SIAM J. SCI. COMPUT.
Vol. 37, No. 5, pp. S90–S112

NESTED KRYLOV METHODS FOR SHIFTED LINEAR SYSTEMS*

MANUEL BAUMANN† AND MARTIN B. VAN GIJZEN†

Abstract. We consider Krylov subspace methods that are designed for sequences of shifted linear systems. For the efficient numerical solution of shifted problems, the shift-invariance property of the corresponding Krylov subspaces is used such that a Krylov basis is computed only once for all shifted systems. Preconditioners in general destroy this shift-invariance property. Known preconditioners that preserve the shift-invariance are the shift-and-invert preconditioner as well as polynomial preconditioners. In this work, we introduce a new approach to the preconditioning of multi-shift Krylov methods. In our new *nested* framework, we use an inner multi-shift Krylov method as a flexible preconditioner within an outer multi-shift Krylov method. In order to preserve the shift-invariance of the underlying Krylov subspaces, we require collinear residuals for the inner Krylov method. This new approach has been implemented for two possible combinations, namely, FOM-FGMRES and IDR-FQMRIDR, and has been tested for various numerical examples arising from geophysical applications.

Key words. Krylov subspace methods, shifted linear systems, flexible preconditioning, inner-outer Krylov methods, induced dimension reduction (IDR) method, time-harmonic wave equation

AMS subject classifications. 65F10, 65F50, 65F08

DOI. 10.1137/140979927

1. Introduction.

We consider shifted linear systems with equal right-hand sides of the form

$$(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}, \quad k = 1, \ldots, N_\sigma, \tag{1.1}$$

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

$$(A - \sigma I)\mathbf{x}^{(\sigma)} = \mathbf{b}, \quad \sigma \in \mathbb{C}, \tag{1.2}$$

keeping in mind that we aim to solve (1.2) 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 early overview of the numerical solution of shifted linear systems using Krylov methods we refer to [15, 30]. Multi-shift variants exist for many Krylov methods, including QMR [9], GMRES(k) [11], FOM(k) [27], BiCGstab(ℓ) [10], CG [35], MINRES [16], and, more recently developed, IDR(s) [5, 38] and QMRCGstab [20].

A known preconditioning technique for shifted linear systems is the so-called shift-and-invert preconditioner of the form $(A - \tau I)$ where the *seed shift* $\tau \in \mathbb{C}$ has to be chosen carefully. This preconditioner has been applied to shifted Helmholtz problems, for example, in [7]. Since the shift-and-invert matrix has to be solved by a direct method, this approach can be computationally costly. This can be overcome by either a multigrid approach [8, 26] or an approximation of the shift-and-invert preconditioner

*Received by the editors July 29, 2014; accepted for publication (in revised form) January 23, 2015; published electronically October 29, 2015.
http://www.siam.org/journals/sisc/37-5/97992.html
†Delft Institute of Applied Mathematics, Faculty EWI, Mekelweg 4, 2628 CD Delft, The Netherlands (m.m.baumann@tudelft.nl, m.b.vangijzen@tudelft.nl). The work of the first author was supported by Shell Global Solutions International B.V.

using a polynomial preconditioner as shown in [1]. Most recently, multiple shift-and-invert preconditioners have been combined in a flexible Krylov method in order to capture a wider range of shifts $\sigma_1, \ldots, \sigma_{N_\sigma}$; cf. [13, 25]. Our article is motivated by the question of whether we can use a Krylov method as a polynomial preconditioner within a flexible method. We will present two new algorithms that are both nested Krylov methods in the sense that an inner, collinear Krylov method is used as a (polynomial) preconditioner to solve shifted linear systems within a flexible outer Krylov iteration.

Our article is organized as follows. In section 2, we review the multi-shift GMRES [11] and the multi-shift QMRIDR [38] algorithms without preconditioning. The flexible versions of both are used as an outer Krylov method for the new nested framework in sections 3.4 and 3.5, respectively. As for the inner, collinear Krylov method, we present the multi-shift full orthogonalization method (FOM) algorithm [27] that automatically leads to collinear residuals in section 3.3.1. In order to use IDR(s) as an inner method, we present a new collinear variant in section 3.3.2. The article concludes 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.,

$$\mathcal{K}_m(A, \mathbf{b}) \equiv \text{span}\{\mathbf{b}, A\mathbf{b}, \ldots, A^{m-1}\mathbf{b}\} = \mathcal{K}_m(A - \sigma I, \mathbf{b}) \quad \forall \sigma \in \mathbb{C}. \tag{2.1}$$

The immediate consequence of this invariance property is that a basis for the underlying Krylov spaces only has to be built once for all shifted systems. Note that relation (2.1) also holds for collinear starting vectors but, in general, different right-hand sides in (1.1) destroy the shift-invariance.

2.1. Multi-shift GMRES. The well-known GMRES method [24] can be adapted to shifted systems in a straightforward way. In [11], a restarted version of multi-shift GMRES has been developed that relies on collinear residuals at restart. In this section, we review the main ideas of [11] and point out how the shift-invariance property (2.1) is used in the algorithm in order to speed up the computational performance when solving shifted systems numerically.

In the classical GMRES method for the unshifted system $A\mathbf{x} = \mathbf{b}$, an orthogonal basis of the mth Krylov subspace is obtained by the Arnoldi method. This leads to the well-known Hessenberg relation [34, eqn. 33.3],

$$AV_m = V_{m+1}\underline{\mathrm{H}}_m, \tag{2.2}$$

where the columns of $V_m \in \mathbb{C}^{N \times m}$ span an orthonormal basis of $\mathcal{K}_m(A, \mathbf{b})$, and $\underline{\mathrm{H}}_m \in \mathbb{C}^{(m+1)\times m}$ 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 mth iteration is given by

$$\mathbf{x}_m = V_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y}\in\mathbb{C}^m}{\text{argmin}} \|\underline{\mathrm{H}}_m \mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1\|, \tag{2.3}$$

with $\mathbf{e}_1$ being the first unit vector of $\mathbb{C}^{m+1}$, and $\mathbf{x}_0 = \mathbf{0}$. For simplicity, we will assume the initial guess to be equal to zero throughout the whole document. The optimization problem in (2.3) can be solved efficiently due to the Hessenberg structure of $\underline{\mathrm{H}}_m$ using,

for instance, Givens rotations; cf. [12, section 5.1.8]. Clearly, we see from (2.3) that $\mathbf{x}_m \in \mathcal{K}_m(A, \mathbf{b})$.

Because of the shift-invariance property (2.1), the matrix V_m which spans the mth Krylov subspace can be reused for any shift σ. Therefore, the Arnoldi iteration in multi-shift GMRES needs to be performed only once, and from the shifted Hessenberg relation,

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

we can derive an approximated solution to the shifted problem (1.2) via

$$\mathbf{x}_m^{(\sigma)} = V_m \mathbf{y}_m^{(\sigma)}, \quad \text{where } \mathbf{y}_m^{(\sigma)} = \underset{\mathbf{y} \in \mathbb{C}^m}{\operatorname{argmin}} \|(\underline{\mathrm{H}}_m - \sigma \underline{\mathrm{I}}_m)\mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1\|, \tag{2.4}$$

where $\underline{\mathrm{I}}_m$ is the identity matrix of size $m \times m$ with an extra zero row appended at the bottom. We note that the matrix $\underline{\mathrm{H}}_m(\sigma) \equiv \underline{\mathrm{H}}_m - \sigma \underline{\mathrm{I}}_m$ 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., $\underline{\mathrm{H}}_m(0) = \underline{\mathrm{H}}_m$.

By comparing (2.3) and (2.4), we note that the mth iterate lies in both cases in 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 $\underline{\mathrm{H}}_m(\sigma)$ 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 [11] require collinear residuals in order to preserve shift-invariance of the respective Krylov spaces after restart; cf. [11, Algorithm 2.4]. A numerically more robust implementation of restarted multi-shift GMRES has recently been proposed in [41].

2.2. Multi-shift QMRIDR. The QMRIDR method presented in [38] is a variant of the induced dimension reduction (IDR) method [31] that makes use of a so-called generalized Hessenberg decomposition and determines the mth iterate via a quasi-minimization of the mth residual. In [14, 38], the following relation is derived:

$$AG_m U_m = G_{m+1}\underline{\mathrm{H}}_m, \tag{2.5}$$

where $U_m \in \mathbb{C}^{m \times m}$ is upper triangular, $\underline{\mathrm{H}}_m \in \mathbb{C}^{(m+1) \times m}$ is of Hessenberg form, and $s+1$ consecutive vectors in G_m belong to the nested Sonneveld spaces $\mathcal{G}_0, \ldots, \mathcal{G}_j$. The entries of U_m, G_m, and $\underline{\mathrm{H}}_m$ are uniquely determined by the specific IDR algorithm [31, 39]; cf. [14] for a detailed derivation. Based on the generalized Hessenberg decomposition (2.5), a multi-shift version of the QMRIDR(s) algorithm has been derived in [38].

