

DELFT UNIVERSITY OF TECHNOLOGY

REPORT 14-01

NESTED KRYLOV METHODS
FOR SHIFTED LINEAR SYSTEMS

M. BAUMANN AND M. B. VAN GIJZEN

ISSN 1389-6520

Reports of the Delft Institute of Applied Mathematics

Delft 2014

Copyright © 2014 by Delft Institute of Applied Mathematics, Delft, The Netherlands.

No part of the Journal may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission from Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands.

Summary

Shifted linear systems are of the form

$$(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}, \quad (1)$$

where $A \in \mathbb{C}^{N \times N}$, $\mathbf{b} \in \mathbb{C}^N$ and $\{\sigma_k\}_{k=1}^{N_\sigma} \in \mathbb{C}$ is a sequence of numbers, called *shifts*. In order to solve (1) for multiple shifts efficiently, shifted Krylov methods make use of the shift-invariance property of their respective Krylov subspaces, i.e.

$$\mathcal{K}_m(A, \mathbf{b}) = \mathcal{K}_m(A - \sigma I, \mathbf{b}), \quad \forall m \in \mathbb{N}, \forall \sigma \in \mathbb{C}, \quad (2)$$

and, therefore, compute one basis of the Krylov subspace (2) for all shifted systems. This leads to a significant speed-up of the numerical solution of the shifted problems because obtaining a basis of (2) is computationally expensive.

However, in practical applications, preconditioning of (1) is required, which in general destroys the shift-invariance. One of the few known preconditioners that lead to a new, preconditioned shifted problem is the so-called shift-and-invert preconditioner which is of the form $(A - \tau I)$ where τ is the *seed* shift. Most recently, multiple shift-and-invert preconditioners have been applied within a flexible GMRES iteration, cf. [10, 20]. Since even the one-time application of a shift-and-invert preconditioner can be computationally costly, polynomial preconditioners have been developed for shifted problems in [1]. They have the advantage of preserving the shift-invariance (2) and being computationally feasible at the same time.

The presented work is a new approach to the iterative solution of (1). We use nested Krylov methods that use an inner Krylov method as a preconditioner for an outer Krylov iteration, cf. [17, 24] for the unshifted case. In order to preserve the shift-invariance property, our algorithm only requires the inner Krylov method to produce collinear residuals of the shifted systems. The collinearity factor is then used in the (generalized) Hessenberg relation of the outer Krylov method. In this article, we present two possible combinations of nested Krylov algorithms, namely a combination of inner multi-shift FOM and outer multi-shift GMRES as well as inner multi-shift IDR(s) and outer multi-shift QMRIDR(s). However, we will point out that in principle every combination is possible as long as the inner Krylov method leads to collinear residuals. Since multi-shift IDR [31] does not lead to collinear residuals by default, the development of a collinear IDR variant that can be applied as an inner method within the nested framework is a second main contribution of this work.

The new nested Krylov algorithm has been tested on several shifted problems. In particular, the inhomogeneous and time-harmonic linear elastic wave equation leads to shifts that directly correspond to different frequencies of the waves.

Keywords: shifted linear systems, flexible preconditioning, inner-outer Krylov methods, Induced Dimension Reduction (IDR) method, time-harmonic elastic wave equation

Nested Krylov methods for shifted linear systems

M. Baumann and M. B. van Gijzen

May 31, 2014

Contents

1	Introduction	7
2	Multi-shift Krylov methods	7
2.1	Multi-shift GMRES	8
2.2	Multi-shift QMRIDR	9
3	Flexible preconditioning for multi-shift Krylov methods	10
3.1	The single shift-and-invert preconditioner for shifted systems	10
3.2	Flexible preconditioning for shifted linear systems	11
3.3	Krylov methods with collinear residuals	12
3.3.1	Collinear residuals in the Full Orthogonalization Method (FOM) . .	12
3.3.2	A variant of multi-shift IDR(s) with collinear residuals	14
3.4	Nested FOM-GMRES for shifted linear systems	18
3.5	Nested IDR-QMRIDR for shifted linear systems	19
4	Numerical experiments	21
4.1	A Helmholtz problem	22
4.1.1	Reflecting boundary conditions	22
4.1.2	Absorbing boundary conditions	24
4.2	The time-harmonic elastic wave equation	26
5	Conclusion	30
A	Numerical solution of the wedge problem	35
B	Numerical solution of the time-harmonic elastic wave equation	36

1 Introduction

We consider shifted linear systems of the form

$$(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}, \quad k = 1, \dots, N_\sigma, \quad (3)$$

where the dimensions are $A \in \mathbb{C}^{N \times N}$, $\mathbf{x}_k \in \mathbb{C}^N$, $\mathbf{b} \in \mathbb{C}^N$ and N_σ denotes the number of distinct shifts $\sigma_k \in \mathbb{C}$. For simplicity, we will often write

$$(A - \sigma I)\mathbf{x}^{(\sigma)} = \mathbf{b}, \quad \sigma \in \mathbb{C}, \quad (4)$$

keeping in mind that we aim to solve (4) for a sequence of many shifts σ and that quantities with a superscript belong to the respective shifted system, i.e. $\mathbf{x}^{(\sigma)}$ is the solution of the linear system with system matrix $(A - \sigma I)$ and right-hand side \mathbf{b} .

For an overview on the numerical solution of shifted linear systems using Krylov methods we refer to [12]. Multi-shift Krylov methods have been developed for QMR [6], GMRES(k) [8], FOM(k) [22], BiCGstab(ℓ) [7], CG [29] and, most recently, for IDR(s) [31].

This paper mostly focuses on preconditioning techniques for shifted linear systems. So far, known preconditioning techniques for shifted linear systems are the so-called shift-and-invert preconditioner of the form $(A - \tau I)$ where τ is a *seed* shift. This preconditioner has been applied to shifted Helmholtz problems for example in [4]. Since the shift-and-invert matrix has to be solved by a direct method, this approach can be computationally very costly. This can be overcome by either a multi-grid approach [5, 21] or by an approximation of the shift-and-invert preconditioner using a polynomial approximation as shown in [1]. Most recently, many shift-and-invert preconditioners have been applied in a flexible Krylov method in order to capture a wider range of shifts, cf. [10, 20]. In this article, we will present two new algorithms that are both nested Krylov methods in the sense that an inner Krylov method is used as a (polynomial) preconditioner to solve shifted linear systems in an outer Krylov iteration.

Our article is organized as follows: In Section 2 we repeat briefly the multi-shift GMRES [8] and the multi-shift QMRIDR [31] algorithm without preconditioning. Both are used as an outer Krylov method for the new nested framework in Section 3.4 and 3.5, respectively. As for the inner, collinear Krylov method, we present the multi-shift FOM algorithm that automatically leads to collinear residuals in Section 3.3.1. In order to use IDR as an inner method, we present a collinear variant of IDR(s) in Section 3.3.2. The article is summarized with various numerical tests in Section 4.

2 Multi-shift Krylov methods

The main property that is used in Krylov subspace methods for shifted linear systems, is the shift-invariance of the Krylov subspaces that are generated by the matrix A and the shifted matrix $(A - \sigma I)$ when the same right-hand side \mathbf{b} is used, i.e.

$$\mathcal{K}_m(A, \mathbf{b}) \equiv \text{span}\{\mathbf{b}, A\mathbf{b}, \dots, A^{m-1}\mathbf{b}\} = \mathcal{K}_m(A - \sigma I, \mathbf{b}) \quad \forall \sigma \in \mathbb{C}. \quad (5)$$

The immediate consequence of this invariance-property is that a basis for the underlying Krylov spaces only has to be build once for all shifted systems.

2.1 Multi-shift GMRES

The well-known GMRES method [19] has first been applied to shifted systems by [8]. In this section, we review the main ideas of [8] and point out how the shift-invariance property (5) is used in the algorithm in order to speed-up the computational performance when solving the shifted systems numerically.

In the classical GMRES method for the unshifted system $A\mathbf{x} = \mathbf{b}$, an orthogonal basis of the m -th Krylov subspace is obtained by the Arnoldi method. This leads to the well-known Hessenberg relation [28, eqn. 33.3],

$$AV_m = V_{m+1}\underline{H}_m, \quad (6)$$

where the columns of $V_m \in \mathbb{C}^{N \times m}$ span an orthonormal basis of $\mathcal{K}_m(A, \mathbf{b})$, and $\underline{H}_m \in \mathbb{C}^{(m+1) \times m}$ is the respective Hessenberg matrix with entries h_{ij} that are uniquely determined by the Arnoldi iteration. Then, in classical GMRES, an approximation to the solution of $A\mathbf{x} = \mathbf{b}$ in the m -th iteration is given by

$$\mathbf{x}_m = V_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y} \in \mathbb{C}^m}{\operatorname{argmin}} \|\underline{H}_m \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1\|, \quad (7)$$

with \mathbf{e}_1 being the first unit vector of \mathbb{C}^m . The optimization problem in (7) can be solved efficiently due to the Hessenberg-structure of \underline{H}_m using for instance Givens rotation, cf. [9, Section 5.1.8]. Clearly, we see from (7) that $\mathbf{x}_m \in \mathcal{K}_m(A, \mathbf{b})$.

Because of the shift-invariance property (5), the matrix V_m which spans the m -th Krylov subspace can be re-used for any shift σ . Therefore, the Arnoldi iteration in multi-shift GMRES only needs to be performed once, and from the shifted Hessenberg relation,

$$(A - \sigma I)V_m = V_{m+1}(\underline{H}_m - \sigma \underline{I}_m),$$

we can derive an approximated solution to the shifted problem (4) via,

$$\mathbf{x}_m^{(\sigma)} = V_m \mathbf{y}_m^{(\sigma)}, \quad \text{where } \mathbf{y}_m^{(\sigma)} = \underset{\mathbf{y} \in \mathbb{C}^m}{\operatorname{argmin}} \|(\underline{H}_m - \sigma \underline{I}_m) \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1\|, \quad (8)$$

where \underline{I}_m is the identity matrix of size $m \times m$ with an extra zero row appended at the bottom. We note that the matrix $\underline{H}_m(\sigma) \equiv \underline{H}_m - \sigma \underline{I}_m$ for the shifted system is of Hessenberg-structure as well. Clearly, we get the original Hessenberg matrix of the unshifted problem back if $\sigma = 0$, i.e. $\underline{H}_m(0) = \underline{H}_m$.

By comparing (7) and (8), we note that the m -th iterate is in both cases spanned by the column space of the matrix V_m and, therefore, lies in the same Krylov subspace $\mathcal{K}_m(A, \mathbf{b})$. Moreover, we note from the derivation of the shifted Hessenberg matrix $\underline{H}_m(\sigma)$ that the shift of the matrix A directly translates into a shift of the Hessenberg matrix.

In order to allow restarting for multi-shift GMRES, the authors of [8] require collinear residuals in order to preserve shift-invariance of the respective Krylov spaces after restarting, cf. [8, Algorithm 2.4]. So far, no preconditioner has been applied and we want to point to [17] for a detailed description of flexible preconditioning for the classical GMRES method.

2.2 Multi-shift QMRIDR

The QMRIDR method presented in [31] is a variant of the Induced Dimension Reduction (IDR) method [25] that makes use of a so-called generalized Hessenberg decomposition and determines the m -th iterate via a quasi-minimization of the m -th residual. In [11, 31], the following relation is derived:

$$AG_m U_m = G_{m+1} \underline{H}_m, \quad (9)$$

where $U_m \in \mathbb{C}^{m \times m}$ is upper triangular, $\underline{H}_m \in \mathbb{C}^{m+1 \times m}$ is of Hessenberg-form, and s consecutive vectors in G_m belong to the nested Sonneveld spaces $\mathcal{G}_0, \dots, \mathcal{G}_j$. The entries of U_m, G_m and \underline{H}_m are uniquely determined by the specific implementation of the IDR algorithm, cf. [11] for a detailed derivation. Based on the generalized Hessenberg decomposition (9), a multi-shift IDR method, the QMRIDR(s) algorithm, can be derived.

