = 1, 2, \ldots, k$, i.e.,

$$\left\| \boldsymbol{b}_j^\delta - \boldsymbol{b}_j \right\| \leq \delta.$$

The same details hold for Algorithm 7 that follows the exposition below.

Similarly to the lo-BGMRES method, the ℓ-shifted lo-BGMRES method determines an approximate solution to (1) by computing an approximate solution to the minimization problem (14). Letting ${}^j\boldsymbol{x}_0^{(\ell)} = \boldsymbol{0} \in \mathbb{R}^n$, the p^{th} iterate of the j^{th} subproblem ${}^j\boldsymbol{x}_p^{(\ell)}$ computed by ℓ-shifted lo-BGMRES satisfies

$$\left\| \boldsymbol{A}\, {}^j\boldsymbol{x}_p^{(\ell)} - \boldsymbol{b}_j^\delta \right\| = \min_{\boldsymbol{x} \in {}^j\mathbb{K}_p(\boldsymbol{A}, \boldsymbol{A}^\ell \boldsymbol{b}_j^\delta)} \left\| \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}_j^\delta \right\|, \tag{18}$$

where ${}^j\mathbb{K}_p(\boldsymbol{A}, \boldsymbol{A}^\ell \boldsymbol{b}_j^\delta) = \operatorname{span}\left\{ \boldsymbol{A}^\ell \boldsymbol{b}_j^\delta, \boldsymbol{A}^{\ell+1} \boldsymbol{b}_j^\delta, \ldots, \boldsymbol{A}^{\ell+p-1} \boldsymbol{b}_j^\delta \right\}$ denotes the j^{th} Krylov subspace. As in the previous sections, when $\ell = 0$ the problem (18) simplifies to (15). Additionally, we again assume that p is small enough so that $\dim({}^j\mathbb{K}_p(\boldsymbol{A}, \boldsymbol{b}_j^\delta)) = p$.

To utilize the ℓ-shifted lo-BGMRES method, the minimization problem (18) may be expressed as

$$\min_{\boldsymbol{y} \in \mathbb{R}^p} \left\| {}^j\boldsymbol{R}_{\ell+p+1,p}^{(\ell+1)} \boldsymbol{y} - \|\boldsymbol{b}_j^\delta\| \left({}^j\boldsymbol{Q}_{\ell+p+1}^{(\ell+1)} \right)^T \boldsymbol{e}_1 \right\| \tag{19}$$

for $j = 1, 2, \ldots, k$, where $\boldsymbol{e_1} \in \mathbb{R}^{\ell+p+1}$ is the first unit axis vector. Full details for the j^{th} subproblem can be found in [11, Section 2]. Denoting the minimizer of (19) by ${}^j\boldsymbol{y}_p^{(\ell)}$ for $j = 1, 2, \ldots, k$, the approximate solution of (14) may be written as ${}^j\boldsymbol{x}_p^{(\ell)} = {}^j\tilde{\boldsymbol{V}}_{\ell+p}\, {}^j\boldsymbol{Q}_{\ell+p,p}^{(\ell)}\, {}^j\boldsymbol{y}_p^{(\ell)}$. With this, the p^{th} approximate solution of (3) via the ℓ-shifted lo-BGMRES method is given by $\boldsymbol{X}_p^{(\ell)} = \left[{}^1\boldsymbol{x}_p^{(\ell)}, {}^2\boldsymbol{x}_p^{(\ell)}, \ldots, {}^k\boldsymbol{x}_p^{(\ell)} \right]$. Algorithm 7 below summarizes the ℓ-shifted lo-BGMRES method. The method is terminated according to the discrepancy principle.

3. Numerical examples

We illustrate the performance of the aforementioned block GMRES methods with several examples. As noted in the previous sections, all algorithms terminate the iterations according to the discrepancy principle. In detail, the iterations are terminated when the relative residual error is smaller than or equal to $\tau\delta$ with $\tau = 1.01$ and $\delta = v$ which is defined in the first example below. To evaluate the quality of the computed solutions, we compute the relative reconstructive error (RRE) defined by

$$\operatorname{RRE}\left(\boldsymbol{X}_p^{(\ell)} \right) = \frac{\left\| \boldsymbol{X}_p^{(\ell)} - \boldsymbol{X}^\dagger \right\|}{\left\| \boldsymbol{X}^\dagger \right\|} \quad \& \quad \operatorname{RRE}\left({}^j\boldsymbol{x}_p^{(\ell)} \right) = \frac{\left\| {}^j\boldsymbol{x}_p^{(\ell)} - \boldsymbol{x}^\dagger \right\|}{\|\boldsymbol{x}^\dagger\|},$$

Algorithm 7: ℓ-shifted lo-BGMRES ($\ell \geq 1$) with discrepancy principle

Input: $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B}^{\delta} \in \mathbb{R}^{n\times k}$
Output: $\boldsymbol{X}_p^{(\ell)} \in \mathbb{R}^{n\times k}$

