SIAM J. MATRIX ANAL. APPL.
Vol. 34, No. 2, pp. 283–311

EIGENVALUE COMPUTATIONS BASED ON IDR*

MARTIN H. GUTKNECHT† AND JENS-PETER M. ZEMKE‡

Abstract. The induced dimension reduction (IDR) method, which has been introduced as a transpose-free Krylov space method for solving nonsymmetric linear systems, can also be used to determine approximate eigenvalues of a matrix or operator. The IDR residual polynomials are the products of a residual polynomial constructed by successively appending linear smoothing factors and the residual polynomials of a two-sided (block) Lanczos process with one right-hand side and several left-hand sides. The Hessenberg matrix of the OrthoRes version of this Lanczos process is explicitly obtained in terms of the scalars defining IDR by deflating the smoothing factors. The eigenvalues of this Hessenberg matrix are approximations of eigenvalues of the given matrix or operator.

Key words. Krylov space method, iterative method, induced dimension reduction, large nonsymmetric eigenvalue problem

AMS subject classifications. Primary, 65F15; Secondary, 65F10, 65F50

DOI. 10.1137/100804012

1. Introduction.

Induced dimension reduction (IDR) is a general concept for defining certain Krylov subspace methods for solving systems of linear equations. While the rationale behind the IDR approach differs considerably from other approaches to Krylov subspace solvers, the resulting methods are mathematically not much different from well-known Krylov methods. In particular, as we will show here, there is also the possibility of extracting eigenvalue information from the recurrence coefficients constructed in IDR methods.

1.1. The IDR approach.

The first IDR method was developed by Sonneveld and published in 1980 in Sections 4 and 5 of the proceedings paper [42] written jointly with Wesseling. It is nearly mathematically equivalent to the BiCGStab method introduced ten years later by van der Vorst and Sonneveld [39] and by van der Vorst [40]. Consider $\mathbf{A}\mathbf{x} = \mathbf{b}$ where $\mathbf{A}$ is a general complex $N \times N$ matrix and $\mathbf{b} \in \mathbb{C}^N$. As for any standard Krylov subspace method, the approximate solutions generated by IDR satisfy

$$\mathbf{x}_n \in \mathbf{x}_0 + \mathcal{K}_n(\mathbf{A}, \mathbf{r}_0)\,, \tag{1.1}$$

where $\mathbf{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$ is the initial residual and

$$\mathcal{K}_n := \mathcal{K}_n(\mathbf{A}, \mathbf{r}_0) := \mathsf{span}\,\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \ldots, \mathbf{A}^{n-1}\mathbf{r}_0\}$$

is the nth Krylov subspace generated by $\mathbf{A}$ from $\mathbf{r}_0$. Clearly, $\mathcal{K}_n \subseteq \mathbb{C}^N$, and, if the data are real, $\mathcal{K}_n \subseteq \mathbb{R}^N$. Relation (1.1) implies that for the nth residual $\mathbf{r}_n := \mathbf{b} - \mathbf{A}\mathbf{x}_n$ there holds

$$\mathbf{r}_n \in \mathbf{r}_0 + \mathbf{A}\mathcal{K}_n(\mathbf{A}, \mathbf{r}_0) \subset \mathcal{K}_{n+1}(\mathbf{A}, \mathbf{r}_0)\,. \tag{1.2}$$

*Received by the editors July 30, 2010; accepted for publication (in revised form) by G.L.G. Sleijpen December 14, 2012; published electronically April 4, 2013.
http://www.siam.org/journals/simax/34-2/80401.html

†Seminar for Applied Mathematics, ETH Zurich, CH-8092 Zurich, Switzerland (mhg@math.ethz.ch). Work started while this author was visiting the TU Berlin, supported by the DFG Forschungszentrum MATHEON and the Mercator Visiting Professorship Program of the DFG.

‡Institut für Mathematik, Technische Universität Hamburg-Harburg, D-21073 Hamburg, Germany (zemke@tu-harburg.de). Part of this work was carried out while the author was visiting Kyushu University in Fukuoka on invitation of Prof. Seiji Fujino and the University of Tokyo.

In the original IDR, it must additionally satisfy

$$\mathbf{r}_n \in \mathcal{G}_j\,, \qquad \text{where} \quad j := \left\lfloor \frac{n}{2} \right\rfloor \tag{1.3}$$

and where the Sonneveld spaces $\mathcal{G}_j$ [33] are here defined by a recurrence of the form

$$\mathcal{G}_j := (\mathbf{I} - \omega_j \mathbf{A})(\mathcal{G}_{j-1} \cap \mathcal{S})\,. \tag{1.4}$$

Here, $\mathcal{S}$ is a hyperplane of $\mathbb{C}^N$, and the parameters ω_j are normally chosen such that the 2-norm of $\mathbf{r}_{2j}$ is minimized within certain constraints. At the beginning, $\mathcal{G}_0 = \mathcal{K}_N$, but no basis of this invariant Krylov subspace is needed. It turns out that the spaces $\mathcal{G}_j$ are nested: $\mathcal{G}_{j+1} \subset \mathcal{G}_j$, and that under mild restrictions on the hyperplane $\mathcal{S}$, in a finite number of steps they reduce to the null space, so that $\mathbf{r}_n = \mathbf{o}$. (The latter is our notation for a zero vector.) It can be shown that the even-indexed IDR residuals are BiCGStab residuals: $\mathbf{r}_{2j} = \mathbf{r}_j^{\text{BiCGStab}}$. In particular, $\Omega_j(z) := (1-\omega_1 z)\cdots(1-\omega_j z)$ is again a factor of the residual polynomial $\mathcal{R}_{2j}$ associated with $\mathbf{r}_{2j}$.

Recently, Sonneveld and van Gijzen [34] considerably generalized and improved the original IDR method. In their IDR(s) the subset $\mathcal{S}$ is a subspace of codimension s, and the indices n and j in (1.3) are linked by $j := \lfloor n/(s+1) \rfloor$. This method can be seen to be related to the ML(k)BiCGStab method with $k = s$ of Yeung and Chan [43] and thus to the nonsymmetric band Lanczos process with one right-hand side starting vector and s left-hand side starting vectors [2]. In the case $s = 1$, the IDR(s) prototype algorithm of [34] differs slightly from the original IDR of [42], but again $\mathbf{r}_{2j} = \mathbf{r}_j^{\text{BiCGStab}}$. Hence, it is still essentially equivalent to BiCGStab. If $s > 1$, there is additional freedom in the choice of the "intermediate" residuals $\mathbf{r}_n$ with $n \neq (s+1)j$. This has been capitalized upon in the IDRBiO variant of IDR(s) described in [41]. Reformulations of the IDR approach have been considered in [32]. Modifications of the basic recurrence (1.4) have led to further similar methods such as IDRStab by Sleijpen and van Gijzen [33], and GIDR(s, L) and GBiCGStab(s, L) by Tanio and Sugihara [37, 38]. We group the methods described in [33, 38] under the common acronym IDRStab, as they differ only in implementation aspects. For an expository description of IDR(s) and IDRBiO, see [18].