The approach of QMRIDR(s) is very similar to the GMRES approach. Therefore, the m -th iterate is constructed by a linear combination of the columns of G_m by putting $\mathbf{x}_m = G_m U_m \mathbf{y}_m$, with a coefficient vector $\mathbf{y}_m \in \mathbb{C}^m$ that is determined via a least-squares problem that involves the Hessenberg matrix \underline{H}_m only. Altogether, the following minimization problem needs to be solved,

$$\mathbf{x}_m = G_m U_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y} \in \mathbb{C}^m}{\operatorname{argmin}} \|\underline{H}_m \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1\|, \quad (10)$$

which is called 'quasi-minimization' of the m -th residual because the matrix G_m does not have orthogonal columns, cf. [31].

For shifted linear systems of the form (4), a very similar relation holds,

$$\mathbf{x}_m^{(\sigma)} = G_m U_m \mathbf{y}_m^{(\sigma)}, \quad \text{where } \mathbf{y}_m^{(\sigma)} = \underset{\mathbf{y} \in \mathbb{C}^m}{\operatorname{argmin}} \|(\underline{H}_m - \sigma \underline{U}_m) \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1\|, \quad (11)$$

and by comparing (10) with (11), we note that the approximate solution to the respective systems lie in the same subspace and the matrix $\underline{H}_m(\sigma) \equiv \underline{H}_m - \sigma \underline{U}_m$ is again of Hessenberg-structure because \underline{U}_m is an upper triangular matrix with an extra zero row.

From the derivations of multi-shift GMRES in Section 2.1 and of multi-shift QMRIDR in this section, we note that in both cases the efficient computation of the Hessenberg matrix of the shifted system $\underline{H}_m(\sigma)$ is crucial for the design of the algorithm. Therefore, we will put emphasis on the computation of $\underline{H}_m(\sigma)$ in the nested Krylov framework in Section 3.4 and Section 3.5 as well.

3 Flexible preconditioning for multi-shift Krylov methods

This section is structured as follows: We will first point out the requirements of a single preconditioner for shifted linear systems. In contrast to [26, 27], we restrict ourselves to such types of preconditioners that preserve the shift-invariance property of the Krylov subspaces in (5). Based on the requirements of a single preconditioner that preserves shift-invariance, a flexible preconditioner is designed that requires collinear residuals for the *inner* iteration. In Subsection 3.3, we present two Krylov methods that lead to collinear residuals; namely the FOM method that produces collinear residuals by construction and a variant of the IDR(s) method where some modifications are necessary in order to obtain collinear residuals. Both methods are used as preconditioners in a nested Krylov method in Subsection 3.4 and 3.5, respectively.

3.1 The single shift-and-invert preconditioner for shifted systems

In order to precondition a shifted linear system (4) without destroying the shift-invariance property of the respective Krylov spaces, we require the following equality after preconditioning,

$$\mathcal{K}_m((A - \sigma I)\mathcal{P}(\sigma)^{-1}, \mathbf{b}) = \mathcal{K}_m(A\mathcal{P}^{-1}, \mathbf{b}), \quad (12)$$

where $\mathcal{P}(\sigma)$ is a different preconditioner for every considered shifted system, and \mathcal{P} is a preconditioner for the unshifted system $A\mathbf{x} = \mathbf{b}$. Relation (12) is satisfied if we find a parameter η that depend on the shift σ and a constant matrix \mathcal{P} such that,

$$(A - \sigma I)\mathcal{P}(\sigma)^{-1} = A\mathcal{P}^{-1} - \eta(\sigma)I, \quad (13)$$

which in fact means that we can write the preconditioned shifted systems as shifted preconditioned systems with new shifts $\eta(\sigma)$. From [14, 15, 23], it is well-known that the so-called shift-and-invert preconditioner $\mathcal{P} \equiv (A - \tau I)$ applied to (13) leads to:

$$\begin{aligned} (A - \sigma I)\mathcal{P}(\sigma)^{-1} &= A(A - \tau I)^{-1} - \eta(\sigma)I \\ &= (1 - \eta(\sigma))\underbrace{\left(A + \frac{\tau\eta(\sigma)}{1 - \eta(\sigma)}I\right)}_{\equiv -\sigma}(A - \tau I)^{-1}. \end{aligned}$$

By choosing $\eta(\sigma)$ in an appropriate way, we can factor out the term $(A - \sigma I)$ on both sides, which yields the following formulas:

$$\eta(\sigma) = \frac{\sigma}{\sigma - \tau}, \quad \mathcal{P}(\sigma) = \frac{1}{1 - \eta(\sigma)}\mathcal{P} = \frac{\tau - \sigma}{\tau}\mathcal{P},$$

where the dependence of $\mathcal{P}(\sigma)$ on the shift becomes explicit. Therefore, only the seed shift τ has to be chosen and in order to invert $\mathcal{P}(\sigma)$ we only need to decompose \mathcal{P} . However,

the one-time decomposition of \mathcal{P} can numerically be very costly and the suitable choice of the seed shift τ is difficult for a large range of shifts $\sigma_1, \dots, \sigma_{N_\sigma}$, as has been pointed out by [20]. In [1], a polynomial preconditioner is suggested that cheaply approximates \mathcal{P}^{-1} .

3.2 Flexible preconditioning for shifted linear systems

Flexible preconditioning of an iterative Krylov subspace method means that a different preconditioner can be used in every iteration j , see [17] where flexible GMRES has been introduced for preconditioning of systems $A\mathbf{x} = \mathbf{b}$. For flexible preconditioning of shifted linear systems, we require a very similar relation to (13), namely,

$$(A - \sigma I)\mathcal{P}_j(\sigma)^{-1} = \alpha_j(\sigma)A\mathcal{P}_j^{-1} - \beta_j(\sigma)I, \quad (14)$$

where the parameters α_j and β_j depend on the shift, and different preconditioners $\mathcal{P}_j, \mathcal{P}_j(\sigma)$ are used at every iteration j . Note that the right-hand side in (14) is again a shifted linear system and, therefore, the shift-invariance is preserved by the flexible preconditioner. Since in a practical algorithm, the preconditioner is always directly applied to a vector \mathbf{v}_j , equation (14) becomes:

$$(A - \sigma I)\mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j = \alpha_j(\sigma)A\mathcal{P}_j^{-1}\mathbf{v}_j - \beta_j(\sigma)\mathbf{v}_j. \quad (15)$$

We will next determine how α_j and β_j have to be chosen such that (14) and (15) hold. We assume the preconditioning to be done by a multi-shift Krylov method itself (the *inner* method) which means that

$$\begin{aligned} \mathbf{z}_j &= \mathcal{P}_j^{-1}\mathbf{v}_j, \\ \mathbf{z}_j^{(\sigma)} &= \mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j, \end{aligned}$$

are computed via a truncated multi-shift Krylov method. More precisely, the vectors \mathbf{z}_j and $\mathbf{z}_j^{(\sigma)}$ denote the approximate (truncated) solution of the linear systems with system matrix A and $(A - \sigma I)$, respectively, and right-hand side \mathbf{v}_j . Therefore, the corresponding (inner) residuals are given by,

$$\mathbf{r}_j = \mathbf{v}_j - A\mathbf{z}_j = \mathbf{v}_j - A\mathcal{P}_j^{-1}\mathbf{v}_j, \quad (16)$$

$$\mathbf{r}_j^{(\sigma)} = \mathbf{v}_j - (A - \sigma I)\mathbf{z}_j^{(\sigma)} = \mathbf{v}_j - (A - \sigma I)\mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j. \quad (17)$$

We require the residuals (16)-(17) of the inner method to be collinear, i.e.

$$\exists \gamma_j^{(\sigma)} \in \mathbb{C} : \quad \gamma_j^{(\sigma)}\mathbf{r}_j = \mathbf{r}_j^{(\sigma)}. \quad (18)$$

Note that the collinearity factor $\gamma_j^{(\sigma)}$ will be different at every iteration j and for every shift σ . Moreover, relation (18) is not a very strong assumption since for example multi-shift FOM [22], multi-shift BiCGstab [7], restarted multi-shift GMRES [8] and multi-shift

IDR [13] lead to collinear residuals. With this assumption, α_j and β_j can be determined from (15) by using the collinearity relation (18) as the following calculation shows:

$$\begin{aligned}
& (A - \sigma I)\mathbf{z}_j^{(\sigma)} = \alpha_j A\mathbf{z}_j - \beta_j \mathbf{v}_j \\
\Leftrightarrow & \mathbf{v}_j - (A - \sigma I)\mathbf{z}_j^{(\sigma)} = \alpha_j \mathbf{v}_j - \alpha_j A\mathbf{z}_j + (1 - \alpha_j + \beta_j)\mathbf{v}_j \\
\Leftrightarrow & \mathbf{r}_j^{(\sigma)} = \underbrace{\alpha_j}_{\equiv \gamma_j^{(\sigma)}} \mathbf{r}_j + \underbrace{(1 - \alpha_j + \beta_j)}_{\equiv 0} \mathbf{v}_j
\end{aligned}$$

Thus, with (15) the residuals are collinear if we choose

$$\alpha_j = \gamma_j^{(\sigma)}, \quad \beta_j = \alpha_j - 1 = \gamma_j^{(\sigma)} - 1$$

in every (outer) iteration $1 \leq j \leq m$. We show the following relation,

$$(A - \sigma I)\mathbf{z}_j^{(\sigma)} = \gamma_j^{(\sigma)} A\mathbf{z}_j - (\gamma_j^{(\sigma)} - 1) \mathbf{v}_j, \quad (19)$$

or, in terms of the flexible preconditioners \mathcal{P}_j and $\mathcal{P}_j(\sigma)$, the following holds:

$$(A - \sigma I)\mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j = \left(\gamma_j^{(\sigma)} A\mathcal{P}_j^{-1} - (\gamma_j^{(\sigma)} - 1) I \right) \mathbf{v}_j, \quad 1 \leq j \leq m.$$

Note that the factors α_j and β_j do depend on σ since the collinearity factor $\gamma_j^{(\sigma)}$ is different for every shift.

3.3 Krylov methods with collinear residuals

In the previous section, we have derived the theoretical basis for a nested Krylov method for shifted linear systems. In order to be able to design a preconditioner that preserves the shift-invariance of the corresponding Krylov spaces, we assumed collinear residuals (18) for the inner multi-shift method. Next, we will present two multi-shift Krylov methods that lead to collinear residuals and, therefore, fulfill assumption (18). It is well-known that the multi-shift version of the Full Orthogonalization Method (FOM) leads to collinear residuals. We will review this result in Subsection 3.3.1. Moreover, we design a new variant of the Induced Dimension Reduction (IDR) method that has collinear residuals in Subsection 3.3.2.

3.3.1 Collinear residuals in the Full Orthogonalization Method (FOM)

The multi-shift FOM method (msFOM) can be derived very similarly to the multi-shift GMRES method in Section 2.1, cf. [22]. In FOM, an orthogonal basis of $\mathcal{K}_m(A, \mathbf{b})$ is obtained via the Arnoldi iteration which yields,

$$V_m^T A V_m = H_m,$$

and can be derived from (6) by left-multiplication with V_m^T . Here, H_m is a square matrix. Assuming H_m is invertible, the m -th iterate is then obtained by

$$\mathbf{x}_m = V_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = H_m^{-1}(\beta \mathbf{e}_1),$$

where, for simplicity, we assumed $\mathbf{x}_0 = \mathbf{0}$ as an initial guess, and $\beta \equiv \|\mathbf{r}_0\| = \|\mathbf{b}\|$. Similarly to multi-shift GMRES, the shifted Hessenberg matrix in **msFOM** can be derived as $H_m(\sigma) = H_m - \sigma I_m$, see [22]. The complete multi-shift FOM algorithm is giving in Algorithm 1.