Set $\boldsymbol{V}_1 = \left[\boldsymbol{b}_1^{\delta}/\|\boldsymbol{b}_1^{\delta}\|, \boldsymbol{b}_2^{\delta}/\|\boldsymbol{b}_2^{\delta}\|, \ldots, \boldsymbol{b}_k^{\delta}/\|\boldsymbol{b}_k^{\delta}\|\right]$ & $\boldsymbol{X}_0^{(\ell)} = \boldsymbol{0}$;
Compute ℓ steps of local block Arnoldi:
$\boldsymbol{A}\,{}^{j}\tilde{\boldsymbol{V}}_{\ell} = {}^{j}\tilde{\boldsymbol{V}}_{\ell+1}\,{}^{j}\boldsymbol{H}_{\ell+1,\ell}$ for $j = 1, 2, \ldots, k$;
for $p = 1, 2, \ldots$ **do**
  Compute next local block Arnoldi step:
  $\boldsymbol{A}\,{}^{j}\tilde{\boldsymbol{V}}_{\ell+p} = {}^{j}\tilde{\boldsymbol{V}}_{\ell+p+1}\,{}^{j}\boldsymbol{H}_{\ell+p+1,\ell+p}$ for $j = 1, 2, \ldots, k$;
  for $j = 1, 2, \ldots, k$ **do**
    Compute QR factorization: $\left[{}^{j}\boldsymbol{Q}_{p+1}^{(1)}, {}^{j}\boldsymbol{R}_{p+1,p}^{(1)}\right] = {}^{j}\boldsymbol{H}_{p+1,p}$;
    for $i = 1, 2, \ldots, \ell$ **do**
      Compute QR factorization:
      $\left[{}^{j}\boldsymbol{Q}_{i+p+1}^{(i+1)}, {}^{j}\boldsymbol{R}_{i+p+1,i+p}^{(i+1)}\right] = {}^{j}\boldsymbol{H}_{i+p+1,i+p}\,{}^{j}\boldsymbol{Q}_{i+p,p}^{(i)}$;
    end
    Compute ${}^{j}\boldsymbol{y}_p^{(\ell)}$ as the minimizer of
    $\left\| {}^{j}\boldsymbol{R}_{\ell+p+1,p}^{(\ell+1)}\boldsymbol{y} - \|\boldsymbol{b}_j^{\delta}\| \left({}^{j}\boldsymbol{Q}_{\ell+p+1}^{(\ell+1)}\right)^{T} \boldsymbol{e}_1 \right\|$;
    Compute ${}^{j}\boldsymbol{x}_p^{(\ell)} = {}^{j}\tilde{\boldsymbol{V}}_{\ell+p}\,{}^{j}\boldsymbol{Q}_{\ell+p,p}^{(\ell)}\,{}^{j}\boldsymbol{y}_p^{(\ell)}$ and update j^{th} column of $\boldsymbol{X}_p^{(\ell)}$ with ${}^{j}\boldsymbol{x}_p^{(\ell)}$;
    Compute $\left\|{}^{j}\boldsymbol{r}_p\right\| = \left\|\boldsymbol{A}\,{}^{j}\boldsymbol{x}_p^{(\ell)} - \boldsymbol{b}_j^{\delta}\right\|$;
    if $\left\|{}^{j}\boldsymbol{r}_p\right\| \leq \tau\delta$ **then**
      Stop;
    end
  end
end

where $\boldsymbol{X}_p^{(\ell)}$ and ${}^j\boldsymbol{x}_p^{(\ell)}$ denote the solutions determined by the appropriate algorithms at iteration p applied to the block linear and linear systems, respectively. Here, the Euclidean vector norm and the Frobenius matrix norm are used in their appropriate contexts. The exact solutions are denoted by $\boldsymbol{X}^\dagger$ and $\boldsymbol{x}^\dagger$ depending on the context. We refer to the RRE of the p^{th} iterate computed by the appropriate method terminated by the discrepancy principle as the breakout RRE value.

To provide a measure of the computational effort required by the block methods in our examples, we tabulate the number of iterations required by each method to terminate according to the discrepancy principle. The number of matrix-vector product evaluations with $\boldsymbol{A}$ may be computed by adding the shifted quantity ℓ to the tabulated number of iterations and then multiplying by the number of right-hand sides k for any shifted or unshifted method. While shifted variants do require additional matrix-block-vector or matrix-vector products with $\boldsymbol{A}$, they may yield computed approximate solutions with lower RRE values, as is illustrated in [11, 19, 15].

We consider three Fredholm integral equations of the first kind in one space-dimension and one linear PDE test problem in two-dimensions as our examples to facilitate our discussion of important considerations when applying block GMRES methods. Block linear systems of equations can arise when considering repeated experimental runs with a linear discrete ill-posed problem with $k = 1$. This is the situation we will focus on. We first consider when the norm of the error in the right-hand side block vector varies between columns and then when the number of right-hand sides varies. Finally, we will investigate the effect of the block size k and of the ill-posedness of a problem when applying the 0-shifted BGMRES method.

Our computational work was carried out in MATLAB R2020b on a MacBook Pro laptop running MacOS Catalina with an i5 Dual-Core Intel processor with @2.7 GHz and 8 GB of RAM. The computations were carried out with about 15 significant decimal digits.