The approach of QMRIDR(s) is very similar to the GMRES approach. Therefore, the mth iterate is constructed as 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 $\underline{\mathrm{H}}_m$ only. Altogether, the following minimization problem needs to be solved:

$$\mathbf{x}_m = G_m U_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y} \in \mathbb{C}^m}{\operatorname{argmin}} \|\underline{\mathrm{H}}_m \mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1\|; \tag{2.6}$$

this is called quasi-minimization of the mth residual because the matrix G_m does not have orthogonal columns (cf. [38]).

For shifted linear systems of the form (1.2), a very similar relation holds,

$$\mathbf{x}_m^{(\sigma)} = G_m U_m \mathbf{y}_m^{(\sigma)}, \quad \text{where } \mathbf{y}_m^{(\sigma)} = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \|(\underline{\mathrm{H}}_m - \sigma\underline{\mathrm{U}}_m)\mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1\|, \tag{2.7}$$

and by comparing (2.6) with (2.7), we note that the approximate solutions to the respective systems lie in the same subspace and the matrix $\underline{\mathrm{H}}_m(\sigma) \equiv \underline{\mathrm{H}}_m - \sigma\underline{\mathrm{U}}_m$ is again of Hessenberg structure since $\underline{\mathrm{U}}_m$ consists of the upper triangular matrix U_m derived in [38] with an extra zero row appended.

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 $\underline{\mathrm{H}}_m(\sigma)$ is crucial for the design of the algorithm. Therefore, we will put emphasis on the computation of $\underline{\mathrm{H}}_m(\sigma)$ as a function of $\underline{\mathrm{H}}_m$ in the nested Krylov framework in sections 3.4 and 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 that preserves the shift-invariance property of the Krylov subspaces in (2.1). In this work, we restrict ourselves to such types of preconditioners that preserve the shift-invariance, but we would like to mention the promising approaches of [4, 32, 33], which do not rely on shift-invariance after preconditioning. Based on the requirements of a single preconditioner that preserves shift-invariance, we design a flexible preconditioner in subsection 3.2 that requires collinear residuals for the *inner* iteration. In subsection 3.3, we present two nested Krylov methods that lead to collinear residuals, namely, FOM, which produces collinear residuals automatically, and a new 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 flexible Krylov method in subsections 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 (1.2) without destroying the shift-invariance property of the respective Krylov spaces, we require the following equality after preconditioning:

$$\mathcal{K}_m((A-\sigma I)\mathcal{P}(\sigma)^{-1}, \mathbf{b}) = \mathcal{K}_m(A\mathcal{P}^{-1}, \mathbf{b}), \tag{3.1}$$

where $\mathcal{P}(\sigma)$ is a different preconditioner for every considered shifted system, and $\mathcal{P}$ is a preconditioner for the unshifted system $A\mathbf{x} = \mathbf{b}$; cf. [1, 15]. Relation (3.1) is satisfied if we find a parameter η that depends on the shift σ and a constant matrix $\mathcal{P}$ such that

$$(A-\sigma I)\mathcal{P}(\sigma)^{-1} = A\mathcal{P}^{-1} - \eta(\sigma) I, \tag{3.2}$$

which in fact means that we can write the preconditioned shifted systems as shifted preconditioned systems with new shifts $\eta(\sigma)$. From [18, 19, 28], it is well known that the so-called shift-and-invert preconditioner $\mathcal{P} \equiv (A - \tau I)$ applied to (3.2) leads to

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

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 $\sigma_1, \ldots, \sigma_{N_\sigma}$, as has been pointed out in [25]. In [1, 41], polynomial preconditioners are suggested that cheaply approximate $\mathcal{P}^{-1}$. Note that the right-hand side in (3.2) defines a new shifted problem, for which reason the single shift-and-invert preconditioner can be applied as a *first layer* in our later algorithm.

We remark that also in the more general case of shifted problems $(A-\sigma B)$ with a mass matrix $B \neq I$, the shift-and-invert preconditioner $\mathcal{P} = A - \tau B$ can be used. Applying this preconditioner yields a sequence of preconditioned shifted systems with a shifted identity in the same way as in the right-hand side of (3.2); cf. [18].

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 [22], where flexible GMRES (FGMRES) has been introduced for preconditioning systems of the form $A\mathbf{x} = \mathbf{b}$. For flexible preconditioning of shifted linear systems, we require a very similar relation to (3.2), namely,

$$(A-\sigma I)\mathcal{P}_j(\sigma)^{-1} = \alpha_j(\sigma)A\mathcal{P}_j^{-1} - \beta_j(\sigma)I, \tag{3.3}$$

where the parameters α_j and β_j will depend on the shift, and different preconditioners $\mathcal{P}_j$ and $\mathcal{P}_j(\sigma)$ are used in every iteration j. Note that the right-hand side in (3.3) is a shifted coefficient matrix and, thus, the shift-invariance is preserved by the flexible preconditioner. Since in a practical algorithm, the preconditioner is always directly applied to a vector $\mathbf{v}_j$, (3.3) translates into

$$(A-\sigma I)\mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j = \alpha_j(\sigma)A\mathcal{P}_j^{-1}\mathbf{v}_j - \beta_j(\sigma)\mathbf{v}_j. \tag{3.4}$$

We will next determine how α_j and β_j have to be chosen such that (3.3) and (3.4) hold. Therefore, we assume the preconditioning to be done by a multi-shift Krylov method itself (the *inner* method), which means that

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

are computed via a truncated multi-shift Krylov method at step j. 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 the same right-hand side $\mathbf{v}_j$. Hence, the corresponding (inner) residuals are given by

$$\mathbf{r}_j = \mathbf{v}_j - A\mathbf{z}_j = \mathbf{v}_j - A\mathcal{P}_j^{-1}\mathbf{v}_j, \tag{3.5}$$

$$\mathbf{r}_j^{(\sigma)} = \mathbf{v}_j - (A-\sigma I)\mathbf{z}_j^{(\sigma)} = \mathbf{v}_j - (A-\sigma I)\mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j. \tag{3.6}$$

We require the residuals (3.5)–(3.6) of the inner method to be collinear, i.e.,

$$\exists \gamma_j^{(\sigma)} \in \mathbb{C}: \quad \mathbf{r}_j^{(\sigma)} = \gamma_j^{(\sigma)}\mathbf{r}_j. \tag{3.7}$$

Note that the collinearity factor $\gamma_j^{(\sigma)}$ will be different at every iteration j and for every shift σ. Moreover, relation (3.7) is not a very strong assumption since, for example, multi-shift FOM [27], multi-shift BiCGstab [10], restarted multi-shift GMRES [11], and multi-shift IDR [5, 17] yield collinear residuals. With this assumption, α_j and β_j can be determined from (3.4) by using the collinearity relation (3.7), as the following calculation shows:

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

Thus, with (3.4) 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 $1 \le j \le m$. We show the relation

$$(A-\sigma I)\mathbf{z}_j^{(\sigma)} = \gamma_j^{(\sigma)} A\mathbf{z}_j - \left(\gamma_j^{(\sigma)} - 1\right)\mathbf{v}_j, \tag{3.8}$$

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

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

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

3.3. Krylov methods with collinear residuals. In the previous section, we 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 (3.7) for the inner multi-shift method. Next, we will present two multi-shift Krylov methods that lead to collinear residuals and, therefore, fulfill assumption (3.7). It is well-known from [27] that the multi-shift version of FOM leads to collinear residuals. We will review this result in subsection 3.3.1. Moreover, we describe a new variant of the IDR method that has collinear residuals in subsection 3.3.2.

3.3.1. Collinear residuals in multi-shift FOM. Multi-shift FOM (msFOM) can be derived very similarly to multi-shift GMRES of section 2.1; cf. [27]. In FOM, an orthogonal basis of $\mathcal{K}_m(A, \mathbf{b})$ is obtained via the Arnoldi iteration, which yields

$$V_m^{\mathsf{H}} A V_m = H_m,$$

and can be derived from (2.2) by left-multiplication with V_m^{H}, where H_m is a square matrix and the superscript H denotes Hermitian transpose. Assuming H_m is invertible, the mth iterate is then obtained by

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

ALGORITHM 1: msFOM FOR $(A-\sigma_k I)\mathbf{x}_k = \mathbf{b},\ k = 1, \ldots, N_\sigma$ [27].