Algorithm 1: **msFOM** for $(A - \sigma_k I) \mathbf{x}_k = \mathbf{b}$, $k = 1, \dots, N_\sigma$, [22]

```

1: // Single FOM:
2: Initialize  $\mathbf{r}_0 = \mathbf{b}$ ,  $\beta = \|\mathbf{r}_0\|$ ,  $\mathbf{v}_1 = \mathbf{r}_0/\beta$ 
3: for  $j = 1$  to  $m$  do
4:   Compute  $\mathbf{w} = A\mathbf{v}_j$ 
5:   for  $i = 1$  to  $j$  do           Arnoldi iteration
6:      $h_{i,j} = \langle \mathbf{w}, \mathbf{v}_i \rangle$ 
7:      $\mathbf{w} = \mathbf{w} - h_{i,j} \mathbf{v}_i$ 
8:   end for
9:   Set  $h_{j+1,j} = \|\mathbf{w}\|$  and  $\mathbf{v}_{j+1} = \mathbf{w}/h_{j+1,j}$ 
10: end for
11: Solve  $\mathbf{y}_m = H_m^{-1}(\beta \mathbf{e}_1)$ 
12: Set  $\mathbf{x}_m = V_m \mathbf{y}_m$ 
13:
14: // Multi-shift FOM:
15: for  $k = 1$  to  $N_\sigma$  do
16:   Construct  $H_m(\sigma_k) = H_m - \sigma_k I_m$ 
17:   Solve  $\mathbf{y}_m^{(\sigma_k)} = H_m(\sigma_k)^{-1}(\beta \mathbf{e}_1)$ 
18:   Set  $\mathbf{x}_m^{(\sigma_k)} = V_m \mathbf{y}_m^{(\sigma_k)}$ 
19: end for

```

It is well-known that the **msFOM** method as presented in Algorithm 1 leads to collinear residuals of the shifted system and the original ($\sigma = 0$) system. We repeat this result in the following Lemma.

Lemma 3.1 (Collinearity of the residuals in Algorithm 1). *We use the notation of Algorithm 1. Let the respective residuals of the original and the shifted system after m steps be*

$$\begin{aligned} \mathbf{r}_m &\equiv \mathbf{b} - A\mathbf{x}_m, \\ \mathbf{r}_m^{(\sigma)} &\equiv \mathbf{b} - (A - \sigma I)\mathbf{x}_m^{(\sigma)}. \end{aligned}$$

*Then there exists a scalar $\gamma^{(\sigma)}$ that depends on the number of performed **msFOM** iterations m , and the shift σ such that*

$$\gamma^{(\sigma)} \mathbf{r}_m = \mathbf{r}_m^{(\sigma)}.$$

Proof. This proof can be found in [18, Proposition 6.7] as well as in [22]. For the residual of the original system it holds after m iterations,

$$\begin{aligned}\mathbf{r}_m &= \mathbf{b} - A\mathbf{x}_m = \mathbf{b} - AV_m\mathbf{y}_m = \mathbf{r}_0 - AV_m\mathbf{y}_m \\ &= \underbrace{\beta\mathbf{v}_1 - V_m H_m \mathbf{y}_m}_{=0} - h_{m+1,m} \mathbf{e}_m^T \mathbf{y}_m \mathbf{v}_{m+1} = -h_{m+1,m} \mathbf{e}_m^T \mathbf{y}_m \mathbf{v}_{m+1}.\end{aligned}$$

Repeating the same calculation for the shifted system yields,

$$\mathbf{r}_m^{(\sigma)} = -h_{m+1,m}^{(\sigma)} \mathbf{e}_m^T \mathbf{y}_m^{(\sigma)} \mathbf{v}_{m+1},$$

with $\mathbf{e}_m \equiv [0, \dots, 0, 1]^T \in \mathbb{C}^m$, and the Arnoldi vector \mathbf{v}_{m+1} is identical to the unshifted case. Since the off-diagonal elements of the shifted Hessenberg matrix are identical to the unshifted Hessenberg matrix in Algorithm 1, and the orthogonal basis vectors \mathbf{v}_i obtained by the Arnoldi iteration are identical, we conclude,

$$\gamma^{(\sigma)} = y_m^{(\sigma)} / y_m, \quad (20)$$

with $y_m, y_m^{(\sigma)}$ being the last entry of the vectors $\mathbf{y}_m, \mathbf{y}_m^{(\sigma)}$, respectively. The residuals are collinear with the collinearity factor $\gamma^{(\sigma)}$ explicitly given by (20). \square

The above lemma shows that **msFOM** is suitable as a preconditioner for the nested framework derived in Section 3.2. The required collinearity factor $\gamma_j^{(\sigma)}$ of the j -th outer iteration in (19) is given by (20), where we assume that **msFOM** is stopped after m inner iterations

3.3.2 A variant of multi-shift IDR(s) with collinear residuals

The IDR method as presented in [25] is a Krylov subspace method that is based on the idea that the residual is forced to lie in spaces \mathcal{G}_j of shrinking dimensions as the number of iterations increases. In more detail, we require the residual \mathbf{r}_{n+1} to fulfill,

$$\mathcal{G}_{j+1} \ni \mathbf{r}_{n+1} = (I - \omega_{j+1}A)\mathbf{v}_n, \quad \text{with } \mathbf{v}_n \in \mathcal{G}_j \cap \mathcal{S}, \quad (21)$$

with $\mathcal{G}_0 \equiv \mathcal{K}_N(A, \mathbf{b})$, and, without loss of generality, let $\mathcal{S} = \mathcal{N}(P^H)$ be the null space of some $N \times s$ matrix $P = [\mathbf{p}_1, \dots, \mathbf{p}_s]$. It is known from the IDR theorem [25, Theorem 2.1] that the spaces \mathcal{G}_j that are generated via the recursive definition (21) are of decreasing dimension and that, eventually, $\mathcal{G}_j = \{\mathbf{0}\}$ for some $j \leq N$. This result guarantees that the residual will be equal to zero in exact arithmetic at some point. Moreover, note that the scalar ω_{j+1} in (21) can be chosen freely which we will exploit in the following in order to derive a collinear variant of the IDR(s) method.

According to [25], the vectors $\mathbf{v}_n \in \mathcal{G}_j \cap \mathcal{S}$ are computed in the following way,

$$\begin{aligned}(P^H \Delta R_n) \mathbf{c} &= P^H \mathbf{r}_n, \\ \mathbf{v}_n &= \mathbf{r}_n - \Delta R_n \mathbf{c},\end{aligned}$$

with $\Delta R_n \equiv [\Delta \mathbf{r}_{n-1}, \dots, \Delta \mathbf{r}_{n-s}]$, and residual difference $\Delta \mathbf{r}_{n-1} \equiv \mathbf{r}_n - \mathbf{r}_{n-1}$. In a similar way, we denote the matrix of the last $s+1$ residuals by $R_n \equiv [\mathbf{r}_n, \dots, \mathbf{r}_{n-s}]$.

A first approach to adapt the IDR method to shifted linear systems that leads to collinear residuals has been done by [13]. However, the algorithm presented in [13] has the disadvantage of evaluating a polynomial recursively. For the collinear IDR-variant presented in this section, we first note that collinear residuals \mathbf{r}_{n+1} and $\mathbf{r}_{n+1}^{(\sigma)}$ will lie in the same Sonneveld spaces. We therefore aim to construct the spaces \mathcal{G}_j only once for all shifted systems. Our approach can be described in two phases: First, we note from (21), that in order to obtain collinear residuals, we need to produce collinear vectors \mathbf{v}_n . Therefore, we want to calculate the intermediate vector $\mathbf{c}^{(\sigma)}$ of the shifted systems such that $\mathbf{v}_n^{(\sigma)} = \alpha \mathbf{v}_n$. The following calculation shows that we can even avoid storing residual differences of the shifted system,

$$\mathbf{v}_n^{(\sigma)} = \mathbf{r}_n^{(\sigma)} - \Delta R_n^{(\sigma)} \mathbf{c}^{(\sigma)} \quad (22)$$

$$= [\mathbf{r}_n^{(\sigma)}, -\Delta R_n^{(\sigma)}] \begin{bmatrix} 1 \\ \mathbf{c}^{(\sigma)} \end{bmatrix} \quad (23)$$

$$= R_n^{(\sigma)} D \begin{bmatrix} 1 \\ \mathbf{c}^{(\sigma)} \end{bmatrix} \quad (24)$$

$$= R_n \tilde{\Gamma}_n^{(\sigma)} D \begin{bmatrix} 1 \\ \mathbf{c}^{(\sigma)} \end{bmatrix} \quad (25)$$

$$= [\mathbf{r}_n, -\Delta R_n] D^{-1} \tilde{\Gamma}_n^{(\sigma)} D \begin{bmatrix} 1 \\ \mathbf{c}^{(\sigma)} \end{bmatrix}. \quad (26)$$

In step (24), we used that the residual updates can be expressed as a product of the actual residuals and a difference matrix D in the following way $[\mathbf{r}_n^{(\sigma)}, -\Delta R_n^{(\sigma)}] = R_n^{(\sigma)} D$, where D is defined as the difference matrix:

$$D \equiv \begin{bmatrix} 1 & -1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & -1 & \\ & & & & 1 \end{bmatrix} \in \mathbb{C}^{s+1 \times s+1}.$$

Then, in step (25), we can map the residuals of the shifted system to the residuals of the original system via a diagonal matrix $\tilde{\Gamma}_n^{(\sigma)} = \text{diag}(\gamma_n^{(\sigma)}, \dots, \gamma_{n-s}^{(\sigma)})$ that consists of the last $s+1$ consecutive collinearity factors. It holds, $R_n^{(\sigma)} = R_n \tilde{\Gamma}_n^{(\sigma)}$. The factor α can then be found by solving,

$$D^{-1} \tilde{\Gamma}_n^{(\sigma)} D \begin{bmatrix} 1 \\ \mathbf{c}^{(\sigma)} \end{bmatrix} = \begin{bmatrix} 1 \\ \mathbf{c} \end{bmatrix},$$

for $[1, \mathbf{c}^{(\sigma)}]^T$ and normalizing the first component. Note that this computation not only determines the factor α but also leads to the intermediate vector $\mathbf{c}^{(\sigma)}$ for the shifted system

without actually storing the residual differences for the shifted system. This is an advantage concerning memory requirements of the new multi-shift algorithm.

In the second step of our approach, we determine the free IDR parameter $\omega_{j+1}^{(\sigma)}$ and the factor $\gamma^{(\sigma)}$ such that the residuals are collinear, i.e.

$$\gamma^{(\sigma)} \mathbf{r}_{n+1} = \mathbf{r}_{n+1}^{(\sigma)}.$$

Therefore, we substitute the definition of the residuals from (21) and use the collinearity of the vector \mathbf{v}_n :

$$\begin{aligned} \gamma^{(\sigma)} \mathbf{r}_{n+1} &= \mathbf{r}_{n+1}^{(\sigma)} \\ \Leftrightarrow \gamma^{(\sigma)} ((I - \omega_{j+1} A) \mathbf{v}_n) &= (I - \omega_{j+1}^{(\sigma)} (A - \sigma I)) \alpha \mathbf{v}_n \\ \Leftrightarrow \gamma^{(\sigma)} \mathbf{v}_n - \gamma^{(\sigma)} \omega_{j+1} A \mathbf{v}_n &= (1 + \omega_{j+1}^{(\sigma)} \sigma) \alpha \mathbf{v}_n - \omega_{j+1}^{(\sigma)} \alpha A \mathbf{v}_n. \end{aligned}$$

By matching the coefficients of the terms that belong to \mathbf{v}_n and $A \mathbf{v}_n$, respectively, we obtain,

$$\gamma^{(\sigma)} = \alpha + \omega_{j+1}^{(\sigma)} \sigma \alpha, \quad \gamma^{(\sigma)} \omega_{j+1} = \omega_{j+1}^{(\sigma)} \alpha,$$

where $\gamma^{(\sigma)}$ and $\omega_{j+1}^{(\sigma)}$ can be calculated as:

$$\omega_{j+1}^{(\sigma)} = \frac{\omega_{j+1}}{1 - \omega_{j+1} \sigma}, \quad \gamma^{(\sigma)} = \frac{\alpha}{1 - \omega_{j+1} \sigma}. \quad (27)$$

Thus, we have derived a formula for the collinearity factor $\gamma^{(\sigma)}$ which can be used in the nested framework in (19). In Algorithm 2, we present the collinear IDR variant (called `collIDR(s)`) using (27) for the choice of the free IDR parameter $\omega_{j+1}^{(\sigma)}$. Note that in (27) the explicit dependence of the collinearity factor on the shift becomes obvious. Since in Algorithm 2 we loop over N_σ distinct shifts, all shifted quantities are labeled with a superscript that depends on this loop index k . Concerning the choice of the s -dimensional shadow space \mathcal{S} and the parameter ω_{j+1} of the unshifted system, we refer to Section 4.1 and 4.2 of [25], respectively.