(***Noise variation***) - Our first two examples consider the situation when the errors in the columns of the right-hand side of the block linear system varies by some known quantity. The first example will consider the Fredholm integral equation of the first kind discussed by Phillips in [20] and the second will focus on an inverse diffusion application utilizing the software package IR Tools [21]. Both examples contain the same number of right-hand sides and have the same noise contamination to compare the performance of the block methods when the singular spectra between examples differs. We comment on this further below.

The integral equation due to Phillips is defined by

$$\int_{-6}^{6} \kappa\left(\omega,\sigma\right)\, x\left(\sigma\right)\, d\sigma \;= b\left(\omega\right), \qquad -6 \leq \omega \leq 6, \tag{20}$$

whose solution, right-hand side, and kernel are given by

$$x(\sigma) = \begin{cases} 1 + \cos\left(\frac{\pi}{3}\sigma\right) & \text{if } |\sigma| < 3, \\ 0 & \text{otherwise,} \end{cases}$$
$$b(\omega) = (6 - |\omega|)\left(1 + \frac{1}{2}\cos\left(\frac{\pi}{3}\omega\right)\right) + \frac{9}{2\pi}\sin\left(\frac{\pi}{3}|\omega|\right),$$
$$\kappa(\omega, \sigma) = \boldsymbol{x}(\omega - \sigma),$$

respectively. The matrix $\boldsymbol{A}$ is obtained by discretizing the integral (20) using a Nyström method based on a composite trapezoidal quadrature rule with 1000 equidistant nodes [22]. This discretization gives a nonsymmetric matrix $\boldsymbol{A} \in \mathbb{R}^{1000\times 1000}$. Application of these equations are described in [20]. The exact solution $\boldsymbol{x}^\dagger$ of the discretized problem satisfies $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}^\dagger$, where $\boldsymbol{x}^\dagger, \boldsymbol{b} \in \mathbb{R}^{1000}$. A vector $\boldsymbol{e} \in \mathbb{R}^{1000}$ is formed with normally distributed random entries with zero mean to simulate noise so that $\boldsymbol{b}^{\boldsymbol{\delta}} = \boldsymbol{b} + \boldsymbol{e}$; it is scaled so as to correspond to a specific noise percentage level

$$v = 100\left(\frac{\|\boldsymbol{e}\|}{\|\boldsymbol{b}\|}\right).$$

We will refer to v as the noise level.

For both examples, we consider the situation when the amount of noise in 6 right-hand side vectors varies between 2-3%. For simplicity we add 2% noise to the first column of the block right-hand side vector and 3% noise to the 6^{th} column with increasing noise increments of 0.2% for the columns in-between. We assume that a fairly accurate error bound is known for each right-hand side so that the discrepancy principle may be used for each right-hand side in the lo-BGMRES methods. The gl-GMRES and BGMRES type methods may only use the larger 3% error bound for the entire block system so as to avoid propagating errors stemming from over-solving some of the right-hand sides with more noise.

With knowledge of the true solutions and the amount of noise in each column of the block linear system, we computed the mean, minimum, maximum, and standard deviation (SD) of the breakout RRE values for each block GMRES method considered using the 6 right-hand sides. These summary statistics along with the iteration count for each method considered are shown in Table 1 for the Phillips example. We note that the 1-shifted lo-BGMRES method performed the best with the smallest mean breakout RRE amongst the block GMRES methods considered. Figure 1 displays the true solutions of each right-hand side as well as the best solutions by each of the considered block methods with smallest mean breakout RRE values. The visually largest variance in the computed solutions is in the tails.

Our next example is an inverse diffusion problem provided by the software package IR Tools. The domain for the problem is $[0, T] \times [0, 1] \times [0, 1]$, and the solution u satisfies

$$\frac{\partial u}{\partial t} = \nabla^2 u \tag{21}$$

Method	Mean RRE	Min RRE	Max RRE	SD RRE	Iter.
0-shifted gl-GMRES	0.1171	0.1007	0.1352	0.0130	3
1-shifted gl-GMRES	0.0835	0.0820	0.0852	0.0012	3
2-shifted gl-GMRES	0.0989	0.0983	0.0995	0.0005	3
0-shifted BGMRES	2.1432	1.0494	3.3764	0.8674	2
1-shifted BGMRES	0.1609	0.1570	0.1640	0.0025	1
2-shifted BGMRES	0.0919	0.0905	0.0936	0.0011	1
0-shifted lo-BGMRES	0.1171	0.1007	0.1353	0.0130	3
1-shifted lo-BGMRES	0.0250	0.0235	0.0272	0.0013	4
2-shifted lo-BGMRES	0.0348	0.0258	0.0557	0.0121	4

Table 1: Results for the *noise variation* Phillips example comparing the accuracy of the block iterative methods for a block vector of size $k = 6$ with noise ranging from $2 - 3\%$. The mean, minimum, maximum, and standard deviation of the RRE values at breakout are provided. In the case of the lo-BGMRES methods, the maximum number of iterations among the right-hand sides is presented.