The convergence behavior of IDR methods is largely not understood. A stochastic analysis of basic IDR variants which relates the convergence behavior to the one of GMRes has recently been given by Sonneveld [35]. A step towards understanding the behavior also in finite precision is our investigation of the relation of IDR to two-sided Lanczos processes.

It is straightforward to extract partial eigenvalue information from a run of BiCGStab, since this method explicitly determines the recurrence coefficients of a nonsymmetric Lanczos process and, thus, of a tridiagonal "projection" of $\mathbf{A}$. In this paper we show how the same eigenvalue information can be extracted from a run of IDR(1), and we investigate the generalization of this eigensolver to the case $s > 1$. We only cover the prototype method of [34], but our approach carries over to other members of the IDR family mentioned above. An extended version of this article with numerical experiments and source codes is our report [14] of 67 pages.

In particular, we consider here the transition from the prototype IDR(s) method to a two-sided (block) Lanczos process with one right-hand side and s left-hand sides in OrthoRes form (defined below in subsection 1.3). The OrthoRes form of this Lanczos process will be denoted for brevity by BiORes($s, 1$), in analogy to the terminology introduced in [4] and [16, 17]. As the IDR(s) variant in [34] has similarities with OrthoRes [44] and is based on a short-term recurrence we denote it by

IDR(s)ORes. The transition from IDR(s)ORes to BiORes($s,1$) is analogous to the corresponding transition from classical IDR [42] or BiCGStab [40] to the BiORes version of BiCG [16, 17]. The first similar transition from a linear equation solver to an eigenvalue solver, namely, from the Hestenes and Stiefel variant of CG, which is the OrthoMin-variant of CG, to the Lanczos method tailored to symmetric matrices was given in the book of Householder [20]. The concurrent solution of linear systems and the corresponding eigenvalue problems has recently attracted interest in connection with using projection based singular preconditioners for deflating unwanted eigenvectors; see, e.g., [1].

1.2. Motivation. There are four reasons to consider eigenvalue computations based on IDR. The obvious reason is to showcase that IDR can be used to compute eigenvalues, and so without the need for the transpose of $\mathbf{A}$ and using recurrences of length $s+1$. The second reason is that the link worked out in [46] between quantities defined in Krylov subspace methods like IDR and interpolation at the computed Ritz values enables us to better understand the convergence of IDR in theory, as well as in finite precision. A third reason is the idea of enhancing existing IDR algorithms in case of slow convergence by utilizing information about the location of the eigenvalues by, e.g., preconditioning based on deflation. This information on the spectrum of $\mathbf{A}$ may especially be useful in a parallel implementation; see the remarks in [47]. It may also be used for the selection of the parameters ω_j; see Simoncini and Szyld [31]. Last but not least, it deepens our understanding of the interrelations between IDR(s), ML(k)BiCGStab, and two-sided (block) Lanczos processes with one right-hand side and several left-hand sides. The price to pay for focusing on all these reasons is that this paper is a rather technical one; subsequent developments are simplified using the techniques presented here.

For the second and third reason and to simplify the presentation, we only consider IDR(s)ORes, the most basic variant of IDR(s). Alternatively, we could have rewritten the recurrences to better suit eigenvalue computations, e.g., by normalizing the residuals, which are only used as *basis vectors* if we do not compute approximate solutions of linear systems. As IDR(s)ORes is of type OrthoRes, we assume that zero is well separated from the field of values, which implies that $\mathbf{A}$ is not too badly conditioned. For the sake of brevity and clarity we only show *how* to compute approximate eigenpairs; the convergence theory and the error analysis of different IDR variants (and the variants of IDR(s)Eig based on it) will be published separately.

1.3. Notation and preliminaries. We use standard notation, and we aim at making it as simple as possible. The identity matrix of size $s \times s$ is denoted by $\mathbf{I} = \mathbf{I}_s$, its column vectors by $\mathbf{e}_j$, and its elements by the Kronecker delta δ_{ij}. The vector of the sums of all columns, i.e., the vector of all ones, is denoted by $\mathbf{e}$. The matrix $\mathbf{O} = \mathbf{O}_s$ denotes the zero matrix of size $s \times s$, the zero column vector of length n is denoted by $\mathbf{o} = \mathbf{o}_n$. The matrix $\mathbf{N}_k$ denotes the nilpotent upshift matrix of size $k \times k$ with elements $\delta_{i,j-1}$. The sizes are omitted if easily deducible from the context. In step $n > s$ of IDR(s)ORes, $\mathbf{R}_{n-s:n}$ denotes the matrix of the last $s+1$ residual vectors,

$$\mathbf{R}_{n-s:n} := \big(\mathbf{r}_{n-s}, \ldots, \mathbf{r}_n\big). \tag{1.5}$$

The matrix of all residual vectors up to step n has $n+1$ columns and is denoted by

$$\mathbf{R}_{n+1} := \big(\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_n\big) = \mathbf{R}_{0:n}. \tag{1.6}$$

We denote the space of polynomials of degree at most n by $\mathbb{P}_n$, and we let $\mathbb{P}_n^\circ$ be the subset of those polynomials that take at 0 the value 1. A superscript † denotes the Moore–Penrose pseudoinverse.

We will repeatedly consider unreduced "extended" Hessenberg matrices that have an extra row at the bottom with a nonzero element only at the end. Let $\mathbf{H}_n \in \mathbb{C}^{n\times n}$ denote an unreduced Hessenberg matrix. Then $\underline{\mathbf{H}}_n \in \mathbb{C}^{(n+1)\times n}$ is used to denote a corresponding unreduced *extended* Hessenberg matrix. The extended identity matrix $\underline{\mathbf{I}}_n$ denotes $\mathbf{I}_n$ with an additional *zero* row at the bottom. Here, most of the extended Hessenberg matrices will have the property that their column sums are zero. We say that these matrices are of OrthoRes-type, since each one defines a Krylov subspace method of the general OrthoRes form [44]. Extended Hessenberg matrices of OrthoRes-type are denoted by appending a superscript $\circ$, like $\underline{\mathbf{Y}}_n^\circ$, $\underline{\mathbf{S}}_n^\circ$, or $\underline{\mathbf{L}}_n^\circ$. The property that the columns of $\underline{\mathbf{S}}_n^\circ$ sum to zero is reflected by $\mathbf{e}^\mathsf{T}\underline{\mathbf{S}}_n^\circ = \mathbf{o}_n^\mathsf{T}$. Such a matrix has an LDMT decomposition $\underline{\mathbf{S}}_n^\circ = \underline{\mathbf{E}}_n^\circ \mathbf{D}_n \mathbf{M}_n^\mathsf{H}$ with an extended bidiagonal unit lower triangular matrix $\underline{\mathbf{E}}_n^\circ \in \mathbb{C}^{(n+1)\times n}$ of OrthoRes-type, a diagonal matrix $\mathbf{D}_n \in \mathbb{C}^{n\times n}$, and a conjugate transposed unit lower triangular matrix $\mathbf{M}_n \in \mathbb{C}^{n\times n}$, where

$$\underline{\mathbf{E}}_n^\circ := \begin{pmatrix} 1 & & \\ -1 & \ddots & \\ & \ddots & 1 \\ & & -1 \end{pmatrix}, \qquad \mathbf{D}_n := -\mathsf{diag}\,(s_{2,1}^\circ, s_{3,2}^\circ \ldots, s_{n+1,n}^\circ). \tag{1.7}$$