// Single FOM:
Initialize $\mathbf{r}_0 = \mathbf{b},\ \beta = \|\mathbf{r}_0\|,\ \mathbf{v}_1 = \mathbf{r}_0/\beta$
for $j = 1$ to m **do**
  Compute $\mathbf{w} = A\mathbf{v}_j$
  for $i = 1$ to j **do**
    $h_{i,j} = \mathbf{v}_i^{\mathsf{H}}\mathbf{w}$
    $\mathbf{w} = \mathbf{w} - h_{i,j}\mathbf{v}_i$
  end for
  Set $h_{j+1,j} = \|\mathbf{w}\|$ and $\mathbf{v}_{j+1} = \mathbf{w}/h_{j+1,j}$
end for
Solve $\mathbf{y}_m = H_m^{-1}(\beta\mathbf{e}_1)$
Set $\mathbf{x}_m = V_m\mathbf{y}_m$
// Multi-shift FOM:
for $k = 1$ to N_σ **do**
  Construct $H_m(\sigma_k) = H_m - \sigma_k I_m$
  Solve $\mathbf{y}_m^{(\sigma_k)} = H_m(\sigma_k)^{-1}(\beta\mathbf{e}_1)$
  Set $\mathbf{x}_m^{(\sigma_k)} = V_m\mathbf{y}_m^{(\sigma_k)}$
end for

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 [27]. The complete multi-shift FOM algorithm is repeated in Algorithm 1. It is well known that msFOM 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*

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

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

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

Proof. This proof can be found in [23, Proposition 6.7] as well as in [27]. For the residual of the original system it holds after m iterations that

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

Repeating the same calculation for the shifted system yields

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

with $\mathbf{e}_m \equiv [0, \ldots, 0, 1]^{\mathsf{H}} \in \mathbb{C}^m$, 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

$$\gamma^{(\sigma)} = y_m^{(\sigma)}/y_m, \tag{3.9}$$

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 (3.9). ∎

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 jth outer iteration in (3.8) is given by (3.9), 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(s) method as presented in [31] 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

$$\mathcal{G}_{j+1} \ni \mathbf{r}_{n+1} = (I - \omega_{j+1}A)\mathbf{v}_n \quad \text{with } \mathbf{v}_n \in \mathcal{G}_j \cap \mathcal{S},\ \omega_{j+1} \in \mathbb{C} \setminus \{0\}, \tag{3.10}$$

with $\mathcal{G}_0 \equiv \mathcal{K}_N(A, \mathbf{b})$, and, without loss of generality, let $\mathcal{S} = \mathcal{N}(P^{\mathsf{H}})$ be the null space of some $N \times s$ matrix $P = [\mathbf{p}_1, \ldots, \mathbf{p}_s]$. It is known from the IDR theorem [31, Theorem 2.1] that the spaces $\mathcal{G}_j$ that are generated via the recursive definition (3.10) are of decreasing dimension and that, eventually, $\mathcal{G}_j = \{\mathbf{0}\}$ for some $j \leq N$. This result guarantees that in exact arithmetic the residual will be equal to zero at some point. Moreover, note that the scalars ω_{j+1} in (3.10) 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 [31], the vectors $\mathbf{v}_n \in \mathcal{G}_j \cap \mathcal{S}$ can be computed in the following way:

$$(P^{\mathsf{H}}\Delta R_n)\mathbf{c} = P^{\mathsf{H}}\mathbf{r}_n, \tag{3.11}$$

$$\mathbf{v}_n = \mathbf{r}_n - \Delta R_n \mathbf{c}, \tag{3.12}$$

with $\Delta R_n \equiv [\Delta\mathbf{r}_{n-1}, \ldots, \Delta\mathbf{r}_{n-s}]$ 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 $R_n \equiv [\mathbf{r}_n, \ldots, \mathbf{r}_{n-s}]$. From the last two definitions, we note that the residual updates can be expressed as a product of the actual residuals and a difference matrix D as

$$[\mathbf{r}_n, -\Delta R_n] = R_n D, \tag{3.13}$$

where D is defined as the invertible matrix:

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

Note that (3.13) holds in the same way for residuals of the shifted systems.

A first approach to adapt the IDR method to shifted linear systems that leads to collinear residuals has been done by [17]. The algorithm presented in [17] follows the

classical approach of evaluating the shifted residual polynomial. For the collinear IDR variant presented in this section, we derive a purely algebraic approach that exploits the IDR theorem and an implementation of the shifted algorithm that is very closely related to the first IDR version in [31]. After finishing this manuscript, we discovered that the same approach but a different implementation of it was independently followed by [5].

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. Moreover, we express the collinearity via a diagonal matrix $\tilde{\Gamma}_n^{(\sigma)} \equiv \text{diag}(\gamma_n^{(\sigma)}, \ldots, \gamma_{n-s}^{(\sigma)})$ that consists of the last $s+1$ consecutive collinearity factors such that it holds that

$$R_n^{(\sigma)} = R_n \tilde{\Gamma}_n^{(\sigma)}, \tag{3.14}$$

where $R_n^{(\sigma)} \equiv [\mathbf{r}_n^{(\sigma)}, \ldots, \mathbf{r}_{n-s}^{(\sigma)}]$ is constructed in the same way as R_n.

Our approach can be described in two phases. First, we note from (3.10) that in order to obtain collinear residuals, we need to produce collinear vectors $\mathbf{v}_n$ and $\mathbf{v}_n^{(\sigma)}$. Therefore, we want to calculate the intermediate vector $\mathbf{c}^{(\sigma)}$ in (3.11) of the shifted systems such that $\mathbf{v}_n^{(\sigma)} = \alpha^{(\sigma)}\mathbf{v}_n$, with $\alpha^{(\sigma)} \in \mathbb{C}$ to be determined. In the following calculation, we make use of (3.13) and (3.14) in order to manipulate relation (3.12) for the shifted system:

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

By comparing with (3.12), we obtain the following $(s+1)\times(s+1)$ system of equations:

$$D^{-1}\tilde{\Gamma}_n^{(\sigma)} D \begin{bmatrix} 1 \\ \mathbf{c}^{(\sigma)} \end{bmatrix} = \alpha^{(\sigma)} \begin{bmatrix} 1 \\ \mathbf{c} \end{bmatrix}, \tag{3.15}$$

where the vector $\mathbf{c}$ is known, and $\mathbf{c}^{(\sigma)} \in \mathbb{C}^s$ and $\alpha^{(\sigma)} \in \mathbb{C}$ can be uniquely determined. Note that in contrast to (3.11)–(3.12), we have computed $\mathbf{c}^{(\sigma)}$ and $\mathbf{v}_n^{(\sigma)} = \alpha^{(\sigma)}\mathbf{v}_n$ of the shifted systems without storing additional residual differences.

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.,

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

Therefore, we substitute the definition of the residuals from (3.10) and use the collinearity of the vectors $\mathbf{v}_n$ and $\mathbf{v}_n^{(\sigma)}$:

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

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

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

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

$$\omega_{j+1}^{(\sigma)} = \frac{\omega_{j+1}}{1-\omega_{j+1}\sigma}, \quad \gamma^{(\sigma)} = \frac{\alpha^{(\sigma)}}{1-\omega_{j+1}\sigma}. \tag{3.16}$$

Thus, we have derived a formula for the collinearity factor $\gamma^{(\sigma)}$ which can be used in the nested framework in (3.8). In Algorithm 2, we present the collinear IDR variant (called collIDR(s)) using (3.16) for the choice of the free IDR parameter $\omega_{j+1}^{(\sigma)}$. Note that in (3.16) 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 sections 4.1 and 4.2 of [31], respectively. As mentioned, the implementation of Algorithm 2 does not require the storage of residual differences for the shifted systems. However, we need to additionally store s updates for the iterates.

3.4. Nested FOM-FGMRES for shifted linear systems. We now present a special case of the nested Krylov framework of section 3.2, namely, a version where msFOM is used as inner preconditioner and FGMRES is used as an outer Krylov iteration. FGMRES has been introduced in [22] for unshifted systems $A\mathbf{x} = \mathbf{b}$ and allows a different preconditioner $\mathcal{P}_j$ in the jth outer iteration. The Hessenberg relation (2.2) therein extends to

$$AZ_m = V_{m+1}\underline{\mathrm{H}}_m, \quad (A-\sigma I)Z_m^{(\sigma)} = V_{m+1}\underline{\mathrm{H}}_m(\sigma), \tag{3.17}$$

where at step $1 \le j \le m$ the (flexible) preconditioning $\mathbf{z}_j \equiv \mathcal{P}_j^{-1}\mathbf{v}_j$, $\mathbf{z}_j^{(\sigma)} \equiv \mathcal{P}_j(\sigma)^{-1}\mathbf{v}_j$ is carried out, and $Z_m \equiv [\mathbf{z}_1, \ldots, \mathbf{z}_m]$ and $Z_m^{(\sigma)} \equiv [\mathbf{z}_1^{(\sigma)}, \ldots, \mathbf{z}_m^{(\sigma)}]$, respectively. Note that (3.17) is formulated for both the shifted and unshifted cases and that one column of relation (3.17) yields

$$A\mathbf{z}_j = V_{m+1}\underline{\mathbf{h}}_j, \quad (A-\sigma I)\mathbf{z}_j^{(\sigma)} = V_{m+1}\underline{\mathbf{h}}_j^{(\sigma)},$$

which we will use next in order to determine the shifted Hessenberg matrix $\underline{\mathrm{H}}_m(\sigma)$ *columnwise*. To this end, we assume that the preconditioner $\mathcal{P}_j(\sigma)$ is equivalent to a truncated msFOM inner iteration with collinear factors $\gamma_j^{(\sigma)}$ for the residuals according to (3.9). From the calculation

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

we can conclude the jth column of the shifted Hessenberg matrix to be

$$\underline{\mathbf{h}}_j^{(\sigma)} = \gamma_j^{(\sigma)}\underline{\mathbf{h}}_j - \left(\gamma_j^{(\sigma)}-1\right)\underline{\mathbf{e}}_j, \quad 1 \le j \le m, \tag{3.18}$$

with $\underline{\mathbf{e}}_j$ being the jth unit vector of $\mathbb{C}^{m+1}$.