Algorithm 2: collIDR(s) for $(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}$, $k = 1, \dots, N_\sigma$

```

1: Set  $\mathbf{r}_0 = \mathbf{b}$ ,  $\mathbf{x}_0 = \mathbf{x}_0^{(\sigma_k)} = 0$ , and  $\gamma_0^{(\sigma_k)} = 1$ 
2: for  $n = 0, \dots, s - 1$  do
3:    $\mathbf{v} = A\mathbf{r}_n$ ;  $\omega = (\mathbf{v}^H \mathbf{r}_n) / (\mathbf{v}^H \mathbf{v})$ 
4:    $\Delta \mathbf{x}_n = \omega \mathbf{r}_n$ ;  $\Delta \mathbf{r}_n = -\omega \mathbf{v}$ 
5:   for  $k = 1, \dots, N_\sigma$  do
6:      $\gamma_{n+1}^{(\sigma_k)} = \frac{\gamma_n^{(\sigma_k)}}{1 - \omega \sigma_k}$ ;  $\omega^{(\sigma_k)} = \frac{\omega}{1 - \omega \sigma_k}$ 
7:      $\Delta \mathbf{x}_n^{(\sigma_k)} = \omega^{(\sigma_k)} \gamma_n^{(\sigma_k)} \mathbf{r}_n$ 
8:      $\mathbf{x}_{n+1}^{(\sigma_k)} = \mathbf{x}_n^{(\sigma_k)} + \Delta \mathbf{x}_n^{(\sigma_k)}$ 
9:   end for
10:  Update:  $\mathbf{x}_{n+1} = \mathbf{x}_n + \Delta \mathbf{x}_n$ ;  $\mathbf{r}_{n+1} = \mathbf{r}_n + \Delta \mathbf{r}_n$ 
11: end for
12:  $\Delta R_{n+1} := (\Delta \mathbf{r}_n, \dots, \Delta \mathbf{r}_0)$ ;  $\Delta X_{n+1} := (\Delta \mathbf{x}_n, \dots, \Delta \mathbf{x}_0)$ 
13:  $\Delta X_{n+1}^{(\sigma_k)} := (\Delta \mathbf{x}_n^{(\sigma_k)}, \dots, \Delta \mathbf{x}_0^{(\sigma_k)})$ ;  $\boldsymbol{\gamma}^{(\sigma_k)} := (\gamma_{n+1}^{(\sigma_k)}, \dots, \gamma_0^{(\sigma_k)})$ 
14:  $n = s$ 
15: while  $\|\mathbf{r}_n\| > TOL$  and  $n < MAXIT$  do
16:   for  $j = 0, \dots, s$  do
17:     Solve  $\mathbf{c}$  from  $P^H \Delta R_n \mathbf{c} = P^H \mathbf{r}_n$ 
18:      $\mathbf{v} = \mathbf{r}_n - \Delta R_n \mathbf{c}$ 
19:     for  $k = 1, \dots, N_\sigma$  do
20:        $\tilde{\Gamma}_n^{(\sigma_k)} := \text{diag}(\boldsymbol{\gamma}^{(\sigma_k)})$ 
21:       Solve  $[1, \mathbf{c}^{(\sigma_k)}]$  from  $D^{-1} \tilde{\Gamma}_n^{(\sigma_k)} D [1, \mathbf{c}^{(\sigma_k)}] = [1, \mathbf{c}]$  such that  $\mathbf{v}^{(\sigma_k)} = \alpha^{(\sigma_k)} \mathbf{v}$ 
22:     end for
23:     if  $j == 0$  then
24:        $\mathbf{t} = A\mathbf{v}$ 
25:        $\omega = (\mathbf{t}^H \mathbf{v}) / (\mathbf{t}^H \mathbf{t})$ ;  $\omega^{(\sigma_k)} = \frac{\omega}{1 - \omega \sigma_k}$ 
26:        $\Delta \mathbf{x}_n = -\Delta X_n \mathbf{c} + \omega \mathbf{v}$ ;  $\Delta \mathbf{r}_n = -\Delta R_n \mathbf{c} - \omega \mathbf{t}$ 
27:     else
28:        $\Delta \mathbf{x}_n = -\Delta X_n \mathbf{c} + \omega \mathbf{v}$ ;  $\Delta \mathbf{r}_n = -A \Delta \mathbf{x}_n$ 
29:     end if
30:     Update:  $\mathbf{x}_{n+1} = \mathbf{x}_n + \Delta \mathbf{x}_n$ ;  $\mathbf{r}_{n+1} = \mathbf{r}_n + \Delta \mathbf{r}_n$ 
31:     for  $k = 1, \dots, N_\sigma$  do
32:        $\gamma_{n+1}^{(\sigma_k)} = \frac{\alpha^{(\sigma_k)}}{1 - \omega \sigma_k}$ 
33:        $\Delta \mathbf{x}_n^{(\sigma_k)} = -\Delta X_n^{(\sigma_k)} \mathbf{c}^{(\sigma_k)} + \omega^{(\sigma_k)} \alpha^{(\sigma_k)} \mathbf{v}$ 
34:        $\mathbf{x}_{n+1}^{(\sigma_k)} = \mathbf{x}_n^{(\sigma_k)} + \Delta \mathbf{x}_n^{(\sigma_k)}$ 
35:     end for
36:     // The IDR-update:
37:      $n = n + 1$ 
38:      $\Delta R_n := (\Delta \mathbf{r}_{n-1}, \dots, \Delta \mathbf{r}_{n-s})$ ;  $\Delta X_n := (\Delta \mathbf{x}_{n-1}, \dots, \Delta \mathbf{x}_{n-s})$ 
39:      $\Delta X_n^{(\sigma_k)} := (\Delta \mathbf{x}_{n-1}^{(\sigma_k)}, \dots, \Delta \mathbf{x}_{n-s}^{(\sigma_k)})$ ;  $\boldsymbol{\gamma}^{(\sigma_k)} := (\gamma_n^{(\sigma_k)}, \dots, \gamma_{n-s}^{(\sigma_k)})$ 
40:   end for
41: end while

```

3.4 Nested FOM-GMRES for shifted linear systems

We now present a special case of the nested Krylov framework of Section 3.2, namely a version where multi-shift FOM (**msFOM**) is used as inner preconditioner and flexible GMRES (**FGMRES**) is used as an outer Krylov iteration. Flexible GMRES has been introduced by [17] for unshifted systems $A\mathbf{x} = \mathbf{b}$ and provides a different preconditioner \mathcal{P}_j in the j -th outer iteration. The Hessenberg relation (6) therein extends to,

$$AZ_m = V_{m+1}\underline{H}_m, \quad (28)$$

where at step $1 \leq j \leq m$ the (flexible) preconditioning $\mathbf{z}_j \equiv \mathcal{P}_j^{-1}\mathbf{v}_j$ is carried out, and $Z_m \equiv [\mathbf{z}_1, \dots, \mathbf{z}_m]$. Note that (28) is formulated for the unshifted case and that one column of relation (28) yields,

$$A\mathbf{z}_j = V_{m+1}\underline{\mathbf{h}}_j,$$

which we will use next in order to determine the shifted Hessenberg matrix. Therefore, we assume that the preconditioner $\mathcal{P}_j(\sigma)$ is equivalent to a truncated **msFOM** inner iteration with collinear factor $\gamma_j^{(\sigma)}$ of the residuals according to (20). From the following calculation

$$\begin{aligned} (A - \sigma I)\mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\sigma)} \\ \Leftrightarrow \gamma_j^{(\sigma)}A\mathbf{z}_j - \left(\gamma_j^{(\sigma)} - 1\right)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\sigma)} \\ \Leftrightarrow \gamma_j^{(\sigma)}V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}\left(\gamma_j^{(\sigma)} - 1\right)\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\sigma)} \\ \Leftrightarrow V_{m+1}\left(\gamma_j^{(\sigma)}\underline{\mathbf{h}}_j - \left(\gamma_j^{(\sigma)} - 1\right)\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\sigma)} \end{aligned}$$

we can conclude the j -th column of the shifted Hessenberg matrix to be,

$$\underline{\mathbf{h}}_j^{(\sigma)} \equiv \gamma_j^{(\sigma)}\underline{\mathbf{h}}_j - \left(\gamma_j^{(\sigma)} - 1\right)\underline{\mathbf{e}}_j, \quad 1 \leq j \leq m, \quad (29)$$

with $\underline{\mathbf{e}}_j$ being the j -th unit vector of \mathbb{C}^{m+1} . Aligning the columns of m outer iterations together, yields the following formula for the shifted Hessenberg matrix,

$$\underline{H}_m(\sigma) = (\underline{H}_m - \underline{I}_m)\Gamma_m^{(\sigma)} + \underline{I}_m, \quad (30)$$

where \underline{I}_m is the $m \times m$ identity matrix with an extra column of zeros attached and $\Gamma_m^{(\sigma)}$ is a diagonal matrix with the collinearity factors on the diagonal, i.e.

$$\Gamma_m^{(\sigma)} \equiv \begin{bmatrix} \gamma_1^{(\sigma)} & & \\ & \ddots & \\ & & \gamma_m^{(\sigma)} \end{bmatrix} \in \mathbb{C}^{m \times m}. \quad (31)$$

We use this notation in order to point out the similarities to the nested algorithm in Section 3.5. Note that for $\sigma = 0$, the shifted Hessenberg matrix (30) reduces to the plain Hessenberg matrix, $\underline{\mathbf{H}}_m(0) = \underline{\mathbf{H}}_m$, because in this case the collinearity factors are all equal to one and, hence, $\Gamma_m^{(0)} = I$. The FOM-GMRES nested Krylov method for shifted linear systems is summarized in Algorithm 3.

Algorithm 3: FGMRES with a msFOM preconditioner for $(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}$, $k = 1, \dots, N_\sigma$

```

1: Initialize  $\mathbf{r}_0 = \mathbf{b}$ ,  $\beta = \|\mathbf{r}_0\|$ ,  $\mathbf{v}_1 = \mathbf{r}_0/\beta$ 
2: for  $j = 1$  to  $m$  do
3:   Preconditioning:  $\mathbf{z}_j^{(\sigma_k)} = \text{msFOM}(A - \sigma_k I, \mathbf{v}_j)$ 
4:   Compute  $\gamma_j^{(\sigma_k)}$  according to (20)
5:   Compute  $\mathbf{w} = A\mathbf{z}_j^{(0)}$ 
6:   for  $i = 1$  to  $j$  do Arnoldi iteration
7:      $h_{i,j} = \langle \mathbf{w}, \mathbf{v}_i \rangle$ 
8:      $\mathbf{w} = \mathbf{w} - h_{i,j}\mathbf{v}_i$ 
9:   end for
10:  Set  $h_{j+1,j} = \|\mathbf{w}\|$  and  $\mathbf{v}_{j+1} = \mathbf{w}/h_{j+1,j}$ 
11:  // Loop over shifted systems:
12:  for  $k = 1$  to  $N_\sigma$  do
13:    Define  $Z_j^{(\sigma_k)} = [\mathbf{z}_1^{(\sigma_k)}, \dots, \mathbf{z}_j^{(\sigma_k)}]$ 
14:    Construct  $\underline{\mathbf{H}}_j(\sigma_k)$  according to (30)
15:    Solve  $\mathbf{y}_j^{(\sigma_k)} = \text{argmin}_{\mathbf{y}} \|\beta\mathbf{e}_1 - \underline{\mathbf{H}}_j(\sigma_k)\mathbf{y}\|$ , with  $\mathbf{e}_1 = [1, 0, \dots, 0]^T \in \mathbb{R}^{j+1}$ 
16:    Set  $\mathbf{x}_j^{(\sigma_k)} = Z_j^{(\sigma_k)}\mathbf{y}_j^{(\sigma_k)}$ 
17:  end for
18: end for

```

We note that in the same way as in flexible GMRES for the unshifted case, extra storage is required because the matrices $Z_j^{(\sigma_k)}$ which span the solution space for every shifted problem need to be stored. This is in fact a drawback of flexible GMRES that has already been pointed out by [17] and applies here for every shift.