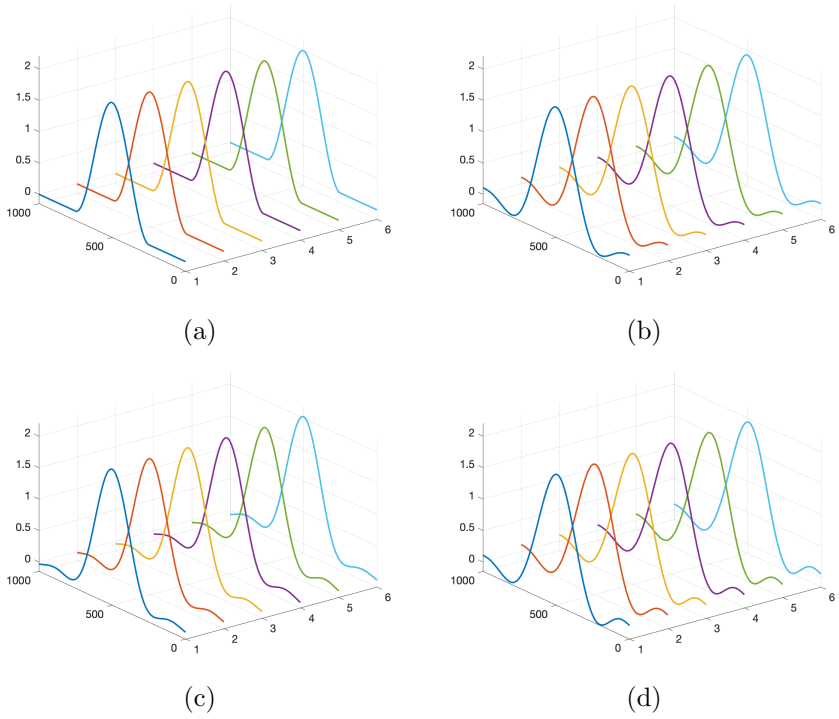


Figure 1: *Noise variation* Phillips example: The true solutions of the Phillips problem are shown in (a). The highest quality solutions in terms of lowest mean RRE values for (b) 1-shifted gl-GMRES, (c) 1-shifted lo-BGMRES, (d) 2-shifted BGMRES reconstructions at breakout.

with homogeneous Neumann boundary conditions and a smooth function u_0 as the initial condition at time $t = 0$. In the forward problem, u_0 is mapped to u_T at time $t = T$. The forward computation given by the function handle $\boldsymbol{A}$ is the numerical solution of the PDE (21). It is computed in IR Tools by the Crank-Nicolson-Galerkin finite element method. The smooth function u is discretized on a uniform finite element mesh. We set $T = 0.01$ and denote the solution of the discretized forward problem by $\boldsymbol{u}_T \in \mathbb{R}^N$. The so obtained noise-contaminated discrete solution is shown in Figure 2(b) where we have added 2% noise. The noise is not visible.

The inverse problem is to reconstruct the initial function $\boldsymbol{u}_0$ from $\boldsymbol{u}_T$. Further details regarding this kind of problems may be found in [21, 23]. In our example, the diffusion time was set to $T = 0.01$, the number of steps to 100, and the discretization carried out on a 64×64 grid which gives $N = 64^2$. The exact solution $\boldsymbol{x}^\dagger = \boldsymbol{u}_0$ of the discretized problem (21) satisfies $\boldsymbol{A}\boldsymbol{x}^\dagger = \boldsymbol{u}_T$ where we set $\boldsymbol{b} = \boldsymbol{u}_T$.

The discretized initial function $\boldsymbol{u}_0$ is displayed in Figure 2(a). Table 2 provides the summary statistics and iteration counts for the block methods applied to the inverse diffusion example. Similarly as for the Phillips problem, the 1-shifted lo-BGMRES method determines approximate solutions with the smallest mean breakout RRE value among the block GMRES methods considered. The solutions for the different right-hand sides computed by the 1-shifted lo-BGMRES method are shown in Figure 3.

The singular spectra of the matrices $\boldsymbol{A} \in \mathbb{R}^{1024\times 1024}$ from the Phillips and inverse diffusion problems are displayed in Figure 4. The two problems have drastically different singular spectra considering that the Phillips problem (green curve) contains no numerically zero singular values, while the inverse diffusion problem (magenta curve) contains several. It is noteworthy that both problems achieved the lowest mean breakout RRE value when using the 1-shifted lo-BGMRES method. This suggests that the robustness of the method may only be mildly affected by the severity of ill-posedness of the problem. We note that the size of the problems were modified from their original definition to make a comparison of their singular spectra possible.