ALGORITHM 2: collIDR(s) FOR $(A-\sigma_k I)\mathbf{x}_k = \mathbf{b},\ k = 1,\ldots,N_\sigma$.

Set $\mathbf{x}_0 = \mathbf{x}_0^{(\sigma_k)} = \mathbf{0}$, $\mathbf{r}_0 = \mathbf{b}$, and $\gamma_0^{(\sigma_k)} = 1$
for $n = 0,\ldots,s-1$ **do**
  $\mathbf{v} = A\mathbf{r}_n$; $\omega = (\mathbf{v}^{\mathsf{H}}\mathbf{r}_n)/(\mathbf{v}^{\mathsf{H}}\mathbf{v})$
  $\Delta\mathbf{x}_n = \omega\mathbf{r}_n$; $\Delta\mathbf{r}_n = -\omega\mathbf{v}$
  for $k = 1,\ldots,N_\sigma$ **do**
    $\gamma_{n+1}^{(\sigma_k)} = \gamma_n^{(\sigma_k)}/(1-\omega\sigma_k)$; $\omega^{(\sigma_k)} = \omega/(1-\omega\sigma_k)$
    $\Delta\mathbf{x}_n^{(\sigma_k)} = \omega^{(\sigma_k)}\gamma_n^{(\sigma_k)}\mathbf{r}_n$
    $\mathbf{x}_{n+1}^{(\sigma_k)} = \mathbf{x}_n^{(\sigma_k)} + \Delta\mathbf{x}_n^{(\sigma_k)}$
  end for
  Update: $\mathbf{x}_{n+1} = \mathbf{x}_n + \Delta\mathbf{x}_n$; $\mathbf{r}_{n+1} = \mathbf{r}_n + \Delta\mathbf{r}_n$
end for
$\Delta R_{n+1} := (\Delta\mathbf{r}_n,\ldots,\Delta\mathbf{r}_0)$; $\Delta X_{n+1} := (\Delta\mathbf{x}_n,\ldots,\Delta\mathbf{x}_0)$
$\Delta X_{n+1}^{(\sigma_k)} := (\Delta\mathbf{x}_n^{(\sigma_k)},\ldots,\Delta\mathbf{x}_0^{(\sigma_k)})$; $\boldsymbol{\gamma}^{(\sigma_k)} := (\gamma_{n+1}^{(\sigma_k)},\ldots,\gamma_0^{(\sigma_k)})$
$n = s$
while $\|\mathbf{r}_n\| > TOL$ **and** $n < MAXIT$ **do**
  for $j = 0,\ldots,s$ **do**
    Solve $\mathbf{c}$ from $P^{\mathsf{H}}\Delta R_n\mathbf{c} = P^{\mathsf{H}}\mathbf{r}_n$
    $\mathbf{v} = \mathbf{r}_n - \Delta R_n\mathbf{c}$
    for $k = 1,\ldots,N_\sigma$ **do**
      $\tilde{\Gamma}_n^{(\sigma_k)} := \mathrm{diag}(\boldsymbol{\gamma}^{(\sigma_k)})$
      Solve $[1,\mathbf{c}^{(\sigma_k)}]$ from $D^{-1}\tilde{\Gamma}_n^{(\sigma_k)}D[1,\mathbf{c}^{(\sigma_k)}] = \alpha^{(\sigma_k)}[1,\mathbf{c}]$ s.t. $\mathbf{v}^{(\sigma_k)} = \alpha^{(\sigma_k)}\mathbf{v}$
    end for
    if j == 0 **then**
      $\mathbf{t} = A\mathbf{v}$
      $\omega = (\mathbf{t}^{\mathsf{H}}\mathbf{v})/(\mathbf{t}^{\mathsf{H}}\mathbf{t})$; $\omega^{(\sigma_k)} = \omega/(1-\omega\sigma_k)$
      $\Delta\mathbf{x}_n = -\Delta X_n\mathbf{c} + \omega\mathbf{v}$; $\Delta\mathbf{r}_n = -\Delta R_n\mathbf{c} - \omega\mathbf{t}$
    else
      $\Delta\mathbf{x}_n = -\Delta X_n\mathbf{c} + \omega\mathbf{v}$; $\Delta\mathbf{r}_n = -A\Delta\mathbf{x}_n$
    end if
    Update: $\mathbf{x}_{n+1} = \mathbf{x}_n + \Delta\mathbf{x}_n$; $\mathbf{r}_{n+1} = \mathbf{r}_n + \Delta\mathbf{r}_n$
    for $k = 1,\ldots,N_\sigma$ **do**
      $\gamma_{n+1}^{(\sigma_k)} = \alpha^{(\sigma_k)}/(1-\omega\sigma_k)$
      $\Delta\mathbf{x}_n^{(\sigma_k)} = -\Delta X_n^{(\sigma_k)}\mathbf{c}^{(\sigma_k)} + \omega^{(\sigma_k)}\alpha^{(\sigma_k)}\mathbf{v}$
      $\mathbf{x}_{n+1}^{(\sigma_k)} = \mathbf{x}_n^{(\sigma_k)} + \Delta\mathbf{x}_n^{(\sigma_k)}$
    end for
    // The IDR-update:
    $n = n+1$
    $\Delta R_n := (\Delta\mathbf{r}_{n-1},\ldots,\Delta\mathbf{r}_{n-s})$; $\Delta X_n := (\Delta\mathbf{x}_{n-1},\ldots,\Delta\mathbf{x}_{n-s})$
    $\Delta X_n^{(\sigma_k)} := (\Delta\mathbf{x}_{n-1}^{(\sigma_k)},\ldots,\Delta\mathbf{x}_{n-s}^{(\sigma_k)})$; $\boldsymbol{\gamma}^{(\sigma_k)} := (\gamma_n^{(\sigma_k)},\ldots,\gamma_{n-s}^{(\sigma_k)})$
  end for
end while

Aligning the columns of m outer iterations together yields the following formula for the shifted Hessenberg matrix:

$$\underline{\mathrm{H}}_m(\sigma) = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m^{(\sigma)} + \underline{\mathrm{I}}_m, \tag{3.19}$$

ALGORITHM 3: NESTED FOM-FGMRES FOR $(A-\sigma_k I)\mathbf{x}_k = \mathbf{b},\ k = 1, \ldots, N_\sigma$.

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: $\mathbf{z}_j^{(\sigma_k)} = msFOM(A - \sigma_k I, \mathbf{v}_j)$
4: Compute $\gamma_j^{(\sigma_k)}$ according to (3.9)
5: Compute $\mathbf{w} = A\mathbf{z}_j^{(0)}$
6: **for** $i = 1$ to j **do**
7: $h_{i,j} = \mathbf{v}_i^{\mathsf{H}}\mathbf{w}$
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_σ **do**
13: Define $Z_j^{(\sigma_k)} = [\mathbf{z}_1^{(\sigma_k)}, \ldots, \mathbf{z}_j^{(\sigma_k)}]$
14: Construct $\underline{\mathrm{H}}_j(\sigma_k)$ according to (3.19)
15: Solve $\mathbf{y}_j^{(\sigma_k)} = \mathrm{argmin}_{\mathbf{y}} \left\|\beta\mathbf{e}_1 - \underline{\mathrm{H}}_j(\sigma_k)\mathbf{y}\right\|$, with $\mathbf{e}_1 = [1, 0, \ldots, 0]^{\mathsf{H}} \in \mathbb{R}^{j+1}$
16: Set $\mathbf{x}_j^{(\sigma_k)} = Z_j^{(\sigma_k)}\mathbf{y}_j^{(\sigma_k)}$
17: **end for**
18: **end for**

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

$$\Gamma_m^{(\sigma)} \equiv \begin{bmatrix} \gamma_1^{(\sigma)} & & \\ & \ddots & \\ & & \gamma_m^{(\sigma)} \end{bmatrix} \in \mathbb{C}^{m\times m}. \tag{3.20}$$

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 (3.19) reduces to the original Hessenberg matrix, $\underline{\mathrm{H}}_m(0) = \underline{\mathrm{H}}_m$, because in this case the collinearity factors are all equal to one and, hence, $\Gamma_m^{(0)} = I$. The FOM-FGMRES nested Krylov method for shifted linear systems is summarized in Algorithm 3.