This has two interesting implications: The determinants of the leading submatrices of $-\mathbf{S}_n^\circ$ are all nonsingular and given by

$$\det(-\mathbf{S}_k^\circ) = \det(-\mathbf{D}_k) = \prod_{j=1}^{k} s_{j+1,j}^\circ =: s_{1:k}^\circ, \quad 1 \leqslant k \leqslant n, \tag{1.8}$$

and the inverses of $\mathbf{S}_k^\circ$ are highly structured,

$$\mathbf{e}_j^\mathsf{T}(\mathbf{S}_k^\circ)^{-1}\mathbf{e}_j = \mathbf{e}_j^\mathsf{T}(\mathbf{S}_k^\circ)^{-1}\mathbf{e}_i = \mathbf{e}_j^\mathsf{T}(\mathbf{S}_k^\circ)^{-1}\mathbf{e}_1, \quad 1 \leqslant i \leqslant j \leqslant k \leqslant n. \tag{1.9}$$

The last equation is verified by noting that in $(\mathbf{S}_k^\circ)^{-1} = (\mathbf{M}_k^\mathsf{H})^{-1}(\mathbf{D}_k)^{-1}(\mathbf{E}_k^\circ)^{-1}$ the first factor, $(\mathbf{M}_k^\mathsf{H})^{-1}$, is upper triangular, the second, $(\mathbf{D}_k)^{-1}$, is diagonal, and $(\mathbf{E}_k^\circ)^{-1}$ is the lower triangular matrix of ones.

Associated with the residual $\mathbf{r}_n$ there is a residual polynomial $\mathcal{R}_n \in \mathbb{P}_n^\circ$ satisfying $\mathbf{r}_n = \mathcal{R}_n(\mathbf{A})\mathbf{r}_0$; see [36]. Whenever a Krylov space solver is of type OrthoRes, $\mathcal{R}_n$ has exactly degree n and is uniquely determined as long as $\mathcal{K}_{n+1}$ has dimension $n+1$. Moreover, in this case there is a corresponding Hessenberg decomposition[1] (of OrthoRes-type)

$$\mathbf{A}\mathbf{R}_n = \mathbf{R}_{n+1}\,\underline{\mathbf{S}}_n^\circ \tag{1.10}$$

with an extended unreduced Hessenberg matrix $\underline{\mathbf{S}}_n^\circ$ of size $(n+1)\times n$. This formula summarizes the recurrences for the residuals; it is mirrored by the formula

$$z\big(\mathcal{R}_0(z), \ldots, \mathcal{R}_n(z)\big) = \big(\mathcal{R}_0(z), \ldots, \mathcal{R}_{n+1}(z)\big)\,\underline{\mathbf{S}}_n^\circ \tag{1.11}$$

[1] We name these relations in honor of Karl Hessenberg. He was to our knowledge the first who considered relations of the type $\mathbf{A}\mathbf{Q}_n = \mathbf{Q}_{n+1}\underline{\mathbf{H}}_n$ with a special unreduced extended Hessenberg matrix $\underline{\mathbf{H}}_n$; see [19]. The names of Lanczos [21, 22] and Arnoldi [3] are associated with special such relations.

describing the recurrences for the residual polynomials. It is well known (see, e.g., [10, section 7.11, p. 252, eq. (8)] or, the probably earliest reference [30, Erste Abtheilung, IV. Abschnitt, section 154, Seite 361, Gleichung (560)]), that $\mathcal{R}_n$ is up to scaling the characteristic polynomial of the $n \times n$ Hessenberg matrix $\mathbf{S}_n^\circ$ and can be expressed as

$$\mathcal{R}_n(z) = \mathsf{det}\,(\mathbf{I}_n - z(\mathbf{S}_n^\circ)^{-1}) = \frac{\mathsf{det}\,(\mathbf{S}_n^\circ - z\mathbf{I}_n)}{\mathsf{det}\,(\mathbf{S}_n^\circ)}. \tag{1.12}$$

In [46] it was shown that this representation cum grano salis carries over to the finite precision case, i.e., modulo additional polynomial error terms.

In sections 3 and 4 the recurrences of IDR(s) will be seen to yield instead of (1.10) a *generalized* Hessenberg decomposition of the form

$$\mathbf{A}\mathbf{Q}_n\mathbf{U}_n = \mathbf{Q}_{n+1}\underline{\mathbf{H}}_n \tag{1.13}$$

with an extended unreduced Hessenberg matrix $\underline{\mathbf{H}}_n^\circ$ and an upper triangular matrix $\mathbf{U}_n$. Such decompositions will be investigated in section 2. For theoretical purposes we may multiply this decomposition from the right-hand side by $\mathbf{U}_n^{-1}$ in case of nonsingular $\mathbf{U}_n$ and define $\underline{\mathbf{S}}_n^\circ := \underline{\mathbf{H}}_n^\circ\mathbf{U}_n^{-1}$ to return to (1.10). However, the point is that both $\underline{\mathbf{H}}_n^\circ$ and $\mathbf{U}_n$ will be banded due to the short recurrences of IDR(s), while the corresponding $\underline{\mathbf{S}}_n^\circ$ will not be. We will repeatedly make use of the fact that the spectrum of the square matrix $\mathbf{S}_n^\circ$ is the same as the spectrum of the pencil $(\mathbf{H}_n^\circ, \mathbf{U}_n)$. We will also encounter examples where $\mathbf{U}_n$ is singular and thus the decomposition can not be transformed in this manner to (1.10).

Often, the residuals $\mathbf{r}_n$ of a Krylov subspace solver satisfy a Bubnov–Galërkin or a Petrov–Galërkin condition. For example, those of the biconjugate gradient (BiCG) method [22, 9], which is closely related to the original IDR and the IDR(1) methods, are characterized by

$$\mathbf{r}_n \in \mathbf{r}_0 + \mathbf{A}\mathcal{K}_n\,, \qquad \mathbf{r}_n \perp \widetilde{\mathcal{K}}_n := \mathcal{K}_n(\mathbf{A}^{\mathsf{H}}, \widetilde{\mathbf{r}}_0)\,,$$

where $\widetilde{\mathbf{r}}_0 \in \mathbb{C}^N$ is an initial shadow residual that can be chosen nearly arbitrarily. The shadow space $\widetilde{\mathcal{K}}_n := \mathcal{K}_n(\mathbf{A}^{\mathsf{H}}, \widetilde{\mathbf{r}}_0)$ is the nth Krylov subspace generated by the adjoint $\mathbf{A}^{\mathsf{H}}$, i.e., the complex conjugate transpose of $\mathbf{A}$, from $\widetilde{\mathbf{r}}_0$.