(***Block size variation***) - We now turn to the situation when the block size is varied. Here, the noise level is held fixed at 1%, but the noise realization is different in each right-hand side. We consider block sizes $k = 3$, 10, and 50. For this example our problem is given by the Fredholm integral equation of the first kind discussed by Shaw in [24]. The integral equation is defined by

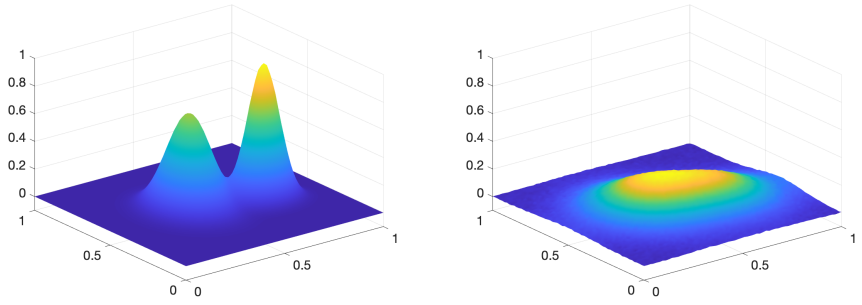
$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \kappa\,(\omega,\sigma)\; x\,(\sigma)\; d\sigma \;= b\,(\omega)\,, \qquad -\frac{\pi}{2} \le \omega \le \frac{\pi}{2}, \tag{22}$$

with kernel

$$\kappa\,(\omega,\sigma) = \Big(\cos\,(\sigma) + \cos\,(\omega)\Big)\left(\frac{\sin\,(\zeta)}{\zeta}\right)^2,$$

where

$$\zeta = \pi\big(\sin(\sigma) + \sin(\omega)\big).$$



(a) (b)

Figure 2: *Noise variation* inverse diffusion example: (a) true solution $\boldsymbol{u}_0$ ($N = 64^2$) and (b) the solution of the forward problem $\boldsymbol{u}_{T=0.01}$ with 2% added noise ($N = 64^2$).

Method	Mean RRE	Min RRE	Max RRE	SD RRE	Iter.
0-shifted gl-GMRES	0.4543	0.4209	0.4786	2.4×10^{-2}	6
1-shifted gl-GMRES	0.4140	0.4129	0.4151	8.8×10^{-4}	8
2-shifted gl-GMRES	0.4183	0.4175	0.4190	5.9×10^{-4}	11
0-shifted BGMRES	2.5521	2.3410	2.9763	2.4×10^{-1}	4
1-shifted BGMRES	0.3934	0.3916	0.3960	1.7×10^{-3}	5
2-shifted BGMRES	0.4515	0.4486	0.4540	2.2×10^{-3}	5
0-shifted lo-BGMRES	0.4698	0.4370	0.4941	2.0×10^{-2}	6
1-shifted lo-BGMRES	0.3888	0.3875	0.3903	1.2×10^{-3}	11
2-shifted lo-BGMRES	0.3956	0.3879	0.4091	1.0×10^{-2}	12

Table 2: Results for the *noise variation* inverse diffusion example comparing the accuracy of the block iterative methods for a block vector of size $k = 6$ with noise ranging from 2% to 3%. The mean, minimum, maximum, and standard deviation of the RRE values at breakout are provided. In the case of the lo-BGMRES methods, the maximum number of iterations among the right-hand sides is presented.