The least-squares problem in line 15 minimizes the residual of each shifted system, as the following calculation proves:

$$\begin{aligned}\mathbf{x}_j^{(\sigma)} &= \underset{\mathbf{x}\in\mathcal{Z}_j^{(\sigma)}}{\mathrm{argmin}} \|\mathbf{b} - (A-\sigma I)\mathbf{x}\| = \underset{\mathbf{y}\in\mathbb{C}^j}{\mathrm{argmin}} \left\|\mathbf{b} - (A-\sigma I)Z_j^{(\sigma)}\mathbf{y}\right\| \\ &= \underset{\mathbf{y}\in\mathbb{C}^j}{\mathrm{argmin}} \left\|\mathbf{b} - V_{j+1}\underline{\mathrm{H}}_j(\sigma)\mathbf{y}\right\| = \underset{\mathbf{y}\in\mathbb{C}^j}{\mathrm{argmin}} \left\|\beta\mathbf{e}_1 - \left(\left(\underline{\mathrm{H}}_j - \underline{\mathrm{I}}_j\right)\Gamma_j^{(\sigma)} + \underline{\mathrm{I}}_j\right)\mathbf{y}\right\|,\end{aligned}$$

where we used the flexible shifted Arnolid relation of (3.17) as well as (3.19).

We note that in the same way as in FGMRES for the unshifted case [22], 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 major drawback of FGMRES that has already been pointed out by [22] and applies here for every shift. We therefore present in section 3.5 a nested algorithm that uses flexible QMRIDR [38, section 4] as an outer method and partly overcomes this storage requirement.

3.5. Nested IDR-FQMRIDR for shifted linear systems. Similar way to the nested FOM-FGMRES algorithm, we present a nested IDR-FQMRIDR method for shifted linear systems where collIDR(s) from Algorithm 2 is used as an inner method. In contrast to the combination in section 3.4, this is a combination of two short recurrence methods. For $V_m \equiv G_m U_m$, relation (2.5) in QMRIDR was given by,

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

In flexible QMRIDR (FQMRIDR) which has been introduced in [38, section 4], this relation is replaced by

$$AZ_m = G_{m+1}\underline{\mathrm{H}}_m, \quad (A-\sigma I)Z_m^{(\sigma)} = G_{m+1}\underline{\mathrm{H}}_m(\sigma), \tag{3.21}$$

with $Z_m, Z_m^{(\sigma)}$ consisting of the respective columns $\mathbf{z}_j \equiv \mathcal{P}_j^{-1}\mathbf{v}_j$, $\mathbf{z}_j^{(\sigma)} \equiv \mathcal{P}(\sigma)_j^{-1}\mathbf{v}_j$ for $1 \le j \le m$, just as before. One column of (3.21) reads

$$A\mathbf{z}_j = G_{m+1}\underline{\mathbf{h}}_j, \quad (A-\sigma I)\mathbf{z}_j^{(\sigma)} = G_{m+1}\underline{\mathbf{h}}_j^{(\sigma)},$$

which we will use in order to derive the shifted Hessenberg matrix of IDR-FQMRIDR.

We assume that the factors $\gamma_j^{(\sigma)}$ are given from (3.16) by an inner collIDR(s) iteration. Then, the FQMRIDR 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

$$\underline{\mathbf{h}}_j^{(\sigma)} = \gamma_j^{(\sigma)}\underline{\mathbf{h}}_j - \left(\gamma_j^{(\sigma)} - 1\right)\underline{\mathbf{u}}_j, \quad 1 \le j \le m, \tag{3.22}$$

where $\underline{\mathbf{u}}_j \equiv [\mathbf{u}_j, 0]^{\mathsf{H}}$ is the vector $\mathbf{u}_j$ from the jth iteration of QMRIDR [38] with an extra zero.

Altogether, we have derived the shifted Hessenberg matrix,

$$\underline{\mathrm{H}}_m(\sigma) = (\underline{\mathrm{H}}_m - \underline{\mathrm{U}}_m)\,\Gamma_m^{(\sigma)} + \underline{\mathrm{U}}_m, \tag{3.23}$$

with $\Gamma_m^{(\sigma)}$ as defined in (3.20), and $\underline{\mathrm{U}}_m \equiv [\underline{\mathbf{u}}_1, \ldots, \underline{\mathbf{u}}_m]$. Here we see the close relation between the two nested methods. By comparing the expression for the shifted Hessenberg matrices in (3.19) and (3.23), 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 [38, 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 (3.22) in order to adapt [38, Algorithm 1] to our nested algorithm.

ALGORITHM 4: NESTED IDR-FQMRIDR FOR $(A-\sigma_k I)\mathbf{x}_k = \mathbf{b},\ k = 1, \ldots, N_\sigma$.

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: $\mathbf{z}_j^{(\sigma_k)} = collIDR(A - \sigma_k I, \mathbf{v}_j)$
4: Compute $\gamma_j^{(\sigma_k)}$ according to (3.16)
5: Compute $\underline{\mathbf{h}}_j, \underline{\mathbf{u}}_j$ as in QMRIDR; see [38, Algorithm 1]
6: Compute $\underline{\mathbf{h}}_j^{(\sigma_k)}$ from (3.22)
7: *// Loop over shifted systems:*
8: **for** $k = 1$ to N_σ **do**
9: Define $Z_j^{(\sigma_k)} = [\mathbf{z}_1^{(\sigma_k)}, \ldots, \mathbf{z}_j^{(\sigma_k)}]$
10: Construct $\underline{\mathrm{H}}_j(\sigma_k)$ according to (3.23)
11: Solve $\mathbf{y}^{(\sigma_k)} = \operatorname{argmin}_{\mathbf{y}} \left\|\beta\mathbf{e}_1 - \underline{\mathrm{H}}_j(\sigma_k)\mathbf{y}\right\|$, with $\mathbf{e}_1 = [1, 0, \ldots, 0]^{\mathsf{H}} \in \mathbb{R}^{j+1}$
12: Set $\mathbf{x}_j^{(\sigma_k)} = Z_j^{(\sigma_k)}\mathbf{y}^{(\sigma_k)}$
13: **end for**
14: **end for**

In line 11 of Algorithm 4, the following *quasi-minimization* of the shifted residual is carried out:

$$\begin{aligned}
\mathbf{x}_j^{(\sigma)} &= \operatorname*{argmin}_{\mathbf{x}\in\mathcal{Z}_j^{(\sigma)}} \|\mathbf{b} - (A-\sigma I)\mathbf{x}\| = \operatorname*{argmin}_{\mathbf{y}\in\mathbb{C}^j} \left\|\mathbf{b} - (A-\sigma I)Z_j^{(\sigma)}\mathbf{y}\right\| \\
&= \operatorname*{argmin}_{\mathbf{y}\in\mathbb{C}^j} \left\|\mathbf{b} - G_{j+1}\underline{\mathrm{H}}_j(\sigma)\mathbf{y}\right\| = \operatorname*{argmin}_{\mathbf{y}\in\mathbb{C}^j} \left\|G_{j+1}\left(\beta\mathbf{e}_1 - \underline{\mathrm{H}}_j(\sigma)\mathbf{y}\right)\right\| \\
&\le \operatorname*{argmin}_{\mathbf{y}\in\mathbb{C}^j} \left\|\beta\mathbf{e}_1 - \underline{\mathrm{H}}_j(\sigma)\mathbf{y}\right\| = \operatorname*{argmin}_{\mathbf{y}\in\mathbb{C}^j} \left\|\beta\mathbf{e}_1 - \left(\left(\underline{\mathrm{H}}_j - \underline{\mathrm{U}}_j\right)\Gamma_j^{(\sigma)} + \underline{\mathrm{U}}_j\right)\mathbf{y}\right\|
\end{aligned}$$

with an estimation for $\|G_{j+1}\|$ given in [38].

4. Numerical experiments. The numerical examples we present are motivated from geophysical applications. We consider the numerical solution of the Helmholtz equation in section 4.1 and of the time-harmonic elastic wave equation (Navier 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 (1.1) and the frequencies of the waves.

All examples have been implemented in MATLAB version R2011B, executed on an Intel Xeon CPU E3-1240 V2 at 3.40 GHz. For a more detailed description of the numerical tests, we refer to the extended report [3].

For the numerical solution of shifted linear systems of the form

$$(A - \sigma_k I)\mathbf{x}_k = \mathbf{b}, \quad k = 1, \ldots, N_\sigma, \tag{4.1}$$

it is of practical use to reformulate the problem (4.1) by the substitutions

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

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

$$(\bar{A} - \bar{\sigma}_k I)\mathbf{x}_k = \mathbf{b}, \quad k = 1, \ldots, N_\sigma, \tag{4.2}$$

the only difference being that in the formulation (4.2) the unshifted ($\bar{\sigma}_k = 0$) 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 Helmholtz equation,

$$-\Delta p - \left(\frac{2\pi f_k}{c(\mathbf{x})}\right)^2 p = s, \quad \text{in } \Omega \subset \mathbb{R}^2, \tag{4.3}$$

where p stands for the pressure and f_k is the kth wave frequency. We consider the *wedge problem* which was introduced in [21]. Therein, the computational domain is $\Omega = [0, 600] \times [0, 1000]$ and the underlying sound velocity $c(\mathbf{x})$ is heterogeneous and represents three different layers; cf. [3, 21] for more details. 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),

$$\frac{\partial p}{\partial n} = 0, \quad \text{on } \partial\Omega, \tag{4.4}$$

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

$$\frac{\partial p}{\partial n} - i\left(\frac{2\pi f_k}{c(\mathbf{x})}\right) p = 0, \quad \text{on } \partial\Omega, \tag{4.5}$$