3.5 Nested IDR-QMRIDR for shifted linear systems

In a similar way to the nested FOM-GMRES algorithm, we present a nested IDR-QMRIDR method for shifted linear systems where `collIDR(s)` is used as inner method. For $V_m \equiv G_m U_m$, relation (9) in QMRIDR was given by,

$$AV_m = G_{m+1}\underline{\mathbf{H}}_m.$$

In flexible QMRIDR (FQMRIDR) which has been introduced in [31, Section 4], this relation is replaced by

$$AZ_m = G_{m+1}\underline{\mathbf{H}}_m, \tag{32}$$

with Z_m consisting of the columns $\mathbf{z}_j \equiv \mathcal{P}_j^{-1}\mathbf{v}_j$ for $1 \leq j \leq m$, just as before. One column of (32) reads,

$$A\mathbf{z}_j = G_{m+1}\underline{\mathbf{h}}_j,$$

which we will use in order to derive the shifted Hessenberg matrix.

We assume that the factor $\gamma_j^{(\sigma)}$ is given from (27) by an inner `collIDR(s)` iteration. Then, the flexible QMRIDR relation applied to a shifted problem yields:

$$\begin{aligned} & (A - \sigma I)\mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j = G_{m+1}\underline{\mathbf{h}}_j^{(\sigma)} \\ \Leftrightarrow & \quad \gamma_j^{(\sigma)}A\mathbf{z}_j - \left(\gamma_j^{(\sigma)} - 1\right)\mathbf{v}_j = G_{m+1}\underline{\mathbf{h}}_j^{(\sigma)} \\ \Leftrightarrow & \quad \gamma_j^{(\sigma)}G_{m+1}\underline{\mathbf{h}}_j - \left(\gamma_j^{(\sigma)} - 1\right)G_m\mathbf{u}_j = G_{m+1}\underline{\mathbf{h}}_j^{(\sigma)} \\ \Leftrightarrow & \quad G_{m+1}\left(\gamma_j^{(\sigma)}\underline{\mathbf{h}}_j - \left(\gamma_j^{(\sigma)} - 1\right)\underline{\mathbf{u}}_j\right) = G_{m+1}\underline{\mathbf{h}}_j^{(\sigma)} \end{aligned}$$

and one column of the shifted Hessenberg matrix is given by,

$$\underline{\mathbf{h}}_j^{(\sigma)} = \gamma_j^{(\sigma)}\underline{\mathbf{h}}_j - \left(\gamma_j^{(\sigma)} - 1\right)\underline{\mathbf{u}}_j, \quad 1 \leq j \leq m, \quad (33)$$

where $\underline{\mathbf{u}}_j \equiv [\mathbf{u}_j, 0]^T$ is the vector \mathbf{u}_j from the j -th iteration of QMRIDR with an extra zero.

Altogether, we have derived the shifted Hessenberg matrix,

$$\underline{\mathbf{H}}_m(\sigma) = (\underline{\mathbf{H}}_m - \underline{\mathbf{U}}_m)\Gamma_m^{(\sigma)} + \underline{\mathbf{U}}_m, \quad (34)$$

with $\Gamma_m^{(\sigma)}$ as defined in (31), and $\underline{\mathbf{U}}_m \equiv [\underline{\mathbf{u}}_1, \dots, \underline{\mathbf{u}}_m]$. Here we see the close relation between the two nested methods: By comparing the expression for the shifted Hessenberg matrices in (30) and (34) we first note that in principle every inner Krylov method can be used that provides collinear residuals. Moreover, we use this factor within the (generalized) Hessenberg relation of the outer Krylov method in a very similar way which shows that in principle also every Krylov method as an outer iteration is suitable.

Note that Algorithm 4 is schematic. For a more detailed and memory-efficient implementation of the flexible QMRIDR(s) routine, see [31, Algorithm 1]. In fact, we only need to apply the `collIDR(s)` routine as a preconditioner in line 18 and change line 32 by the formula (33) in order to adapt [31, Algorithm 1] to our nested method.

Algorithm 4: FQMRIDR with a collIDR(s) preconditioner for $(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}$, $k = 1, \dots, N_\sigma$

- 1: Initialize $\mathbf{r}_0 = \mathbf{b}$, $\beta = \|\mathbf{r}_0\|$, $\mathbf{v}_1 = \mathbf{r}_0/\beta$
 - 2: **for** $j = 1$ to m **do**
 - 3: Preconditioning: $\mathbf{z}_j^{(\sigma_k)} = \text{collIDR}(A - \sigma_k I, \mathbf{v}_j)$
 - 4: Compute $\gamma_j^{(\sigma_k)}$ according to (27)
 - 5: Compute $\underline{\mathbf{h}}_j, \underline{\mathbf{u}}_j$ as in QMRIDR, see [31, Algorithm 1]
 - 6: Compute $\underline{\mathbf{h}}_j^{(\sigma_k)}$ from (33)
 - 7: // Loop over shifted systems:
 - 8: **for** $k = 1$ to N_σ **do**
 - 9: Define $Z_j^{(\sigma_k)} = [\mathbf{z}_1^{(\sigma_k)}, \dots, \mathbf{z}_j^{(\sigma_k)}]$
 - 10: Construct $\underline{\mathbf{H}}_j(\sigma_k)$ according to (34)
 - 11: Solve $\mathbf{y}^{(\sigma_k)} = \text{argmin}_{\mathbf{y}} \|\beta \mathbf{e}_1 - \underline{\mathbf{H}}_j(\sigma_k) \mathbf{y}\|$, with $\mathbf{e}_1 = [1, 0, \dots, 0]^T \in \mathbb{R}^{j+1}$
 - 12: Set $\mathbf{x}_j^{(\sigma_k)} = Z_j^{(\sigma_k)} \mathbf{y}^{(\sigma_k)}$
 - 13: **end for**
 - 14: **end for**
-

4 Numerical experiments

The numerical examples we present are motivated in two respects: First, we consider examples taken from earlier publications on nested Krylov method and multi-shift Krylov methods, respectively. We will compare how the new nested framework compares to the results presented earlier. We consider the numerical solution of the Helmholtz equation in Section 4.1 and of the time-harmonic elastic wave equation in Section 4.2. In both cases, we consider the numerical solution at multiple frequencies that arise from a frequency-domain model of acoustic and elastic waves, respectively. We will point out that there exists a one-to-one relation between the considered shifts in (3) and the frequencies of the waves.

All examples have been implemented in MATLAB R2011B on a Intel(R) Xeon(R) CPU E3-1240 V2 @ 3.40GHz CPU.

For the numerical solution of shifted linear systems of the form

$$(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}, \quad k = 1, \dots, N_\sigma, \quad (35)$$

it is of practical use to re-formulate the problem (35) by the following substitution:

$$\begin{aligned} \tilde{\sigma}_k &\equiv \sigma_k - \sigma^*, \\ \tilde{A} &\equiv A - \sigma^* I, \end{aligned}$$

for some $\sigma^* \in \{\sigma_1, \dots, \sigma_{N_\sigma}\}$. This way, it is equivalent to solve the shifted linear systems

$$(\tilde{A} - \tilde{\sigma}_k I)\mathbf{x}_k = \mathbf{b}, \quad k = 1, \dots, N_\sigma, \quad (36)$$

with the only difference that in the formulation (36) the unshifted ($\tilde{\sigma}_k = 0$) solution corresponds to one of the N_σ solutions we are interested in.

4.1 A Helmholtz problem

As a first example, we consider acoustic wave propagation which can be modeled by the so-called Helmholtz equation,

$$-\Delta p - \left(\frac{2\pi f_k}{c(\mathbf{x})} \right)^2 p = s, \quad \text{in } \Omega \subset \mathbb{R}^2, \quad (37)$$

where p stands for the pressure and f_k is the k -th wave frequency. We consider the wedge problem which was introduced in [16]. Therein, the computational domain Ω is $\Omega = [0, 600] \times [0, 1000]$ and the underlying sound velocity $c(\mathbf{x})$ is heterogeneous as displayed in Figure 10, Appendix A. Moreover, the sound source is given by a point source at the top of the computational domain, $s = \delta(x - 300, z)$. We distinguish between reflecting boundary conditions (homogeneous Neumann boundary conditions),

$$\frac{\partial p}{\partial n} = 0, \quad \text{on } \partial\Omega, \quad (38)$$

and absorbing boundary conditions (so-called Sommerfeld radiation boundary conditions) of the form,

$$\frac{\partial p}{\partial n} - i \left(\frac{2\pi f_k}{c(\mathbf{x})} \right) p = 0, \quad \text{on } \partial\Omega, \quad (39)$$

where i is the imaginary unit. Preconditioning techniques for the Helmholtz problem in the single-shift case ($N_\sigma = 1$) are for instance discussed in [3, 21, 30].

4.1.1 Reflecting boundary conditions

When including homogeneous Neumann boundary conditions (38) in the discretization of (37), we derive a systems of the form,

$$(K - (2\pi f_k)^2 M)\mathbf{p} = \mathbf{s}, \quad (40)$$

where K is the (negative) discrete Laplacian, M is a diagonal mass matrix, and the unknown vector \mathbf{p} consists of discrete pressure values. Since M is diagonal, we can cheaply re-write (40) as a shifted linear system,

$$(A - \sigma_k I)\mathbf{p}_k = \mathbf{b}, \quad \sigma_k = (2\pi f_k)^2,$$

with $A = M^{-1}K$ and $\mathbf{b} = M^{-1}\mathbf{s}$. In the same way as in [31], we consider in the following $N_\sigma = 4$ different frequencies, $f_k = \{1, 2, 4, 8\}Hz$. In the considered setting, this requires a system size of almost 4,000 equations. Moreover, we applied a shift-and-invert preconditioner as described in Section 3.1 in order to speed-up convergence for all considered solvers, cf. Table 1 where the seed shift τ is specified.

First, we apply the multi-shift GMRES (see Section 2.1) and the multi-shift QMRIDR (see Section 2.2) algorithm without nested preconditioning to the problem. As shown in

Figure 1, the convergence behavior is quite similar for the different frequencies. Moreover, we note a phase of stagnation in the convergence which in general does not give rise to a promising application of nested methods.

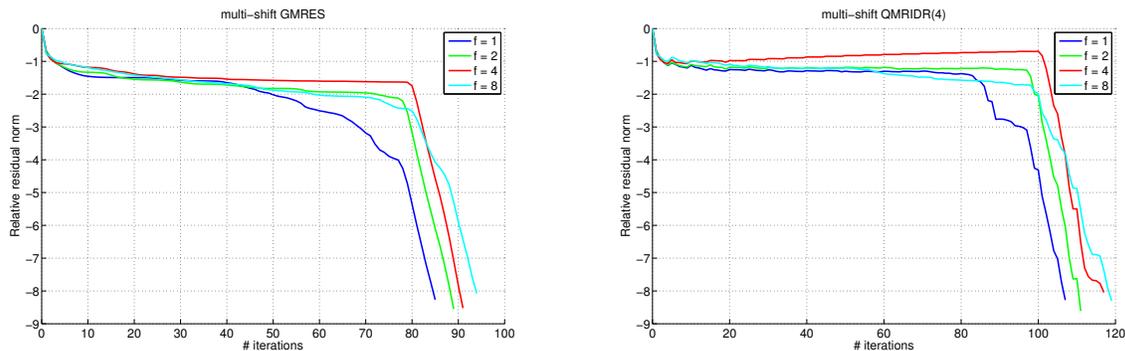


Figure 1: Convergence behavior of multi-shift GMRES (left) and multi-shift QMRIDR(4) (right) for (37)-(38) at four different frequencies.