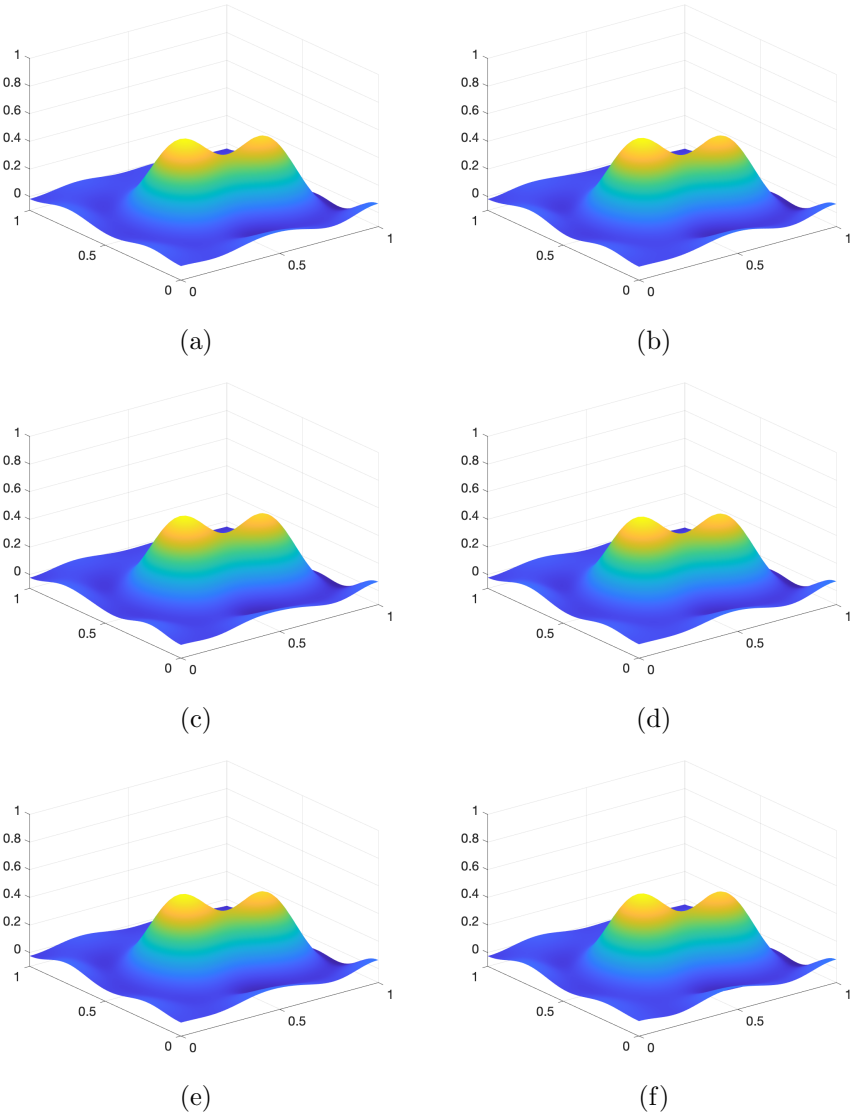


Figure 3: *Noise variation* inverse diffusion example: Lowest mean RRE approximate solutions determined by 1-shifted lo-BGMRES with noise levels (a) 2%, (b) 2.2%, (c) 2.4%, (d) 2.6%, (e) 2.8%, and (f) 3%.

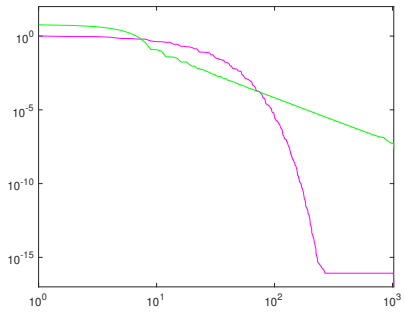


Figure 4: *Noise variation* example: Singular values of the system matrices of size 1024×1024 for the Phillips problem (green curve) and for the inverse diffusion problem (magenta curve).

The matrix $\boldsymbol{A}$ is again obtained by discretizing the integral (22) using the same Nyström method that was used to discretize equation (20). This gives a non-symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{1000\times 1000}$. The right-hand side function $b(\omega)$ is chosen so that the solution $x(\sigma)$ is the sum of two Gaussian functions. Applications of this equation are described in [24].

The basic summary statistics of the breakout RRE values and iteration counts for this example are shown in Table 3 for the methods considered. In the case when $k = 3$, the mean RRE values for all 2-shifted methods are competitive, with 2-shifted BGMRES performing the best. However, we found that as the number of right-hand sides increases to 10, the BGMRES methods perform poorly, while the lo-BGMRES and gl-GMRES variants still tend to be competitive. Finally, when $k = 50$ the mean RRE values of the BGMRES methods fail to be relevant, while the 2-shifted lo-BGMRES performs the best on average amongst the methods considered.

The iteration count for the BGMRES methods decreased as k was increased from 3 to 10. We have found this behavior to hold for various discrete ill-posed problems as the block size increases. Furthermore, we found (not shown here) that an additional forced iteration with the BGMRES methods is usually perilous to the RRE values, indicating that the discrepancy principle is a suitable termination criterion for the BGMRES methods.

(***Ill-conditioning***) - Our final example seeks to shed light on the poor performance of the BGMRES methods in the last example when $k \geq 3$. As we saw, the larger the block size, the larger the mean RRE values of the solutions computed by the BGMRES methods. We offer the explanation that as the block size increases, the conditioning of the block Hessenberg matrix $\boldsymbol{H}$ determined by BGMRES also increases.

To illustrate this observed behavior, we consider the Fredholm integral equation of the first kind

$$\int_0^{\pi} \kappa\left(\omega,\sigma\right)\, x\left(\sigma\right)\, d\sigma \;= b\left(\omega\right), \qquad 0 \leq \omega \leq \frac{\pi}{2}, \tag{23}$$

Method	Mean RRE	Min RRE	Max RRE	SD RRE	Iter.
		$k = 3$			
0-shifted gl-GMRES	0.1566	0.1561	0.1575	0.0007	6
1-shifted gl-GMRES	0.1216	0.1212	0.1223	0.0006	5
2-shifted gl-GMRES	0.0603	0.0529	0.0692	0.0083	7
0-shifted BGMRES	0.5603	0.4655	0.6772	0.1075	3
1-shifted BGMRES	0.2017	0.1574	0.2606	0.05314	2
2-shifted BGMRES	0.0543	0.0533	0.0554	0.0011	2
0-shifted lo-BGMRES	0.1789	0.1244	0.2652	0.0756	6
1-shifted lo-BGMRES	0.1215	0.1209	0.1223	0.0007	5
2-shifted lo-BGMRES	0.0926	0.0599	0.1484	0.0486	6
		$k = 10$			
0-shifted gl-GMRES	0.1278	0.1252	0.1312	0.0021	6
1-shifted gl-GMRES	0.1228	0.1204	0.1256	0.0017	5
2-shifted gl-GMRES	0.1448	0.1346	0.1552	0.0067	6
0-shifted BGMRES	15.370	1.9556	30.258	9.7273	2
1-shifted BGMRES	14.979	6.9506	26.310	7.1952	1
2-shifted BGMRES	0.3602	0.0896	0.6828	0.1994	1
0-shifted lo-BGMRES	0.2198	0.1226	0.3616	0.0994	5
1-shifted lo-BGMRES	0.1226	0.1199	0.1253	0.0018	5
2-shifted lo-BGMRES	0.1190	0.0562	0.1484	0.0397	5
		$k = 50$			
0-shifted gl-GMRES	0.1266	0.1121	0.1375	0.0052	6
1-shifted gl-GMRES	0.1228	0.1090	0.1345	0.0050	5
2-shifted gl-GMRES	0.1448	0.1327	0.1589	0.0073	6
0-shifted BGMRES	9.9×10^8	1.0×10^8	3.5×10^9	6.7×10^8	2
1-shifted BGMRES	1.3×10^9	1.2×10^8	4.7×10^9	8.1×10^8	1
2-shifted BGMRES	7.2×10^4	1.7×10^4	1.8×10^5	3.8×10^4	1
0-shifted lo-BGMRES	0.2593	0.1157	0.3774	0.0929	6
1-shifted lo-BGMRES	0.1315	0.1144	0.2175	0.0246	5
2-shifted lo-BGMRES	0.1149	0.0529	0.1619	0.0369	5