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 [6, 7, 8, 26, 37]. When nonhomogeneous Neumann boundary conditions (4.5) are included in (4.3), we end up with a discretization of the form

$$(K + i(2\pi f_k)C - (2\pi f_k)^2 M)\mathbf{p} = \mathbf{s}, \quad \sigma_k = 2\pi f_k, \tag{4.6}$$

where C represents the boundary conditions (4.5), K is the (negative) discrete Laplacian, and M is a mass matrix. The unknown vector $\mathbf{p}$ consists of discrete pressure values. Note that (4.6) is quadratic in σ_k and can be linearized as

$$\left\{ \begin{bmatrix} iC & K \\ I & 0 \end{bmatrix} - \sigma_k \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \right\} \begin{bmatrix} \sigma_k \mathbf{p} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{s} \\ 0 \end{bmatrix}, \quad k = 1, \ldots, N_\sigma, \tag{4.7}$$

as suggested in [27, 28], which yields a shifted problem. For reflecting boundary conditions (4.4), we get $C = 0$, and a shifted linear system can be obtained by simply multiplying (4.6) with M^{-1}, which we avoid in the following. The shift-and-invert preconditioner (as introduced in section 3.1) for (4.7) is given by

$$\mathcal{P} = \begin{bmatrix} iC & K \\ I & 0 \end{bmatrix} - \tau \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} iC - \tau M & K \\ I & -\tau I \end{bmatrix}, \tag{4.8}$$

with damping parameter $\tau \in \mathbb{C}$, and can be seen as the analogue of the shifted Laplace preconditioner of [6, 7, 8] applied to the block system (4.7). We remark that (4.8) can be decomposed and inverted in the following way:

$$\mathcal{P}^{-1} = \begin{bmatrix} I & \tau I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & (K + i\tau C - \tau^2 M)^{-1} \end{bmatrix} \begin{bmatrix} 0 & I \\ I & -iC + \tau M \end{bmatrix}. \tag{4.9}$$

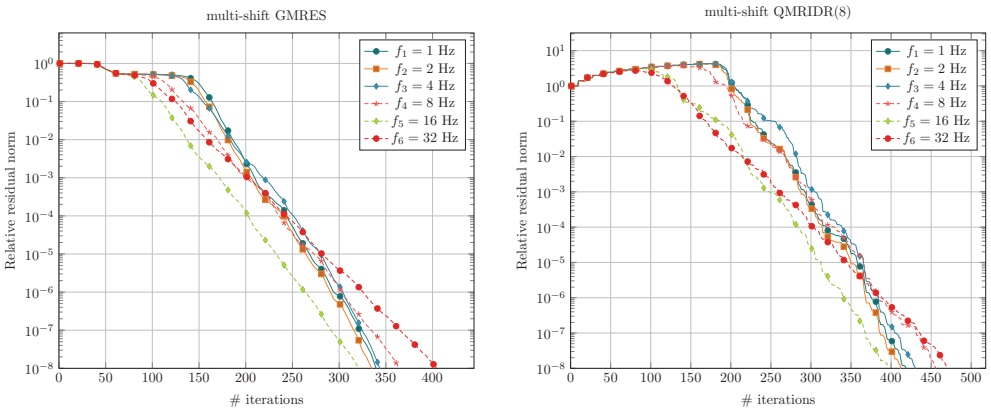

FIG. 1. *Convergence behavior of multi-shift GMRES (left) and multi-shift QMRIDR*(8) *(right) for* (4.3)–(4.5).

Therefore, from a computational point of view, applying the preconditioner (4.8) to (4.7) reduces to efficiently solving systems with the shifted Laplacian $(K+i\tau C-\tau^2 M)$ in (4.9) which is of the same dimension as (4.6).

For large-scale applications, a multigrid approach as in [6, 8, 26] can be used to approximate the shifted Laplace preconditioner. This gives rise to the question to which accuracy a multigrid method needs to be applied. Using the nested method of section 3.4 or section 3.5 for the preconditioned shifted problem yields a setting similar to [36], where the authors analyze inexact nested Krylov methods for the unshifted case. The insight of [36] on the accuracy that is needed for matrix-vector multiplications may give guidelines for an extension to the shifted framework.

In the present test case, we aim to solve (4.7) for a range of six frequencies $f_k = \{1, 2, 4, 8, 16, 32\}$ Hz and restrict ourselves to a direct method for the shift-and-invert preconditioner. For the resolution of high frequencies and due to the doubling of unknowns from (4.6) to (4.7), the system size becomes more than 30,000 equations. We present the convergence behavior for multi-shift GMRES and multi-shift QMRIDR with $s = 8$ using only the single shift-and-invert preconditioner (4.9) in Figure 1, respecively. The convergence curves show that the residual norms first stagnate or even increase in the case of QMRIDR(8). Moreover, the convergence rates are nearly linear as soon as the residual norms start decreasing. Therefore, it is intuitive to truncate the inner iterations in this region. The convergence plots of nested FOM-FGMRES and nested IDR-FQMRIDR(8) are presented in Figures 2 and 3, respectively.

In both nested algorithms, the number of *inner* iterations is chosen in such a way that the relative residual norms are of 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.

In Table 1, we want to point out the CPU 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 (FOM-FGMRES), 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 Arnoldi orthogonalization process. This also explains why we observe no speed-up for multi-shift QMRIDR, which is a short recurrence method by design.

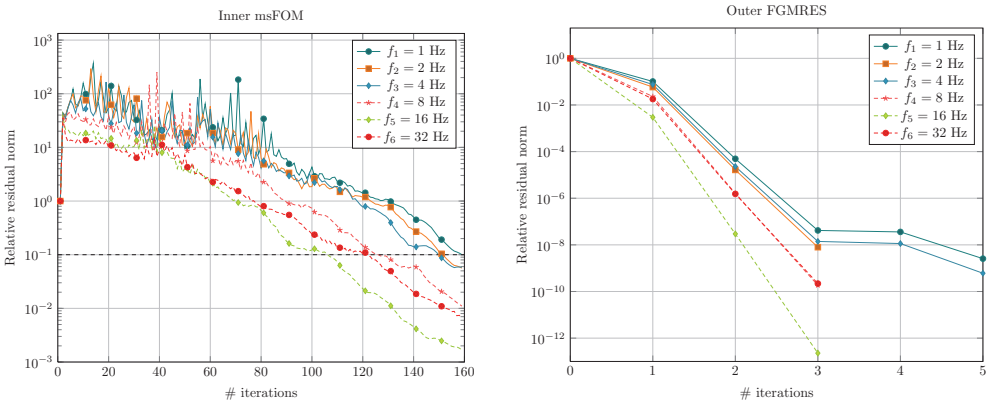

FIG. 2. *Convergence behavior of FOM-FGMRES for* (4.3)–(4.5)*: typical inner convergence (left) and outer convergence (right).*

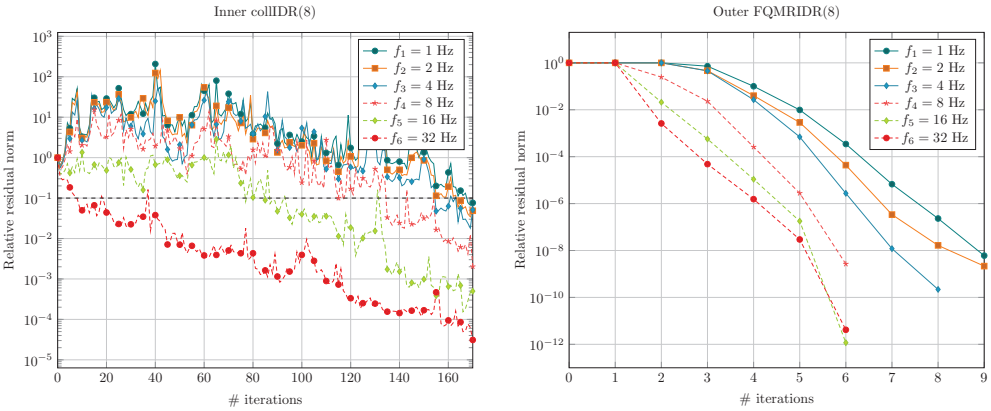

FIG. 3. *Convergence behavior of IDR-FQMRIDR* (8) *for* (4.3)–(4.5)*: typical inner convergence (left) and outer convergence (right).*

TABLE 1

Comparison of multi-shift and nested multi-shift algorithms for the wedge problem (4.3) *with absorbing boundary conditions* (4.5)*. Here,* $\hat{\sigma} \equiv \max\{\sigma_k, k = 1, \ldots, N_\sigma\}$.