Because of the slow convergence behavior in Figure 1, we conclude that a large number of inner iteration is required in the nested framework. Therefore, we applied the nested FOM-GMRES algorithm (Algorithm 3) with a relatively large number of inner iterations, cf. Table 1 and Figure 2. We observe from Figure 2 that once the inner FOM method converges below a relative tolerance of 0.1, the outer multi-shift GMRES algorithm converges within a few iterations. An optimal combination of inner-outer iterations has been evaluated and is presented in Table 1. We note that a computational speed-up of a factor two is obtained. Due to the stagnation phase of multi-shift GMRES, we were not able to obtain convergence for the restarted multi-shift GMRES version of [8, Algorithm 2.4].

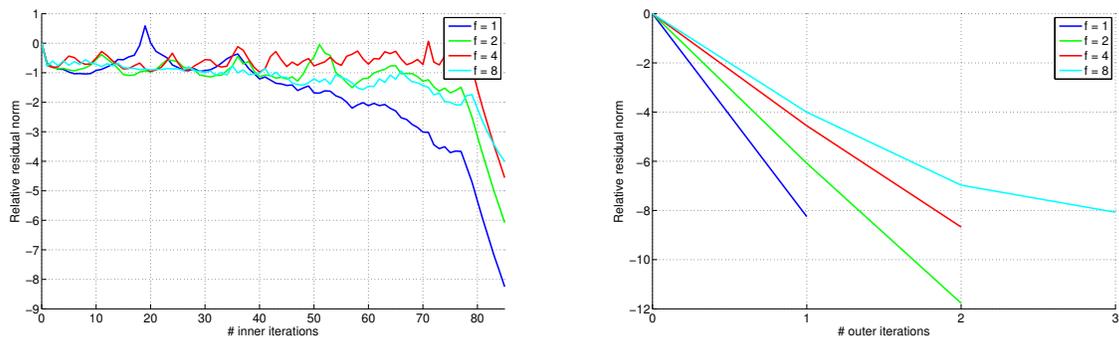


Figure 2: Convergence behavior of FOM-GMRES for (37)-(38): Typical inner convergence (left) and outer convergence (right).

	multi-shift Krylov		nested multi-shift Krylov	
	msGMRES	QMRIDR(4)	FOM-GMRES	IDR-QMRIDR(4)
# inner iterations	-	-	85	104
# outer iterations	94	119	3	4
seed shift τ	0.7-0.7i	0.7-0.7i	0.7-0.7i	0.7-0.7i
CPU time	3.24s	1.27s	1.62s	3.28s

Table 1: Comparison of multi-shift and nested multi-shift algorithms for the wedge problem with reflecting boundary conditions.

Moreover, we applied the collinear IDR variant as a flexible preconditioner to multi-shift QMRIDR (Algorithm 4). The results of Table 1 and Figure 3 show that again a large number of inner iterations is required until the inner Krylov method acts as a useful (flexible) preconditioner. This results again in a small number of required outer iterations but the overall time that is required does not show an improvement compared to the plain QMRIDR method.

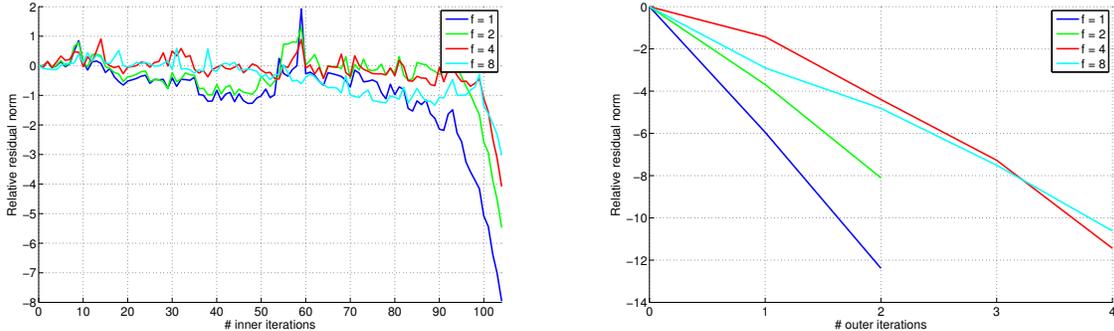


Figure 3: Convergence behavior of IDR-QMRIDR(4) for (37)-(38): Typical inner convergence (left) and outer convergence (right).

4.1.2 Absorbing boundary conditions

When non-homogeneous Neumann boundary conditions (39) are included in (37), we end up with a discretization of the form,

$$(K + i(2\pi f_k)C - (2\pi f_k)^2 M)\mathbf{p} = \mathbf{s}, \quad \sigma_k = 2\pi f_k, \quad (41)$$

where C represents the boundary conditions. Note that (41) is a quadratic eigenvalue problem in σ_k which according to [23] can be re-written as a shifted problem, see (45) in the next section for more details. In the present setting, we aim to solve a wider range of frequencies $f_k = \{1, 2, 4, 8, 16, 32\}Hz$. For the resolution of high frequencies and due to the doubling of unknowns in (45), the system size becomes more than 30,000 equations.

We present the convergence behavior for multi-shift GMRES and multi-shift QMRIDR(s) using only the single shift-and-invert preconditioner in Figure 4, respectively. The convergence curves are similar to the ones presented in Figure 1 in the sense that the residual norms first stagnate or even increase in the case of QMRIDR(s). However, the convergence curves are much more flat as soon as the residual norms decrease. Therefore, it is intuitively to truncate the inner iterations in this region.

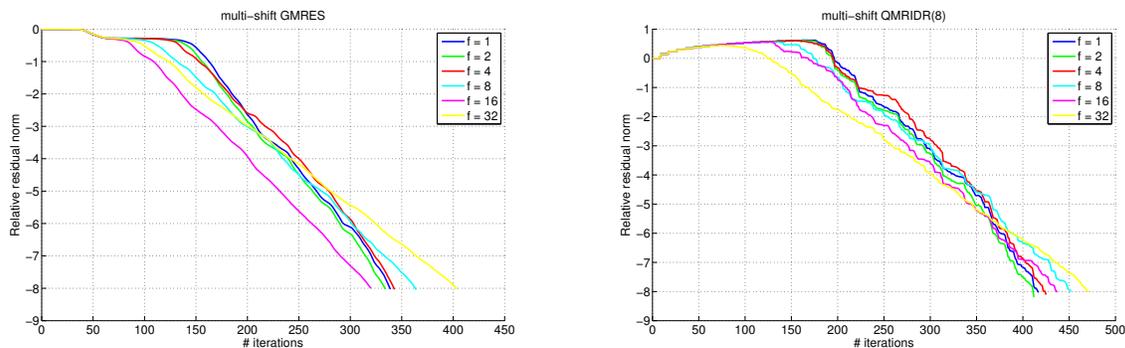


Figure 4: Convergence behavior of multi-shift GMRES (left) and multi-shift QMRIDR(8) (right) for (37)-(39).

The convergence plots of nested FOM-GMRES and nested IDR-QMRIDR(8) are presented in Figure 5 and Figure 6, respectively. In both cases, the number of inner iterations is chosen in such a way that the relative residual norms are of the size 0.1 or smaller which seems to be a good choice for truncation of the inner algorithm. The convergence rate of the outer Krylov method is in both cases very fast.

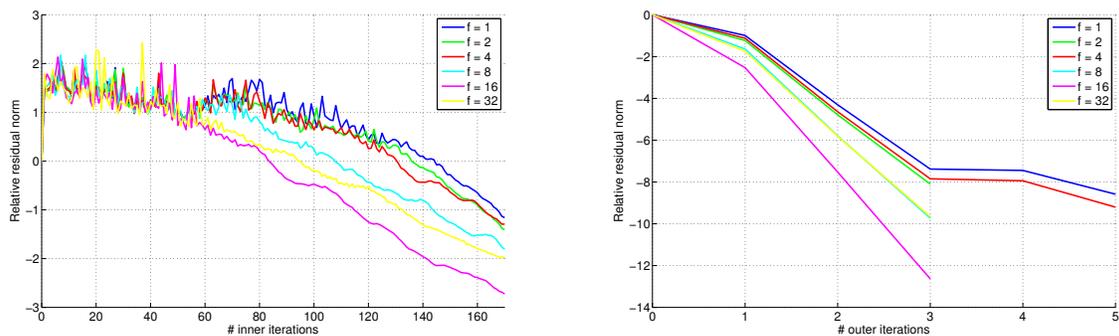


Figure 5: Convergence behavior of FOM-GMRES for (37)-(39): Typical inner convergence (left) and outer convergence (right).

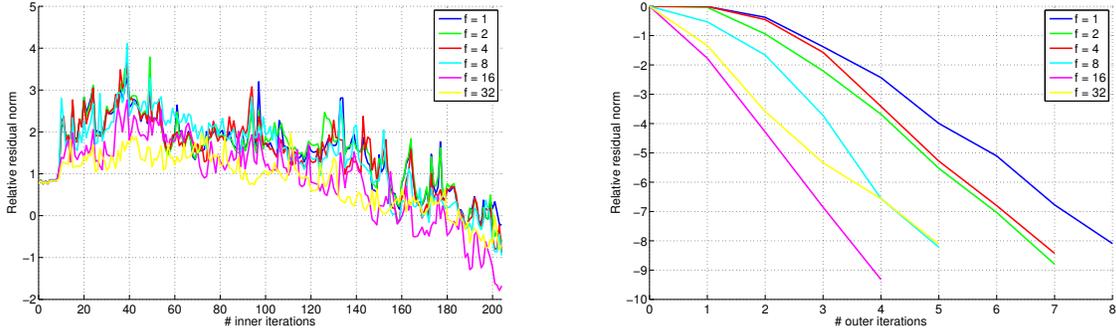


Figure 6: Convergence behavior of IDR-QMRIDR(8) for (37)-(39): Typical inner convergence (left) and outer convergence (right).

In Table 2, we want to point out the time that is required in order to solve all six shifted systems up to a relative tolerance of 10^{-8} . Comparing multi-shift GMRES and multi-shift GMRES preconditioned by a nested FOM method (called FOM-GMRES), we observe that the nested method is more than five times faster. Since the total number of iterations, and therefore, the number of matrix-vector multiplications is larger in the nested method, we conclude that the observed speed-up is due to shorter recurrence of the orthogonalization process. This also explains why we observe no speed-up for QMRIDR which is a short recurrence method by design.

	multi-shift Krylov		nested multi-shift Krylov	
	msGMRES	QMRIDR(8)	FOM-GMRES	IDR-QMRIDR(8)
# inner iterations	-	-	170	200
# outer iterations	404	471	5	8
seed shift τ	0.7-0.7i	0.7-0.7i	0.7-0.7i	0.7-0.7i
CPU time	193.69s	73.12s	36.66s	75.61s

Table 2: Comparison of multi-shift and nested multi-shift algorithms for the wedge problem with absorbing boundary conditions.