Table 3: Results for the *block size variation* example that compare the accuracy of the block iterative methods for block vector sizes $k = 3$, 10, and 50 with 1% noise. The mean, minimum, maximum, and standard deviation of the RRE values at breakout are shown. In the case of the ℓ-shifted lo-BGMRES methods, the maximum number of iterations among the right-hand sides is displayed.

with $\kappa(\omega,\sigma) = \exp\Big(\omega\cos(\sigma)\Big)$, $b(\omega) = 2\sinh(\omega)/\omega$, and $x(\sigma) = \sin(\sigma)$. This equation is discussed by Baart in [25]. The matrix $\boldsymbol{A}$ is obtained by discretizing the integral (23) using a Galerkin method with 1000 orthonormal box functions as test and trial functions; see [26]. This discretization gives a nonsymmetric matrix $\boldsymbol{A} \in \mathbb{R}^{1000\times 1000}$ and a scaled discrete approximation $\boldsymbol{x}^\dagger \in \mathbb{R}^{1000}$ such that $\boldsymbol{A}\boldsymbol{x}^\dagger = \boldsymbol{b}$ with $\boldsymbol{b} \in \mathbb{R}^{1000}$.

We will explore the conditioning of the upper block Hessenberg matrices $\boldsymbol{H}$ determined by the BGMRES methods using the three Fredholm integral problems from this section, where we refer to specific problems by the name of the authors: Phillips, Shaw, and Baart. Specifically, we determine the condition number of the final upper block Hessenberg matrix at breakout of the 0-shifted BGMRES method for all three problems for different block sizes.

The condition number of a general matrix $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ with $m \geq n$ is given by σ_1/σ_n, where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$ are the singular values of $\boldsymbol{M}$. For each of the problems will keep the noise level fixed at 1%, and use different realizations of the noise, similarly as in the previous examples.

Figure 5(a) shows the block size of the specified problem in color solved by 0-shifted BGMRES versus the condition number of the upper block-Hessenberg matrix upon termination of the method by the discrepancy principle. We recall that the mean RRE results from the previous example for Shaw when $k = 10$ or $k = 50$ were both poor, with $k = 50$ significantly worse. From this plot we notice that the computed condition numbers for the Shaw problem with $k = 10$ and $k = 50$ are approximately 10^5 and 10^{16}, respectively. The latter value is likely affected by the use of finite-dimensional arithmetic with about 15 decimal digits. The condition number for the matrices of the Baart problem increases rapidly with k; the increase is much slower for matrices associated with the Phillips problem. In general, the worse the conditioning of the matrix of a linear system of equations, the more sensitive is the computed solution to errors in the right-hand side.

The singular values of the matrices $\boldsymbol{A} \in \mathbb{R}^{1000\times 1000}$ for each problem are shown in Figure 5(b). One can see that Baart (blue curve) has only few singular values that are not numerically zero, Shaw (red curve) has more singular values that are not numerically zero, and Phillips (green curve) has no tiny singular value. The computed smallest singular value for Phillips is approximately 10^{-7}, while that for Baart and Shaw are 10^{-16}; the latter value is likely affected by the fact that arithmetic is carried out with about 15 significant decimal digits. The plots in Figure 5 suggest that severely ill-posed problems solved by BGMRES methods are vulnerable to ill-conditioning issues already when the block size is modest (4-10).

4. Conclusions

We considered several aspects of the application of block GMRES methods to the solution of linear discrete ill-posed problems when there is more than

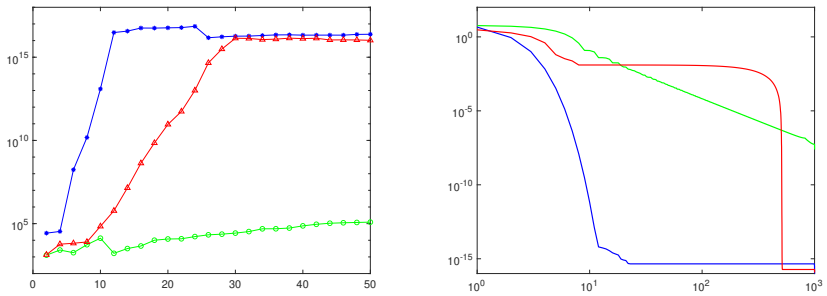


Figure 5: *Ill-conditioning* example: (a) Block size vs. condition number of the block upper-Hessenberg matrix for the Shaw (red triangles), Phillips (green circles), and Baart (blue stars) problems. (b) Singular values of the matrices of the Shaw (red curve), Phillips (green curve), and Baart (blue curve) problems.

one right-hand side. In particular, we explored the performance of block methods when the right-hand sides of a block linear system are contaminated by noise of various norm and when the block size is varied. We showed that our new lo-BGMRES methods can provide superior RRE compared to both the gl-GMRES and BGMRES type methods. Additionally, we illustrated that the condition number of the upper block-Hessenberg matrix of the 0-shifted BGMRES method is a function of the block size and the severity of the ill-posedness of the underlying problem.