In the ML(s)BiCG method[2] that Yeung and Chan [43] proposed as a theoretical tool for deriving their ML(s)BiCGStab method, this shadow space is replaced by a block Krylov space. ML(s)BiCG is not exactly a generalization of BiCG since it does not construct a pair of biorthogonal or block biorthogonal bases, but the nonsymmetric band Lanczos process [2, 12, 13, 11] or the block Lanczos process [5, 23] can be adapted to yield such generalizations of BiCG; see section 3.3 of Loher [23]. In particular, we could develop a generalization of the BiORes version of BiCG that is directly based on recurrences for residuals satisfying relations similar to those of ML(s)BiCG—if the method does not break down early. The aforementioned method BiORes($s, 1$) is a more general scheme, also based on s left starting vectors (shadow residuals) and only one right starting vector, the initial residual $\mathbf{r}_0$. In BiORes($s, 1$) we allow far more flexibility as in ML(s)BiCG.

1.4. Outline. In addition to the IDR residuals $\mathbf{r}_n$ we will encounter in this paper other related sets of residuals and the corresponding residual polynomials. We will also use the Hessenberg pencils and Hessenberg matrices associated with the

[2] Yeung and Chan [43] call the parameter k, but we are interested in the case where $k = s$.

recurrences for both the residual vectors and the residual polynomials. The transition from IDR(s)ORes to BiORes($s,1$) proceeds in two steps, a *purification step* and a *deflation step*. In the purification step, the $\lfloor n/(s+1)\rfloor$ known roots ω_k^{-1} of the IDR residual polynomial are moved to infinity. The size of the matrices is not reduced, and the $\lfloor n/(s+1)\rfloor$-fold infinite eigenvalue prevents the existence of a corresponding Hessenberg matrix. The deflation step can be split up into two substeps: first, a basis transformation is applied so that the infinite eigenspace is spanned by standard unit vectors. Then this eigenspace is eliminated, and hereby the problem size is reduced.

The various quantities will be defined later; an overview of all of them is given in the appendix. In particular, Table A.1 lists the residuals that play a role in this paper together with the corresponding extended Hessenberg pencils and extended Hessenberg matrices, and Table A.2 lists the residual polynomials together with the corresponding residuals and some relations between them. Moreover, in Figure A.1 we depict, in a small example, which entries of the Hessenberg pencils are modified in these transitions.[3]

The paper is organized as follows. In section 2 we define a refinement of the concept of Hessenberg decompositions necessary for the treatment of IDR(s). In section 3 we consider the application of IDR(s) to eigenvalue computations for the case $s=1$. In section 4 we generalize and extend the results for $s=1$ to the general case $s \geqslant 1$. And in section 5 we indicate how to obtain eigenvector approximations and an estimator on the accuracy of the approximate eigenpair. An interesting, partly surprising numerical experiment is presented in section 6. In this example we apply 50 sweeps with $s=4$ to a linear system of size 236 and obtain useful spectral information despite the fact that the residual has not yet converged at all. Finally, before we come to the appendix, we draw some conclusions.

2. Generalized Hessenberg decompositions. We define a generalized Hessenberg decomposition to denote a matrix equation of the type (1.13). When $\mathbf{U}_n = \mathbf{I}_n$, this definition collapses to the definition of a Hessenberg decomposition. We remark that here $\mathbf{Q}_{n+1} \in \mathbb{C}^{N\times(n+1)}$ denotes a generic matrix of basis vectors.

Equation (1.13) corresponds in the case of full rank of $\mathbf{Q}_{n+1}$ and $\mathbf{U}_n$ to an oblique Petrov–Galërkin projection of the pencil $(\mathbf{A},\mathbf{I})$, since with $\widehat{\mathbf{Q}}_n^{\mathsf{H}} := \underline{\mathbf{I}}_n^{\mathsf{T}}\mathbf{Q}_{n+1}^{\dagger}$ we have

$$\widehat{\mathbf{Q}}_n^{\mathsf{H}}(\mathbf{A},\mathbf{I})\mathbf{Q}_n\mathbf{U}_n = \widehat{\mathbf{Q}}_n^{\mathsf{H}}(\mathbf{Q}_{n+1}\underline{\mathbf{H}}_n, \mathbf{Q}_n\mathbf{U}_n) = (\underline{\mathbf{I}}_n^{\mathsf{T}}\underline{\mathbf{H}}_n, \mathbf{U}_n) = (\mathbf{H}_n, \mathbf{U}_n). \tag{2.1}$$

If $\mathbf{U}_n$ is singular, we will refer to (1.13) as a *singular projection*.

In the case of IDR(s)ORes the matrix $\mathbf{Q}_n$ is the previously defined $\mathbf{R}_n$, the matrix of all residual vectors. As IDR(s)ORes is a Krylov method, these can be characterized using residual polynomials. We need another expression for the residual polynomials solely based on the matrices $\mathbf{U}_n$, $\underline{\mathbf{H}}_n$ defining the decomposition (1.13). To achieve this, we prove that the columns of matrices $\mathbf{Q}_n$ satisfying a generalized Hessenberg decomposition (1.13) can be described with the aid of determinants of leading principal submatrices of these pencils. The proof is based on a slight generalization of the proofs of [45, Lemma 3.1];[4] see also [7, 8] and [10, section 7.11] and [46, Theorem 2.1], where the case $\mathbf{U}_n = \mathbf{I}_n$ has been treated. We omit the index n for simplicity and consider an unreduced Hessenberg/upper triangular pencil ${}^z\mathbf{H} := z\mathbf{U} - \mathbf{H}$ with unreduced Hessenberg matrix $\mathbf{H} \in \mathbb{C}^{n\times n}$ and upper triangular

[3] Readers may want to look up these tables and the figure occasionally in order not to get lost.
[4] An earlier similar result can be found in [6].

matrix $\mathbf{U} \in \mathbb{C}^{n\times n}$. Similar to [45, eq. (3.1), p. 595] and [46, eq. (1.9), p. 410] we define the scalars $h_{i:j} := \prod_{\ell=i}^{j} h_{\ell+1,\ell}$ and polynomial vectors $\boldsymbol{\nu}(z)$ and $\check{\boldsymbol{\nu}}(z)$: Let ${}^z\mathbf{H}_{i:j}$ denote the principal submatrix of ${}^z\mathbf{H}$ consisting of the elements indexed by rows and columns i to j and define $\chi_{i:j}(z)$, $\boldsymbol{\nu}(z)$, and $\check{\boldsymbol{\nu}}(z)$ by

$$\chi_{i:j}(z) := \begin{cases} \mathsf{det}\,({}^z\mathbf{H}_{i:j}), & 1 \leqslant i \leqslant j \leqslant n, \\ 1, & i-1=j, \end{cases} \tag{2.2}$$

$$\boldsymbol{\nu}(z) := \left(\frac{\chi_{i+1:n}(z)}{h_{i:n-1}}\right)_{i=1}^{n}, \quad \text{and} \quad \check{\boldsymbol{\nu}}(z) := \left(\frac{\chi_{1:j-1}(z)}{h_{1:j-1}}\right)_{j=1}^{n} \tag{2.3}$$

with the usual convention that the empty product is one.