	Multi-shift Krylov methods			
	msGMRES	rest_msGMRES	msQMRIDR(8)	collIDR(8)
# inner iterations	-	200	-	-
# outer iterations	404	4	471	584
Seed shift τ	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$
CPU time	157.33s	39.4s	62.51s	79.13s
	Nested multi-shift Krylov methods			
	FOM-FGMRES	IDR(8)-FGMRES	FOM-FQMRIDR(8)	IDR(8)-FQMRIDR(8)
# inner iterations	160	250	100	170
# outer iterations	5	13	8	9
Seed shift τ	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$
CPU time	28.56s	216.5s	24.4s	75.41s

Moreover, we included other possible combinations of the nested algorithm which show that a combination of inner msFOM with outer FQMRIDR(8) performs best. Table 1 also shows the performance of collIDR(8) as a stand-alone multi-shift

TABLE 2
Value of constant parameters taken from [2].

$\rho\ [kg/m^3]$	$c_p\ [m/s]$	$c_s\ [m/s]$	$\lambda\ [Pa]$	$\mu\ [Pa]$
$2.7 \cdot 10^3$	$6.8983 \cdot 10^3$	$4.3497 \cdot 10^3$	$2.6316 \cdot 10^{10}$	$5.1084 \cdot 10^{10}$

algorithm as well as restarted multi-shift GMRES [11] (rest_msGMRES), which can be seen as a nested combination of multi-shift GMRES with an outer Richardson iteration. For all considered cases, the seed shift τ of the shift-and-invert preconditioner (4.8) has been tuned such that optimal convergence was obtained.

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 $\sigma_k = 2\pi f_k, k = 1, \ldots, N_\sigma$. The scattering of time-harmonic waves is described in [2] by a Navier equation,

$$-\sigma_k^2 \rho(\mathbf{x})\mathbf{u} - \nabla \cdot \tau(\mathbf{u}) = \mathbf{s}, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^2, \tag{4.10}$$

where $\mathbf{u} : \Omega \to \mathbb{R}^2$ 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 tensors are derived from Hooke's law and are given by

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

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 2.

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

$$i\gamma(\mathbf{x})\sigma_k\rho(\mathbf{x})B\mathbf{u} + \tau(\mathbf{u})\mathbf{n}(\mathbf{x}) = \mathbf{0}, \quad \mathbf{x} \in \partial\Omega, \tag{4.11}$$

where γ is the absorption coefficient, $i \equiv \sqrt{-1}$, and $B_{i,j}(\mathbf{x}) \equiv c_p(\mathbf{x})n_i n_j + c_s(\mathbf{x})t_i t_j$. 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 at 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 2). In the following, we prescribed absorbing boundary conditions on whole $\partial\Omega$ by setting $\gamma \equiv 1$.

From a discretization of (4.10)–(4.11) using linear finite elements, we obtain the linear systems

$$(K + i\sigma_k C - \sigma_k^2 M)\underline{\mathbf{u}} = \underline{\mathbf{s}}, \quad k = 1, \ldots, N_\sigma, \tag{4.12}$$

with K, C, M being symmetric and sparse block matrices and $\underline{\mathbf{u}}, \underline{\mathbf{s}}$ being the discretized counterpart of $\mathbf{u}, \mathbf{s}$ in lexicographical order. Here, C contains the boundary conditions (4.11), and K and M are a stiffness and mass matrix, respectively. Re-formulation of (4.12) in the same way as (4.7) yields a block system of the form

$$\left\{ \begin{bmatrix} iC & K \\ I & 0 \end{bmatrix} - \sigma_k \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \right\} \begin{bmatrix} \sigma_k\underline{\mathbf{u}} \\ \underline{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{s}} \\ 0 \end{bmatrix}, \quad k = 1, \ldots, N_\sigma, \tag{4.13}$$

which is again a shifted linear system with shifts $\sigma_1, \ldots, \sigma_{N_\sigma}$.

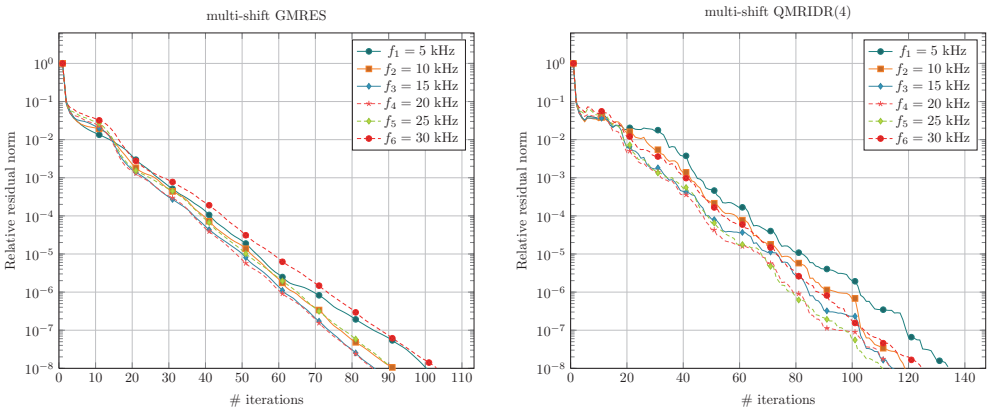

FIG. 4. *Convergence behavior of multi-shift GMRES (left) and multi-shift QMRIDR*(4) *(right) for* (4.13).

TABLE 3

Comparison of multi-shift and nested multi-shift algorithms for the linear elastic wave equation (4.10)–(4.11). *Again,* $\hat{\sigma} \equiv \max\{\sigma_k, k = 1, \ldots, N_\sigma\}$.

	Multi-shift Krylov methods			
	msGMRES	rest_msGMRES	msQMRIDR(4)	collIDR(4)
# inner iterations	-	20	-	-
# outer iterations	103	7	136	134
Seed shift τ	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$
CPU time	17.71s	6.13s	22.35s	22.58s
	Nested multi-shift Krylov methods			
	FOM-FGMRES	IDR(4)-FGMRES	FOM-FQMRIDR(4)	IDR(4)-FQMRIDR(4)
# inner iterations	20	25	30	30
# outer iterations	7	9	5	15
Seed shift τ	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$	$(0.7 - 0.7i)\hat{\sigma}$
CPU time	9.12s	32.99s	8.14s	58.36s

The considered numerical setting is taken from [2, 18]. Therein, the parameters are set to the values presented in Table 2, 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 $f_1 = 5,000$ Hz to $f_6 = 30,000$ Hz in uniform steps. For more details on the discretization and the numerical results, we refer to our technical report [3].

We again ran our numerical experiments using an additional shift-and-invert preconditioner (4.9) with seed shift τ as shown in Table 3. In Figure 4, we present the convergence curves of multi-shift GMRES and multi-shift QMRIDR(4) without nested preconditioning. In this experiment, we observe a flat convergence behavior that gives rise to an early truncation in the nested framework.

For nested FOM-FGMRES, we chose the number of inner msFOM iterations such that the relative residual norms are below a threshold tolerance of 0.1; cf. Figure 5. Our numerical experiments have proved that this leads to a rapid convergence in only 7 iterations in the outer FGMRES 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-FGMRES, we observe a speed-up of almost two; cf. Table 3.

Moreover, we applied nested IDR-FQMRIDR(4) to (4.13) in Figure 6. From the convergence behavior of the inner collIDR iteration (Algorithm 2), we note that the

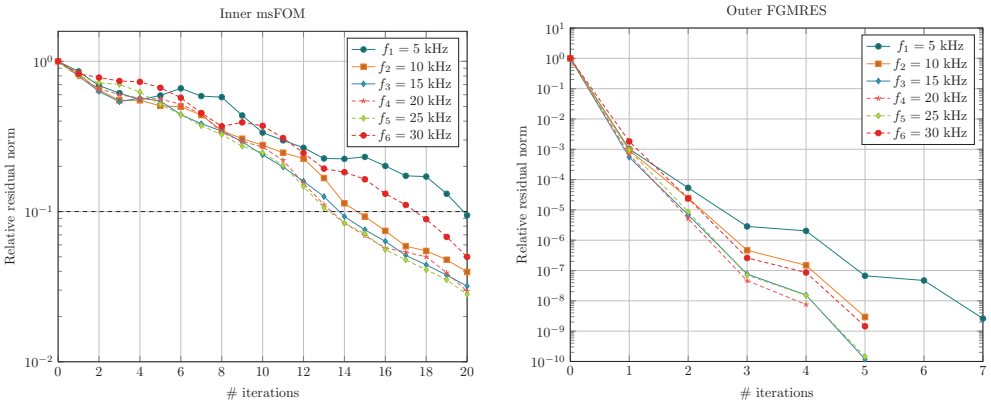

FIG. 5. *Convergence behavior of FOM-FGMRES for* (4.13)*: typical inner convergence (left) and outer convergence (right).*

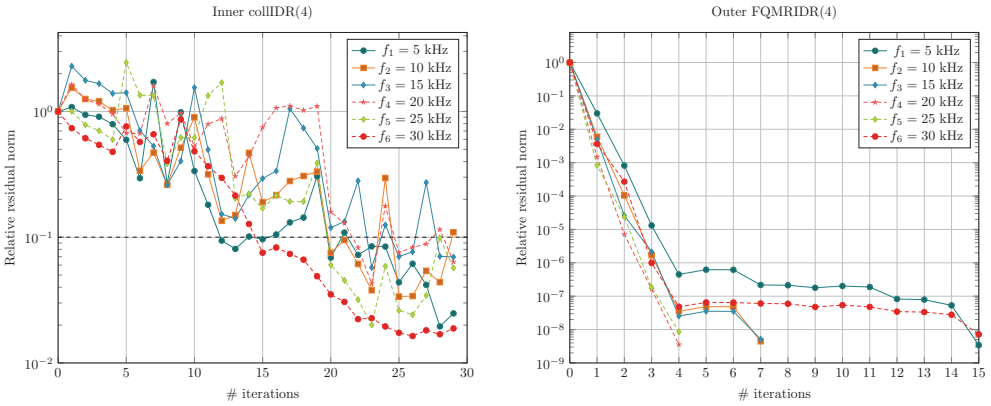

FIG. 6. *Convergence behavior of IDR-FQMRIDR* (4) *for* (4.13)*: typical inner convergence (left) and outer convergence (right).*

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 short recurrence of QMRIDR(s).