Acknowledgments

The authors would like to thank the referees for comments.

Competing interests

The authors declare that they have no conflict of interest.

References

[1] H. W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer, Doordrecht, 1996.

[2] P. C. Hansen, Rank Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, SIAM, Philadelphia, 1998.

[3] C. Brezinski, M. Redivo-Zaglia, G. Rodriguez, S. Seatzu, Multi-parameter regularization techniques for ill-conditioned linear systems, Numer. Math. 94 (2003) 203–228.

[4] C. Brezinski, G. Rodriguez, S. Seatzu, Error estimates for linear systems with applications to regularization, Numer. Algorithms 49 (2008) 85–104.

[5] P. C. Hansen, T. K. Jensen, G. Rodriguez, An adaptive pruning algorithm for the discrete L-curve criterion, J. Comput. Appl. Math. 198 (2007) 483–492.

[6] S. Kindermann, Convergence analysis of minimization-based noise level-free parameter choice rules for linear ill-posed problems, Electron. Trans. Numer. Anal. 38 (2011) 233–257.

[7] S. Kindermann, K. Raik, A simplified L-curve method as error estimator, Electron. Trans. Numer. Anal. 53 (2020) 217–238.

[8] L. Reichel, G. Rodriguez, Old and new parameter choice rules for discrete ill-posed problems, Numer. Algorithms 63 (2013) 65–87.

[9] L. Reichel, G. Rodriguez, S. Seatzu, Error estimates for large-scale ill-posed problems, Numer. Algorithms 51 (2009) 341–361.

[10] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd. ed., SIAM, Philadelphia, 2003.

[11] A. Buccini, L. Onisk, L. Reichel, Range restricted iterative methods for linear discrete ill-posed problems, Electron. Trans. Numer. Anal. 58 (2023) 348–377.

[12] F. P. A. Beik, M. Najafi-Kalyani, L. Reichel, Iterative tikhonov regularization of tensor equations based on the arnoldi process and some of its generalizations, Appl. Numer. Math. 151 (2020) 425–447.

[13] L. Reichel, U. O. Ugwu, Tensor arnoldi-tikhonov and gmres-type methods for ill-posed problems with a t-product structure, J. Sci. Comput. 90 (2022) Art. 59.

[14] L. Elbouyahyaoui, A. Messaoudi, H. Sadok, Algebraic properties of the block gmres and block arnoldi methods, Electron. Trans. Numer. Anal. 33 (2009) 207–220.

[15] L. Dykes, L. Reichel, A family of range restricted iterative methods for linear discrete ill-posed problems, Dolomites Research Notes on Approximation 6 (2013) 27–36.

[16] K. Jbilou, A. Messaoudi, H. Sadok, Global fom and gmres algorithms for matrix equations, Appl. Numer. Math. 31 (1999) 49–63.

[17] K. Jbilou, H. Sadok, A. Tinzefte, Oblique projection methods for linear systems with multiple right-hand sides, Electron. Trans. Numer. Anal. 20 (2005) 119–138.

[18] B. Vital, Etude de quelques methodes de resolution de problemes lineaires de grande taille sur multiprocesseur, Ph.D. thesis, Univ. de Rennes, 1990.

[19] D. Calvetti, B. Lewis, L. Reichel, On the choice of subspace for iterative methods for linear discrete ill-posed problems, Int. J. Appl. Math. Comput. Sci. 11 (2001) 1069–1092.

[20] D. L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, J. ACM 9 (1962) 84–97.

[21] S. Gazzola, P. C. Hansen, J. G. Nagy, IR Tools: A MATLAB package of iterative regularization methods and large-scale test problems, Numerical Algorithms 81 (2019) 773–811.

[22] A. Neuman, L. Reichel, H. Sadok, Algorithms for range restricted iterative methods for linear discrete ill-posed problems, Numer. Algorithms 59 (2012) 325–331.

[23] T. Min, B. Geng, J. Ren, Inverse estimation of the initial condition for the heat equation, Intl. J. Pure Appl. Math. 82 (2013) 581–593.

[24] C. Shaw Jr., Improvements of the resolution of an instrument by numerical solution of an integral equation, J. Math. Anal. Appl. 37 (1972) 83–112.

[25] M. L. Baart, The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned least-squares problems, IMA J. Numer. Anal. 2 (1982) 241–247.

[26] P. C. Hansen, Regularization tools version 4.0 for Matlab 7.3, Numerical Algorithms 46 (2007) 189–194.