First we generalize [45, Lemma 3.1] to unreduced Hessenberg/upper triangular pencils as already indicated in [45, section 5, p. 605]. This proves among others that if θ is an eigenvalue of the pencil ${}^z\mathbf{H}$, then the vectors $\boldsymbol{\nu}(\theta)$ and $\check{\boldsymbol{\nu}}(\theta)$ are right and left eigenvectors, respectively. The proof is a variation of the proof of [45, Lemma 3.1] and is omitted here.

LEMMA 2.1. *Let $\mathbf{H} \in \mathbb{C}^{n\times n}$ be unreduced Hessenberg and $\mathbf{U} \in \mathbb{C}^{n\times n}$ upper triangular. Denote ${}^z\mathbf{H} := z\mathbf{U} - \mathbf{H}$, and let $\boldsymbol{\nu}(z)$ and $\check{\boldsymbol{\nu}}(z)$ be defined by* (2.2)–(2.3). *Define $\chi_n(z) := \mathsf{det}\,({}^z\mathbf{H}) = \mathsf{det}\,(z\mathbf{U} - \mathbf{H})$. Then*

$$({}^z\mathbf{H})\boldsymbol{\nu}(z) = \mathbf{e}_1 \frac{\chi_n(z)}{h_{1:n-1}}, \quad \check{\boldsymbol{\nu}}(z)^{\mathsf{T}}({}^z\mathbf{H}) = \frac{\chi_n(z)}{h_{1:n-1}}\,\mathbf{e}_n^{\mathsf{T}}. \tag{2.4}$$

Next we generalize [46, Theorem 2.1] to obtain an expression for the columns of $\mathbf{Q}_n$ as polynomials evaluated at $\mathbf{A}$ times the first column $\mathbf{q}_1$.

THEOREM 2.2. *Let the columns of $\mathbf{Q}_{n+1}$ be defined by a generalized Hessenberg decomposition* (1.13) *with unreduced extended Hessenberg matrix $\underline{\mathbf{H}}_n$. Let $\chi_n(z)$ be defined as in Lemma* 2.1. *Then*

$$\mathbf{q}_{n+1} = \frac{\chi_n(\mathbf{A})}{h_{1:n}}\mathbf{q}_1. \tag{2.5}$$

Proof. To prove (2.5), we start with the generalized Hessenberg decomposition (1.13). First we subtract both sides from the trivial equation $z\mathbf{Q}_n\mathbf{U}_n = z\mathbf{Q}_n\mathbf{U}_n$ to introduce a dependency on the variable z,

$$z\mathbf{Q}_n\mathbf{U}_n - \mathbf{A}\mathbf{Q}_n\mathbf{U}_n = \mathbf{Q}_n(z\mathbf{U}_n - \mathbf{H}_n) - \mathbf{q}_{n+1}h_{n+1,n}\mathbf{e}_n^{\mathsf{T}}. \tag{2.6}$$

We multiply (2.6) by the vector

$$\boldsymbol{\nu}(z) = \sum_{k=1}^{n} \mathbf{e}_k\, \mathbf{e}_k^{\mathsf{T}}\boldsymbol{\nu}(z) = \sum_{k=1}^{n} \mathbf{e}_k\, \nu_k(z) \tag{2.7}$$

and utilize (2.4) and $\nu_n(z) \equiv 1$ to obtain

$$\begin{aligned} \sum_{k=1}^{n}\Big(z\mathbf{Q}_n\mathbf{U}_n\mathbf{e}_k\,\nu_k(z) - \mathbf{A}\mathbf{Q}_n\mathbf{U}_n\mathbf{e}_k\,\nu_k(z)\Big) & \\ = \sum_{k=1}^{n}\Big(z\nu_k(z)\mathbf{Q}_n\mathbf{U}_n\mathbf{e}_k - \mathbf{A}\nu_k(z)\mathbf{Q}_n\mathbf{U}_n\mathbf{e}_k\Big) & \\ = \sum_{k=1}^{n}\Big(z\nu_k(z) - \mathbf{A}\nu_k(z)\Big)\mathbf{Q}_n\mathbf{U}_n\mathbf{e}_k = \mathbf{q}_1\,\frac{\chi_n(z)}{h_{1:n-1}} - \mathbf{q}_{n+1}h_{n+1,n}. & \end{aligned} \tag{2.8}$$

Substituting $\mathbf{A}$ for z in the last line of (2.8), which is possible as only *scalar* polynomials occur, gives $\mathbf{o}_n$ on the left-hand side, since the term in parentheses is zero for every k, $k = 1, \ldots, n$. After reformulating the right-hand side of the last line of (2.8) we have proven the theorem. ∎

Thus, whenever we have computed a generalized Hessenberg decomposition with a matrix $\mathbf{Q}_{n+1}$ composed of residual vectors, we have the residual polynomials to hand: they are the characteristic polynomials of the unreduced Hessenberg/upper triangular pencil $(\mathbf{H}_n, \mathbf{U}_n)$ scaled by the product of the off-diagonal elements of $\underline{\mathbf{H}}_n$; see (1.8) and (1.12). The new feature is that when $\mathbf{U}_n$ is singular, the degree of the polynomial is no longer equal to the dimension of the space spanned thus far. We will come back to this point in section 4. A side effect is that the roots of the residual polynomials can be computed as the eigenvalues of the pencil $(\mathbf{H}_n, \mathbf{U}_n)$.

3. The case $s = 1$. If $s = 1$, the recurrences of IDR(s) in [34] are mathematically equivalent to[5]

$$\begin{aligned} \mathbf{v}_{n-1} &:= (1-\gamma_n)\mathbf{r}_{n-1} + \gamma_n \mathbf{r}_{n-2}\,, & \mathbf{x}'_{n-1} &:= (1-\gamma_n)\mathbf{x}_{n-1} + \gamma_n \mathbf{x}_{n-2}\,, \\ \mathbf{r}_n &:= (\mathbf{I} - \omega_j \mathbf{A})\,\mathbf{v}_{n-1}\,, & \mathbf{x}_n &:= \mathbf{x}'_{n-1} + \omega_j \mathbf{v}_{n-1}\,, \end{aligned} \tag{3.1}$$

where $n > 1$, $j = \lfloor n/2 \rfloor$. Here, γ_n has to be chosen such that $\mathbf{v}_{n-1} \perp \mathbf{p}$, where $\mathbf{p}$ is a basis vector of the one-dimensional orthogonal complement of $\mathcal{S}$, so that $\mathbf{v}_{n-1} \in \mathcal{S}$. If n is even, i.e., $j = n/2$, then ω_j is usually chosen such that $\|\mathbf{r}_n\|$ is minimal, but, basically, any nonzero value is acceptable.