4.2 The time-harmonic elastic wave equation

Our second example considers the wave propagation of sound waves through an elastic medium. We are interested in the numerical solution of time-harmonic waves at multiple (angular) frequencies $\sigma_k = 2\pi f_k, k = 1, \dots, N_\sigma$. The scattering of time-harmonic waves is described in [2] by a Navier equation,

$$-\sigma_k^2 \rho(\mathbf{x}) \mathbf{u} - \nabla \cdot \tau(\mathbf{u}) = \mathbf{s}, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^2, \quad (42)$$

where $\mathbf{u} : \Omega \rightarrow \mathbb{R}^2$ is the unknown displacement vector, \mathbf{s} is typically a point source, and $\rho(\mathbf{x})$ is the density of the material which is assumed to be space-dependent. The strain

and stress tensor are derived from Hooke's law and are given by,

$$\tau(\mathbf{u}) \equiv \lambda(\mathbf{x}) (\nabla \cdot \mathbf{u}) + 2\mu(\mathbf{x})\epsilon(\mathbf{u}), \quad \epsilon(\mathbf{u}) \equiv \frac{1}{2} \left(\nabla \mathbf{u} + (\nabla \mathbf{u})^T \right).$$

Note that the underlying density $\rho(\mathbf{x})$ as well as the Lamé parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ have to be prescribed in the considered model, see Table 3.

In contrast to the example in Section 4.1, we will consider more realistic boundary conditions. Therefore, the following impedance boundary condition is prescribed,

$$i\gamma(\mathbf{x})\sigma_k\rho(\mathbf{x})B\mathbf{u} + \tau(\mathbf{u})\mathbf{n}(\mathbf{x}) = \mathbf{0}, \quad \mathbf{x} \in \partial\Omega, \quad (43)$$

where γ is the absorbency coefficient, $i \equiv \sqrt{-1}$, and $B_{i,j}(\mathbf{x}) \equiv c_p(\mathbf{x})n_in_j + c_s(\mathbf{x})t_it_j$. Here, n_i and t_i are the components of the (outer) normal vector \mathbf{n} and the tangential vector \mathbf{t} , respectively. For $\Omega \subset \mathbb{R}^2$ we consequently get a 2×2 matrix B for every boundary point $\mathbf{x} \in \partial\Omega$. The quantities c_p and c_s are the speed of pressure wave and shear wave, respectively (see Table 3). In the following, we prescribed absorbing boundary conditions on whole $\partial\Omega$ by setting $\gamma \equiv 1$.

ρ [kg/m^3]	c_p [m/s]	c_s [m/s]	λ [Pa]	μ [Pa]
$2.7 \cdot 10^3$	$6.8983 \cdot 10^3$	$4.3497 \cdot 10^3$	$2.6316 \cdot 10^{10}$	$5.1084 \cdot 10^{10}$

Table 3: Value of constant parameters taken from [2].

From a discretization of (42)-(43) using linear finite elements, we obtain the linear systems

$$(K + i\sigma_k C - \sigma_k^2 M)\mathbf{u} = \mathbf{s}, \quad (44)$$

with K, C, M being symmetric and sparse matrices, and \mathbf{u}, \mathbf{s} being the discretized counterpart of \mathbf{u}, \mathbf{s} in lexicographical order. Here, C contains the boundary conditions (43) and M is a diagonal mass matrix. Re-formulation of (44) yields

$$\left\{ \begin{bmatrix} iM^{-1}C & M^{-1}K \\ I & 0 \end{bmatrix} - \sigma_k \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \right\} \begin{bmatrix} \sigma_k \mathbf{u} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} M^{-1} \mathbf{s} \\ 0 \end{bmatrix}, \quad (45)$$

which is a standard step for the numerical treatment of quadratic eigenvalue problems, cf. [23]. Note that (45) is a block system of the form

$$(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}, \quad \sigma_k = 1, \dots, N_\sigma,$$

which is again a shifted linear system.

The considered numerical setting is described in [2, 14]. Therein, the parameters are set to the values presented in Table 3, and the unit square is considered as computational domain Ω . The angular frequencies σ_k are given by $\sigma_k = 2\pi f_k$ and range from $f_1 = 5,000Hz$ until $f_6 = 30,000Hz$ in uniform steps. For more details, see Appendix B.

We were again running our numerical experiments using an additional shift-and-invert preconditioner with seed shift τ as shown in Table 4. In Figure 7, we present the convergence curves of multi-shift GMRES and multi-shift QMRIDR(s) without nested preconditioning. In this experiments, we observe a flat convergence behavior that gives rise to an early truncation in the nested framework.

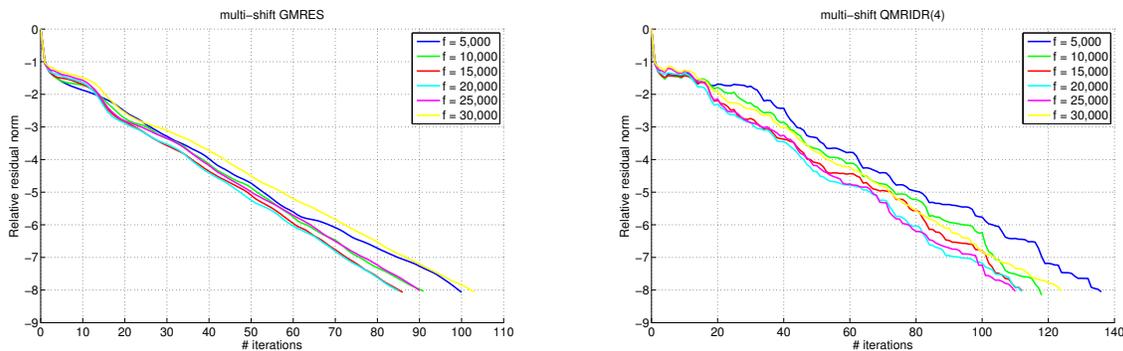


Figure 7: Convergence behavior of multi-shift GMRES (left) and multi-shift QMRIDR(4) (right) for (45).

For nested FOM-GMRES, we chose the number of inner FOM iterations such that the relative residual norms are below a threshold tolerance of 0.1, cf. Figure 8. Our numerical experiments have proven that this leads to a rapid convergence in only 7 iterations in the outer GMRES loop. When measuring the actual CPU time that is required to solve all $N_\sigma = 6$ shifted systems with multi-shift GMRES and nested FOM-GMRES, we observe a speed-up of almost two, cf. Table 4.

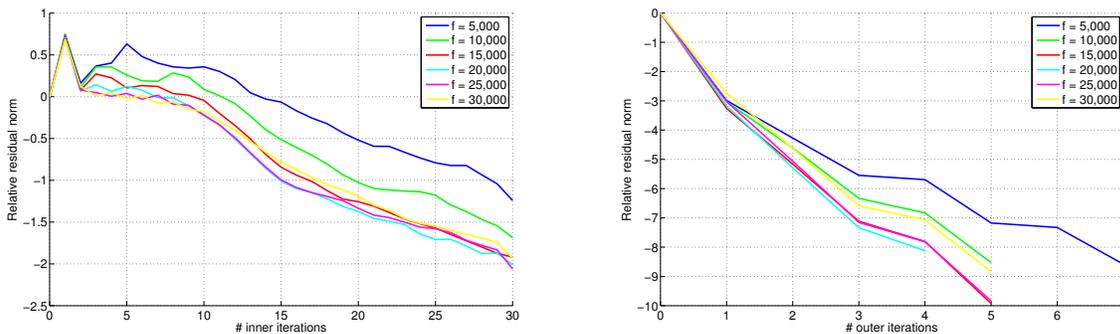


Figure 8: Convergence behavior of FOM-GMRES for (45): Typical inner convergence (left) and outer convergence (right).

	multi-shift Krylov methods			
	msGMRES	rest_msGMRES	QMRIDR(4)	collIDR(4)
# inner iterations	-	20	-	-
# outer iterations	103	7	136	134
seed shift τ	0.7-0.7i	0.7-0.7i	0.7-0.7i	0.7-0.7i
CPU time	17.71s	6.13s	22.35s	22.58s
	nested multi-shift Krylov methods			
	FOM-GMRES	IDR(4)-GMRES	FOM-QMRIDR(4)	IDR(4)-QMRIDR(4)
# inner iterations	30	25	30	30
# outer iterations	7	9	5	15
seed shift τ	0.7-0.7i	0.7-0.7i	0.7-0.7i	0.7-0.7i
CPU time	9.62s	42.19s	8.14s	58.36s

Table 4: Comparison of multi-shift and nested multi-shift algorithms for the linear elastic wave equation.

Moreover, we applied nested IDR-QMRIDR(s) to (45). From the convergence behavior of the inner IDR iteration (Algorithm 2), we note that the convergence curves show irregular jumps which makes a proper truncation of the inner preconditioner rather difficult. As in the previous tests, we do not observe a speed-up in CPU time for the nested algorithm which is mostly due to the fact of short recurrence of QMRIDR(s).

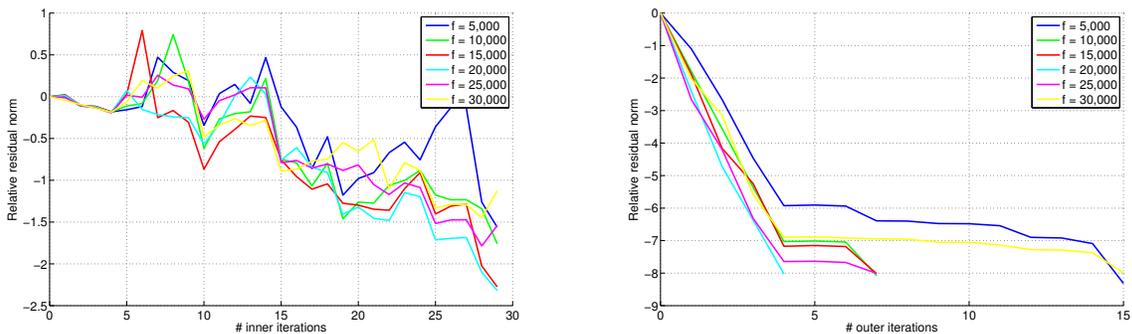


Figure 9: Convergence behavior of IDR-QMRIDR(4) for (45): Typical inner convergence (left) and outer convergence (right).

In Table 4 we present as well numerical results for an implementation of restarted multi-shift GMRES (`rest_msGMRES`) and our IDR-variant that exploits collinear residuals (`collIDR(s)`) which can be seen as a multi-shift Krylov method when being applied as a stand-alone algorithm as presented in Algorithm 2. Moreover, we compare performance of the nested Krylov methods with different inner-outer methods combined. For the specific setting considered in Table 4, we first note that QMRIDR(4) and `collIDR(4)` require similar CPU times. Moreover, we observe that a combination of inner `msFOM` and outer QMRIDR(4)

perform best among the nested algorithms. The restarted version of multi-shift GMRES performs best in this setting but did not converge in the examples considered in Section 4.1.

5 Conclusion

This work presents an algorithmic framework for the numerical solution of shifted linear systems (3) with inner-outer Krylov methods that allow flexible preconditioning. In this sense, it can be seen as a generalization of the work of [24] to sequences of shifted problems. The most general algorithm of this paper can be summarized in the following way,

1. the flexible preconditioner $\mathcal{P}_j(\sigma)$ is itself an *inner* multi-shift Krylov method which produces collinear residuals in the sense of (18),
2. the collinearity factor is used in the j -th iteration of an *outer* Krylov method in order to derive the Hessenberg matrix of the shifted systems, cf. (29), (33).

We call this new framework a *nested* Krylov method for shifted linear systems since the inner Krylov iteration is considered as a flexible preconditioner for the outer Krylov method.

This general framework has been illustrated and tested for two possible combinations of inner-outer Krylov methods. We present a combination of inner FOM and outer GMRES in Algorithm 3. Therefore, the collinearity factor for the inner Krylov method (multi-shift FOM) is given by (20) without any further manipulations. When combining multi-shift IDR(s) and QMRIDR(s) as presented in Algorithm 4, a new variant of IDR(s) has been developed which leads to collinear residuals with collinearity factor given by (27), cf. Algorithm 2. In both cases, a shifted Hessenberg matrix has been derived using the respective collinearity factors. This has been done for flexible GMRES (30) and flexible QMRIDR(s) (34), respectively.