In Table 3 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 3, we first note that QMRIDR(4) and collIDR(4) require similar CPU times. Moreover, we observe that a combination of inner msFOM and outer FQMRIDR(4) perform best among the nested algorithms. The restarted version of multi-shift GMRES performs best in this setting but did not converge in some examples described in [3].

5. Conclusion. This work presents an algorithmic framework for the numerical solution of shifted linear systems (1.1) with inner-outer Krylov methods that allow

flexible preconditioning. In this sense, it can be seen as a generalization of the work of [22, 29, 40] 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 (3.7).
2. The collinearity factor is used in the jth iteration of an *outer* Krylov method in order to derive the Hessenberg matrix of the shifted systems; cf. (3.18), (3.22).

We call this new framework a *nested* Krylov method for shifted linear systems since the inner Krylov iteration is considered a flexible preconditioner for the outer Krylov method. Moreover, the fact that we can use a Krylov method as flexible preconditioner shows that a Krylov polynomial can be used as a polynomial preconditioner in the sense of [1].

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 FGMRES in Algorithm 3. Therefore, the collinearity factor for the inner Krylov method (multi-shift FOM) is given by (3.9) without any further manipulations. When combining multi-shift IDR(s) and FQMRIDR(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 (3.16); cf. Algorithm 2. In both cases, a shifted Hessenberg matrix has been derived using the respective collinearity factors. This has been done for FGMRES (3.19) and FQMRIDR(s) (3.23), 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 $\|\mathbf{r}_j^{(\sigma_k)}\|/\|\mathbf{r}_0^{(\sigma_k)}\| < 0.1$ at every (outer) iteration j and for all shifts $\sigma_1, \ldots, \sigma_{N_\sigma}$. This way, we were able to obtain a computational speed-up up to a factor of five when comparing multi-shift GMRES with nested FOM-FGMRES in section 4.1.

Acknowledgment. We want to thank René-Édouard Plessix 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, Philadelphia, 2013.

[2] T. Airaksinen, A. Pennanen, and J. Toivanen, *A damping preconditioner for time-harmonic wave equations in fluid and elastic material*, J. Comput. Phys., 228 (2009), pp. 1466–1479.

[3] M. Baumann and M.B. van Gijzen, *Nested Krylov Methods for Shifted Linear Systems*, DIAM report 14-01, TU Delft, 2014.

[4] M. Benzi and D. Bertaccini, *Approximate inverse preconditioning for shifted linear systems*, BIT, 43 (2003), pp. 231–244.

[5] L. Du, T. Sogabe, and S.-L. Zhang, *IDR(s) for solving shifted nonsymmetric linear systems*, J. Comput. Appl. Math., 274 (2015), pp. 35–43.

[6] Y.A. Erlangga, C.W. Oosterlee, and C. Vuik, *A novel multigrid based preconditioner for heterogeneous Helmholtz problems*, SIAM J. Sci. Comput., 27 (2006), pp. 1471–1492.

[7] Y.A. Erlangga, C. Vuik, and C.W. Oosterlee, *On a class of preconditioners for solving the Helmholtz equation*, Appl. Numer. Math., 50 (2004), pp. 409–425.

[8] Y.A. Erlangga, C. Vuik, and C.W. Oosterlee, *Comparison of multigrid and incomplete LU shifted-Laplace preconditioners for the inhomogeneous Helmholtz equation*, Appl. Numer. Math., 56 (2006), pp. 648–666.

[9] R. Freund, *Solution of shifted linear systems by quasi-minimal residual iterations*, in Numerical Linear Algebra, deGruyter, Berlin, 1993, pp. 101–121.

[10] A. Frommer, *BiCGStab(ℓ) for families of shifted linear systems*, Computing, 70 (2003), pp. 87–109.
[11] A. Frommer and U. Glässner, *Restarted GMRES for Shifted Linear Systems*, SIAM J. Sci. Comput., 19 (1998), pp. 15–26.
[12] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins Stud. Math. Sci., Johns Hopkins University Press, Baltimore, 1996.
[13] G.-D. Gu, X.L. Zhou, and L. Lin, *A flexible preconditioned Arnoldi method for shifted linear systems*, J. Comput. Math., 25 (2007), pp. 522–530.
[14] M.H. Gutknecht and J.-P.M. Zemke, *Eigenvalue computations based on IDR*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 283–311.
[15] B. Jegerlehner, *Krylov Space Solvers for Shifted Linear Systems*, ar Xiv:hep-lat/9612014, 1996.
[16] M.E. Kilmer and E. de Sturler, *Recycling subspace information for diffuse optical tomography*, SIAM J. Sci. Comput., 27 (2006), pp. 2140–2166.
[17] S. Kirchner, *IDR-Verfahren zur Lösung von Familien geshifteter linearer Gleichungssysteme*, M.S. thesis, Bergische Universität Wuppertal, 2011.
[18] K. Meerbergen, *The solution of parametrized symmetric linear systems*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 1038–1059.
[19] K. Meerbergen and Z. Bai, *The Lanczos method for parameterized symmetric linear systems with multiple right-hand sides*, SIAM J. Matrix Anal. Appl., 4 (2010), pp. 1642–1662.
[20] J. Meng, P. Zhu, and H. Li, *QMRCGstab algorithm for families of shifted systems*, in Proceedings of the Ninth International Conference on Computational Intelligence and Security, 2013.
[21] R.-E. Plessix and W.A. Mulder, *Separation-of-variables as a preconditioner for an iterative Helmholtz solver*, Appl. Numer. Math., 44 (2003), pp. 383–400.
[22] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.
[23] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[24] Y. Saad and M. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[25] 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 (2013), pp. 3001–3023.
[26] A. H. Sheikh, D. Lahaye, and C. Vuik, *A Scalable Helmholtz Solver Combining the Shifted Laplace Preconditioner with Multigrid Deflation*, DIAM report 11-01, TU Delft, 2011.
[27] V. Simoncini, *Restarted full orthogonalization method for shifted linear systems*, BIT, 43 (2003), pp. 459–466.
[28] 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 (2002), pp. 1875–1897.
[29] V. Simoncini and D. Szyld, *Flexible inner-outer Krylov subspace methods*, SIAM J. Numer. Anal., 40 (2003), pp. 2219–2239.
[30] V. Simoncini and D. Szyld, *Recent computational developments in Krylov subspace methods for linear systems*, Numer. Linear Algebra Appl., 14 (2007), pp. 1–59.
[31] 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 (2008), pp. 1035–1062.
[32] K. Soodhalter, *Two recursive GMRES-type methods for shifted linear systems with general preconditioning*, arXiv:1403.4428, 2014.
[33] K. Soodhalter, D.B. Szyld, and F. Xue, *Krylov subspace recycling for sequences of shifted linear systems*, Appl. Numer. Math., 81 (2014), pp. 105–118.
[34] L.N. Trefethen and D. Bau, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
[35] J. van den Eshof and G.L.G. Sleijpen, *Accurate conjugate gradient methods for families of shifted systems*, Appl. Numer. Math., 49 (2004), pp. 17–37.
[36] J. van den Eshof, G.L.G. Sleijpen, and M.B. van Gijzen, *Relaxation strategies for nested Krylov methods*, Int. J. Comput. Appl. Math., 177 (2005), pp. 347–365.
[37] 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 (2007), pp. 1942–1958.
[38] M.B. van Gijzen, G.L.G. Sleijpen, and J.-P.M. Zemke, *Flexible and multi-shift induced dimension reduction algorithms for solving large sparse linear systems*, Numer. Linear Algebra Appl., 22 (2015), pp. 1–25.

[39] M.B. van Gijzen and P. Sonneveld, *An elegant IDR(s) variant that efficiently exploits biorthogonality properties*, ACM Trans. Math. Software, 38 (2011), pp. 5:1–5:19.
[40] C. Vuik and H. van der Vorst, *GMRESR: A family of nested GMRES methods*, Numer. Linear Algebra Appl., 1 (1994), pp. 369–386.
[41] G. Wu, Y.-C. Wang, and X.-Q. Jin, *A preconditioned and shifted GMRES algorithm for the Pagerank problem with multiple damping factors*, SIAM J. Sci. Comput., 34 (2012), pp. 2558–2575.