For $n = 1$ we may write $\mathbf{r}_1 = (\mathbf{I} - \omega_0 \mathbf{A})\mathbf{r}_0$. Again ω_0 can be chosen arbitrarily nonzero, e.g., such that $\|\mathbf{r}_1\|$ is minimal. By induction it is easily seen from (3.1) that $\mathbf{r}_n = (\mathbf{I} - \omega_j \mathbf{A})\mathbf{v}_{n-1} = \Omega_j(\mathbf{A})\mathbf{w}_n$, where $\Omega_j(z) := (1 - \omega_1 z) \cdots (1 - \omega_j z)$, and that $\mathbf{w}_n \in \mathcal{K}_{n+1-j}$. So, we have

$$\begin{aligned} \mathbf{r}_n = \Omega_j(\mathbf{A})\mathbf{w}_n &= \begin{cases} \Omega_j(\mathbf{A})\rho_j(\mathbf{A})\mathbf{r}_0 & \text{if } n = 2j\,, \\ \Omega_j(\mathbf{A})\widehat{\rho}_{j+1}(\mathbf{A})\mathbf{r}_0 & \text{if } n = 2j+1\,, \end{cases} \\ \mathbf{v}_{n-1} = \Omega_{j-1}(\mathbf{A})\mathbf{w}_n &= \begin{cases} \Omega_{j-1}(\mathbf{A})\rho_j(\mathbf{A})\mathbf{r}_0 & \text{if } n = 2j\,, \\ \Omega_{j-1}(\mathbf{A})\widehat{\rho}_{j+1}(\mathbf{A})\mathbf{r}_0 & \text{if } n = 2j+1\,. \end{cases} \end{aligned} \tag{3.2}$$

Here ρ_j denotes the jth BiCG residual polynomial, which is the characteristic polynomial of the tridiagonal matrix of the Lanczos process, and is often referred to as a Lanczos polynomial, scaled by $\rho_j(0) = 1$, while $\widehat{\rho}_{j+1}$ denotes another residual polynomial, which has degree $j+1$.

Inserting these formulas into $\mathbf{v}_{n-1} = (1-\gamma_n)\mathbf{r}_{n-1} + \gamma_n \mathbf{r}_{n-2}$ we get, after a short calculation, for $n = 2j$ and $n = 2j+1$, respectively,

$$\begin{aligned} \rho_j(z) &:= (1-\gamma_{2j})\,\widehat{\rho}_j(z) + \gamma_{2j}\,\rho_{j-1}(z)\,, \\ \widehat{\rho}_{j+1}(z) &:= (1-\gamma_{2j+1})\,(1-\omega_j z)\,\rho_j(z) + \gamma_{2j+1}\,\widehat{\rho}_j(z) \end{aligned} \qquad (j = 1, 2, \ldots). \tag{3.3}$$

The initial settings are $\rho_0(z) := 1$ and $\widehat{\rho}_1(z) := (1 - \omega_0 z)$.

In matrix-vector notation, due to $\mathbf{w}_{2j} = \rho_j(\mathbf{A})\mathbf{r}_0$ and $\mathbf{w}_{2j+1} = \widehat{\rho}_{j+1}(\mathbf{A})\mathbf{r}_0$, the recurrences (3.3) can alternatively be expressed by

$$\begin{aligned} \mathbf{w}_{2j} &:= (1-\gamma_{2j})\,\mathbf{w}_{2j-1} + \gamma_{2j}\,\mathbf{w}_{2j-2}\,, \\ \mathbf{w}_{2j+1} &:= (1-\gamma_{2j+1})\,(1-\omega_j \mathbf{A})\,\mathbf{w}_{2j} + \gamma_{2j+1}\,\mathbf{w}_{2j-1} \end{aligned} \qquad (j = 1, 2, \ldots). \tag{3.4}$$

[5] Compared to [34] and [18] the index of γ_n has been changed here by 1.

It describes a mixture of a classical Krylov subspace method given by a three-term recurrence and the construction of a new residual based on a weighting process. We can incorporate the latter into the three-term recurrence and thus remove the vectors $\mathbf{w}_{2j+1}$ with odd indices from the recurrence. In the following, this is described using the language of *polynomials*. The general case $s \geqslant 1$ will be treated in section 4 using the language of *matrix recurrences*.

Let us rewrite the recurrences (3.3). We let $\kappa_j := \gamma_{2j}$ and $\widehat{\kappa}_j := \gamma_{2j+1}$, and we move the scalars to the right-hand side of the polynomials:

$$\begin{aligned} &\rho_j(z) - \rho_{j-1}(z)\,\kappa_j = \widehat{\rho}_j(z)\,(1-\kappa_j)\,, \\ &\widehat{\rho}_{j+1}(z) - \widehat{\rho}_j(z)\,\widehat{\kappa}_j = \rho_j(z)\,(1-\widehat{\kappa}_j) - z\rho_j(z)\,\omega_j(1-\widehat{\kappa}_j) \end{aligned} \qquad (j = 1, 2, \dots)\,. \tag{3.5}$$

Next we gather some of the coefficients in, respectively, a lower and an upper bidiagonal matrix:

$$\underline{\mathbf{B}}_m := \begin{pmatrix} -\kappa_1 & & & \\ 1 & -\kappa_2 & & \\ & \ddots & \ddots & \\ & & 1 & -\kappa_m \\ & & & 1 \end{pmatrix}, \quad \widehat{\mathbf{B}}_m := \begin{pmatrix} 1 & -\widehat{\kappa}_1 & & \\ & 1 & \ddots & \\ & & \ddots & -\widehat{\kappa}_{m-1} \\ & & & 1 \end{pmatrix}.$$