Various numerical tests have been performed that showed an optimal performance of the nested algorithm when the inner Krylov method was truncated as the relative residuals satisfy $\|\mathbf{r}_j^{(\sigma_k)}\|/\|\mathbf{r}_0^{(\sigma_k)}\| < 10^{-1}$ at every (outer) iteration j and for all shifts $\sigma_1, \dots, \sigma_{N_\sigma}$. This way, we were able to obtain a computational speed-up up to a factor of five when comparing multi-shift GMRES to nested FOM-GMRES.

Acknowledgments

We want to thank René-Édouard Plessix from Shell Global Solutions International for helpful discussions.

References

- [1] M. I. Ahmad, D. B. Szyld, and M.B. van Gijzen. Preconditioned multishift BiCG for \mathcal{H}_2 -optimal model reduction. Technical report, Temple University, 2013.
- [2] T. Airaksinen, A. Pennanen, and J. Toivanen. A damping preconditioner for time-harmonic wave equations in fluid and elastic material. *Journal of Computational Physics*, 228(5):1466:1479, 2009.
- [3] Y. A. Erlangga, C. W. Oosterlee, and C. Vuik. A novel multigrid based preconditioner for heterogeneous Helmholtz problems. *SIAM J. Sci. Comput.*, 27:1471–1492, 2006.
- [4] Y.A. Erlangga, C. Vuik, and C.W. Oosterlee. On a class of preconditioners for solving the Helmholtz equation. *Appl. Num. Math.*, 50:409–425, 2004.
- [5] Y.A. Erlangga, C. Vuik, and C.W. Oosterlee. Comparison of multigrid and incomplete LU shifted-Laplace preconditioners for the inhomogeneous Helmholtz equation. *Applied Numerical Mathematics*, 56:648–666, 2006.
- [6] R. Freund. Solution of shifted linear systems by quasi-minimal residual iterations. *Numerical linear algebra, deGruyter, Berlin*, pages 101–121, 1993.
- [7] A. Frommer. BiCGStab(ℓ) for families of shifted linear systems. *Computing*, 70:87–109, 2003.
- [8] A. Frommer and U. Glässner. Restarted GMRES for Shifted Linear Systems. *SIAM J. Sci. Comput.*, 19(1):15–26, January 1998.
- [9] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 3rd edition, 1996.
- [10] G.-D. Gu, X.L. Zhou, and L. Lin. A flexible preconditioned Arnoldi method for shifted linear systems. *Journal of Computational Mathematics*, 25:522–530, 2007.
- [11] M. H. Gutknecht and J.-P. M. Zemke. Eigenvalue computations based on IDR. *SIAM J. Matrix Analysis Applications*, 34(2):283–311, 2013.
- [12] B. Jegerlehner. Krylov space solvers for shifted linear systems. *HEP-LAT*, hep-lat/9612014, 1996.
- [13] S. Kirchner. IDR-Verfahren zur Lösung von Familien geshifteter linearer Gleichungssysteme (in German). Master’s thesis, Bergische Universität Wuppertal, 2011.
- [14] K. Meerbergen. The solution of parametrized symmetric linear systems. *SIAM J. Matrix Anal. Appl.*, 24:1038:1059, 2003.

- [15] K. Meerbergen and Z. Bai. The Lanczos method for parameterized symmetric linear systems with multiple right-hand sides. *SIAM J. Matrix Anal. Appl.*, 4:1642–1662, 2010.
- [16] R. E. Plessix and W. A. Mulder. Separation-of-variables as a preconditioner for an iterative Helmholtz solver. *Appl. Num. Math.*, 44:383–400, 2003.
- [17] Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM J. Sci. Comput.*, 14:461–469, 1993.
- [18] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2nd edition, 2003.
- [19] Y. Saad and M. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7:856–869, 1986.
- [20] A. Saibaba, T. Bakhos, and P. Kitanidis. A flexible Krylov solver for shifted systems with application to oscillatory hydraulic tomography. *SIAM J. Sci. Comput.*, 35:3001–3023, 2013.
- [21] A. H. Sheikh, D. Lahaye, and C. Vuik. A scalable Helmholtz solver combining the shifted Laplace preconditioner with multigrid deflation. Technical report, DIAM report 11-01, 2011.
- [22] V. Simoncini. Restarted full orthogonalization method for shifted linear systems. *BIT Numerical Mathematics*, 43:459–466, 2003.
- [23] V. Simoncini and F. Perotti. On the numerical solution of $(\lambda^2 A + \lambda B + C)x = b$ and application to structural dynamics. *SIAM J. Sci. Comput.*, 23:1875:1897, 2002.
- [24] V. Simoncini and D. Szyld. Flexible inner-outer Krylov subspace methods. *SIAM J. Numer. Anal.*, 40:2219–2239, 2003.
- [25] P. Sonneveld and M. B. van Gijzen. IDR(s): A family of simple and fast algorithms for solving large nonsymmetric systems of linear equations. *SIAM J. Sci. Comput.*, 31:1035:1062, 2008.
- [26] K. Soodhalter. Two recursive GMRES-type methods for shifted linear systems with general preconditioning. *Review by Journal*, 2014.
- [27] K. Soodhalter, D. B. Szyld, and F. Xue. Krylov subspace recycling for sequences of shifted linear systems. *Applied Numerical Mathematics*, 81:105–118, 2014.
- [28] L. N. Trefethen and D. Bau. *Numerical Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 1997.
- [29] J. van den Eshof and G. L. G. Sleijpen. Accurate conjugate gradient methods for families of shifted systems. *Appl. Numer. Math.*, 49(1):17–37, April 2004.

- [30] M. B. van Gijzen, Y. A. Erlangga, and C. Vuik. Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian. *SIAM J. Sci. Comput.*, 29:1942–1958, 2007.
- [31] M. B. van Gijzen, G. L. G. Sleijpen, and J.-P. M. Zemke. Flexible and multi-shift Induced Dimension Reduction algorithm for solving large sparse linear systems. Technical report, DIAM, TU Delft, 2011.

Appendices

The focus of this work is not on the discretization of the numerical examples presented in Section 4.1 and Section 4.2, respectively. However, for the sake of completeness, we present the actual simulation results of these examples in the appendix.

A Numerical solution of the wedge problem

We present the numerical solution of problem (37) with absorbing boundary conditions (39) as discussed in Section 4.1.2 at a frequency of $32Hz$. The problem has been discretized with linear finite elements on a uniform triangulation of $\Omega = [0, 600] \times [0, 1000]$. We used a constant grid size of $h = 6.25m$ in both spatial dimensions. For the problem presented in Figure 10, this implies the solution of shifted linear systems roughly at the size of 30,000 unknowns. In Figure 10, we see that different layers for the sound velocity (left) lead to refraction of the wave which is emitted on the top of the domain (right).

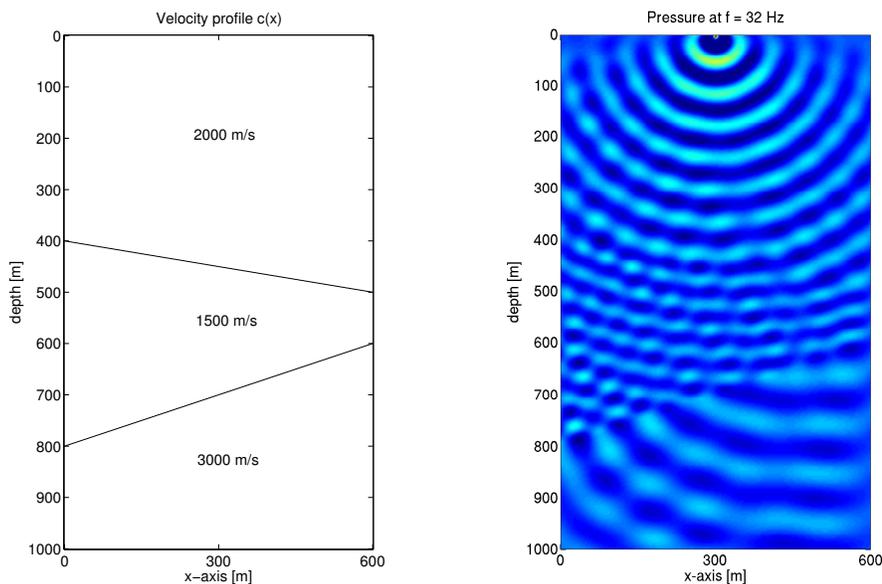


Figure 10: Left: sound velocity profile $c(\mathbf{x})$ of the wedge problem as described in [31]. Right: numerical solution of (37) with absorbing boundary conditions (39) at $f = 32Hz$.

B Numerical solution of the time-harmonic elastic wave equation

The linear elastic wave equation (42)-(43) has been solved on the unit square domain, $\Omega = [0, 1] \times [0, 1]$, using the parameter setting of Table 3 as suggested in [2]. We used linear finite elements for the discretization which leads to a straight-forward implementation of the boundary conditions (43). For the discretization, we used a triangulation of Ω with mesh points equidistantly spaced in both dimensions with mesh size h . For the resolution of the wave, we require at most 10 grid points per wavelength λ , i.e.

$$h < \frac{\lambda}{10}, \quad \text{where } \lambda = \min_{k \in \{1, \dots, N_\sigma\}} \left(\frac{c_s}{f_k} \right). \quad (46)$$

Here, we used $N_\sigma = 6$ different frequencies $f_k = \{5, \dots, 30\} \text{kHz}$ which leads to $h = 0.01$ to be a sufficiently small grid size according to (46). Hence, we use 10,201 grid points and, therefore, the system size for all shifted systems in (45) equals $N = 40,804$.

We used the results presented in [2] for the validation of our implementation when absorbing boundary conditions (43) and a point source in the middle of Ω has been prescribed, i.e. $\mathbf{s} = \delta(\mathbf{x} - (0.5, 0.5)^T)$. The numerical solution for the frequencies used in Section 4.2 are presented in Figure 11.

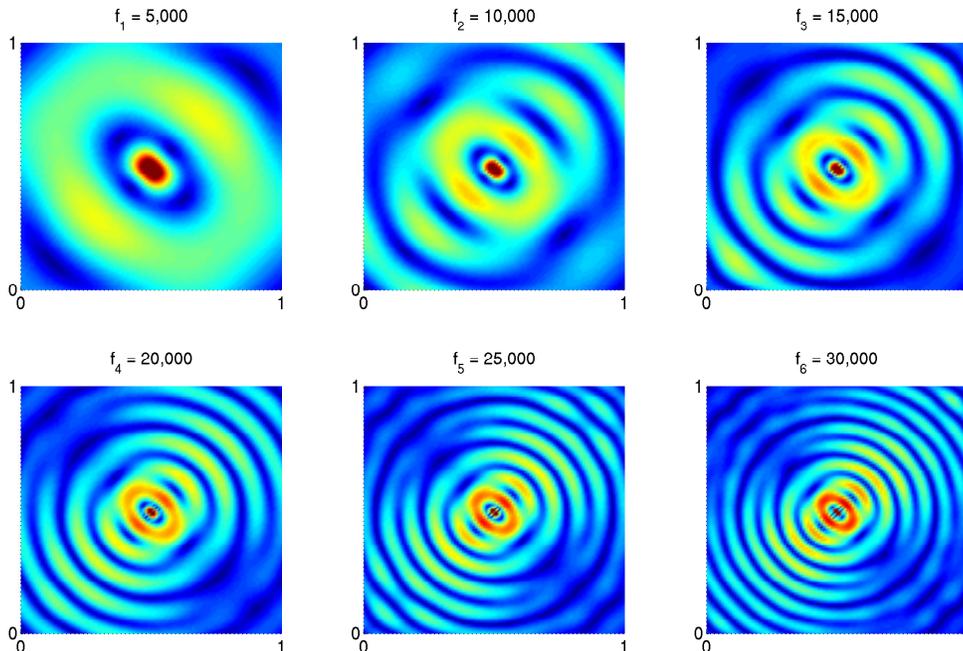


Figure 11: Numerical solution of the linear elastic wave equation (42) at multiple frequencies. We display the absolute value of the real part of $\mathbf{u}(\mathbf{x})$.