Here, $\underline{\mathbf{B}}_m$ is underlined because it is an $(m+1) \times m$ matrix, i.e., it contains an extra row at the bottom. We further need the three $m \times m$ diagonal matrices

$$\begin{aligned} \mathbf{D}_{\omega;m} &:= \mathsf{diag}\,(\omega_0, \omega_1, \dots, \omega_{m-1})\,, \\ \mathbf{D}_{\kappa;m} &:= \mathsf{diag}\,(1-\kappa_1, 1-\kappa_2, \dots, 1-\kappa_m)\,, \\ \mathbf{D}_{\widehat{\kappa};m} &:= \mathsf{diag}\,(1, 1-\widehat{\kappa}_1, \dots, 1-\widehat{\kappa}_{m-1}) \end{aligned}$$

and the row vectors

$$\boldsymbol{\rho}_m^{\mathsf{T}}(z) := \begin{pmatrix} \rho_0(z) & \cdots & \rho_{m-1}(z) \end{pmatrix}, \qquad \widehat{\boldsymbol{\rho}}_m^{\mathsf{T}}(z) := \begin{pmatrix} \widehat{\rho}_1(z) & \cdots & \widehat{\rho}_m(z) \end{pmatrix}.$$

Then the recurrences (3.3) for $0 \leqslant j < m$ (including the relation $\widehat{\rho}_1(z) = (1 - \omega_0 z)\rho_0(z)$) can be summarized as

$$\begin{aligned} \boldsymbol{\rho}_{m+1}^{\mathsf{T}}(z)\,\underline{\mathbf{B}}_m &= \widehat{\boldsymbol{\rho}}_m^{\mathsf{T}}(z)\,\mathbf{D}_{\kappa;m}\,, \\ z\,\boldsymbol{\rho}_m^{\mathsf{T}}(z)\,\mathbf{D}_{\omega;m}\,\mathbf{D}_{\widehat{\kappa};m} &= \boldsymbol{\rho}_m^{\mathsf{T}}(z)\,\mathbf{D}_{\widehat{\kappa};m} - \widehat{\boldsymbol{\rho}}_m^{\mathsf{T}}(z)\,\widehat{\mathbf{B}}_m\,. \end{aligned} \tag{3.6}$$

Eliminating $\widehat{\boldsymbol{\rho}}_m^{\mathsf{T}}(z)$ with the help of the first equation leads to

$$z\,\boldsymbol{\rho}_m^{\mathsf{T}}(z)\,\mathbf{D}_{\omega;m}\,\mathbf{D}_{\widehat{\kappa};m} = \boldsymbol{\rho}_m^{\mathsf{T}}(z)\,\mathbf{D}_{\widehat{\kappa};m} - \boldsymbol{\rho}_{m+1}^{\mathsf{T}}(z)\,\underline{\mathbf{B}}_m\,\mathbf{D}_{\kappa;m}^{-1}\,\widehat{\mathbf{B}}_m \tag{3.7}$$

or

$$z\,\boldsymbol{\rho}_m^{\mathsf{T}}(z)\,\mathbf{D}_{\omega;m}\,\mathbf{D}_{\widehat{\kappa};m} = \boldsymbol{\rho}_{m+1}^{\mathsf{T}}(z)\left(\begin{pmatrix} \mathbf{D}_{\widehat{\kappa};m} \\ \mathbf{o}_m^{\mathsf{T}} \end{pmatrix} - \underline{\mathbf{B}}_m\,\mathbf{D}_{\kappa;m}^{-1}\,\widehat{\mathbf{B}}_m\right). \tag{3.8}$$

This is the polynomial form of a Hessenberg relation. The standard form is obtained by multiplying from the right-hand side with the inverse of $\mathbf{D}_{\omega;m}\,\mathbf{D}_{\widehat{\kappa};m}$, followed by inserting $z := \mathbf{A}$ and applying both sides to $\mathbf{r}_0$. The $(m+1) \times m$ matrix

$$\underline{\mathbf{T}}_m := \left(\begin{pmatrix} \mathbf{I}_m \\ \mathbf{o}_m^{\mathsf{T}} \end{pmatrix} - \underline{\mathbf{B}}_m\,\mathbf{D}_{\kappa;m}^{-1}\,\widehat{\mathbf{B}}_m\mathbf{D}_{\widehat{\kappa};m}^{-1}\right)\mathbf{D}_{\omega;m}^{-1} \tag{3.9}$$

is tridiagonal, and its leading $m \times m$ principal submatrix $\mathbf{T}_m$ can be understood as an oblique projection of $\mathbf{A}$ (an orthogonal one, diagonally scaled such that it is of OrthoRes-type, if $\mathbf{A}$ is Hermitian). Since for n even, $\mathbf{w}_n = \mathbf{w}_{2j} = \rho_j(\mathbf{A})\mathbf{r}_0$ is known to be a BiCG residual, $\mathbf{T}_m$ must be the same matrix as the one obtained by

m steps of BiORes, the ORes-variant of BiCG [17]. The fact that the recurrence is of OrthoRes-type is clearly visible in (3.8). Since all residual polynomials have a constant term equal to one (see the recurrences (3.3)), setting $z = 0$ proves that the columns of $\underline{\mathbf{T}}_m$ sum to zero.

4. The case $s \geqslant 1$. The relations of importance for eigenvalue computations in IDR(s)ORes are sketched in Algorithm 1. We have omitted the rules for computing the scalars ω_j and the approximate solutions $\mathbf{x}_n$ of the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$. Every rule of computation of the vectors $\mathbf{c}_n$ and the scalars ω_j defines one particular instance of a corresponding IDR algorithm from the family of IDR methods.

Algorithm 1. IDR(s)ORes.

input : $\mathbf{A}$, $\mathbf{b}$, $\mathbf{x}_0$, s, $\mathbf{P}$
output: $\mathbf{R}_{n+1}$, $\mathbf{c}_{s+1}, \mathbf{c}_{s+2}, \ldots$, $\omega_1, \omega_2, \ldots$
1 $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$
2 compute $\mathbf{R}_{s+1} = \mathbf{R}_{0:s} = \big(\mathbf{r}_0, \ldots, \mathbf{r}_s\big)$ using, e.g., OrthoRes
3 $\nabla\mathbf{R}_{1:s} = \big(\nabla\mathbf{r}_1, \ldots, \nabla\mathbf{r}_s\big)$
4 $n \leftarrow s+1$, $j \leftarrow 1$
5 **while** not converged **do**
6 for $k = 0, \ldots, s$ compute $\mathbf{c}_n$
7 $\mathbf{c}_n = (\mathbf{P}^{\mathsf{H}}\nabla\mathbf{R}_{n-s:n-1})^{-1}\mathbf{P}^{\mathsf{H}}\mathbf{r}_{n-1}$
8 $\mathbf{v}_{n-1} = \mathbf{r}_{n-1} - \nabla\mathbf{R}_{n-s:n-1}\mathbf{c}_n$
9 if $k = 0$ compute ω_j
10 $\nabla\mathbf{r}_n = -\nabla\mathbf{R}_{n-s:n-1}\mathbf{c}_n - \omega_j\mathbf{A}\mathbf{v}_{n-1}$
11 $\mathbf{r}_n = \mathbf{r}_{n-1} + \nabla\mathbf{r}_n$
12 $n \leftarrow n+1$
13 $\nabla\mathbf{R}_{n-s:n-1} = \big(\nabla\mathbf{r}_{n-s}, \ldots, \nabla\mathbf{r}_{n-1}\big)$
14 $j \leftarrow j+1$
15 **end**

A few remarks are in order. The original IDR(s)ORes used, in line 2, OrthoRes(1) to compute the residuals $\mathbf{r}_1$ to $\mathbf{r}_s$. Since we can use other Krylov subspace methods, as long as they correspond to a Hessenberg decomposition, we advocate the use of full OrthoRes [44] or GMRes [27]. They do not require extra memory space compared to the later IDR steps. The forward difference operator Δ in the original algorithm has been replaced by the backward difference operator ∇, as this appears to be more natural. The forward and backward difference operators Δ and ∇ are defined by

$$\Delta\mathbf{r}_n := \mathbf{r}_{n+1} - \mathbf{r}_n \quad \text{and} \quad \nabla\mathbf{r}_n := \mathbf{r}_n - \mathbf{r}_{n-1}, \tag{4.1}$$

respectively. These finite difference operators are applied columnwise to matrices. Closely related to the forward and backward difference operators is the MATLAB `diff` operator defined columnwise by

$$\mathbf{c} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_s \end{pmatrix} \in \mathbb{C}^s \quad \Rightarrow \quad \texttt{diff}(\mathbf{c}) = \begin{pmatrix} \gamma_2 - \gamma_1 \\ \vdots \\ \gamma_s - \gamma_{s-1} \end{pmatrix} \in \mathbb{C}^{s-1}. \tag{4.2}$$

We have shifted the index of the residuals and residual differences by -1 compared to [34]. Thus, in line 4 of Algorithm 1, we initialize n to be $s+1$ in place of s. It

turns out that this shifted index n determines the nth column of certain Hessenberg and banded matrices to be introduced later on. There exist several alternative, but mathematically equivalent, ways to update the approximate solutions $\mathbf{x}_n$ and the corresponding residuals $\mathbf{r}_n$ [34].

We need to access the vectors $\mathbf{v}_{n-1}$, $\mathbf{c}_n$ and the scalars ω_j in the eigenvalue computations. Thus, we have appended indices $n-1$ and n to $\mathbf{v}_{n-1}$ and $\mathbf{c}_n$, respectively, and introduced the index j of ω_j not present in [34]. Obviously, j is given by $j = \lfloor n/(s+1) \rfloor$ and remains constant for every $s+1$ steps. The vectors $\mathbf{c}_n$ have the elements

$$\mathbf{c}_n = \begin{pmatrix} \gamma_1^{(n)} & \dots & \gamma_s^{(n)} \end{pmatrix}^{\mathsf{T}}; \tag{4.3}$$

additionally, we define $\gamma_0^{(n)} := 0$, $\gamma_{s+1}^{(n)} := 0$ for all $n \in \mathbb{N}$. We stress that the ordering of the elements in $\mathbf{c}_n$ has changed compared to the IDR algorithm in [34, Figure 3.1].

We further remark that for reasons of numerical stability, line 10 of Algorithm 1 cannot be found "as such" in [34]. In the original IDR(s)ORes the update was given by $\nabla\mathbf{x}_n = -\nabla\mathbf{X}_{n-s:n-1}\mathbf{c}_n + \omega_j\mathbf{v}_{n-1}$ followed by $\nabla\mathbf{r}_n = -\mathbf{A}\nabla\mathbf{x}_n$. As we removed in our variant the recurrences for the iterates from the original IDR(s)ORes algorithm, we update $\mathbf{r}_n$ in Algorithm 1 according to

$$\nabla\mathbf{r}_n = -\mathbf{A}\nabla\mathbf{x}_n = \mathbf{A}\nabla\mathbf{X}_{n-s:n-1}\mathbf{c}_n - \omega_j\mathbf{A}\mathbf{v}_{n-1} = -\nabla\mathbf{R}_{n-s:n-1}\mathbf{c}_n - \omega_j\mathbf{A}\mathbf{v}_{n-1}\,. \tag{4.4}$$

The last equality sign in (4.4) is justified in exact arithmetic because in this case we can ensure that for all $n - s \leqslant k < n$ indeed

$$-\nabla\mathbf{r}_k = -\mathbf{r}_k + \mathbf{r}_{k-1} = -\mathbf{b} + \mathbf{A}\mathbf{x}_k + \mathbf{b} - \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}\nabla\mathbf{x}_k\,. \tag{4.5}$$

In finite precision the gap between the negative *computed* residual differences $-\nabla\mathbf{r}_k$ and the *computed* differences $\nabla\mathbf{x}_k$ of the iterates multiplied by $\mathbf{A}$ has to be monitored. It indicates whether we can trust the eigenpair approximations obtained using IDR.

4.1. The original IDR Hessenberg recurrence. We reformulate the recurrences slightly by expressing everything in terms of residual vectors instead of residual differences, which amounts to a discrete partial integration, i.e., application of Abel's summation formula,

$$\begin{aligned} &\mathbf{r}_{n-1} - \nabla\mathbf{R}_{n-s:n-1}\mathbf{c}_n \\ &\quad = \mathbf{r}_{n-1} - \sum_{j=0}^{s-1}(\mathbf{r}_{n-s+j} - \mathbf{r}_{n-s+j-1})\gamma_{j+1}^{(n)} \\ &\quad = \mathbf{r}_{n-1} - \sum_{j=0}^{s}\mathbf{r}_{n-s+j-1}(\gamma_j^{(n)} - \gamma_{j+1}^{(n)}) = \mathbf{R}_{n-s-1:n-1}\,\mathtt{diff}\begin{pmatrix}0\\ \mathbf{c}_n\\ 1\end{pmatrix}. \end{aligned} \tag{4.6}$$

Eliminating the residual differences, we observe that in all inner steps we have the residual recurrence

$$\begin{aligned} \mathbf{r}_n &= (\mathbf{I} - \omega_j\mathbf{A})\mathbf{v}_{n-1} = (\mathbf{I} - \omega_j\mathbf{A})(\mathbf{r}_{n-1} - \nabla\mathbf{R}_{n-s:n-1}\mathbf{c}_n) \\ &= (\mathbf{I} - \omega_j\mathbf{A})\,\mathbf{R}_{n-s-1:n-1}\,\mathtt{diff}\begin{pmatrix}0\\ \mathbf{c}_n\\ 1\end{pmatrix} =: (\mathbf{I} - \omega_j\mathbf{A})\,\mathbf{R}_{n-s-1:n-1}\mathbf{y}_n, \end{aligned} \tag{4.7}$$

as in the left set of equations in the recurrence (3.1) for the case $s = 1$.

Classical Krylov subspace theory would embark upon the construction of a Hessenberg decomposition with the residuals as a "basis." The leading submatrices of the constructed Hessenberg matrix are oblique "projections" of the original matrix; thus we could use the eigenvalues of these submatrices as approximations of eigenvalues, the so-called Ritz values, and the prolongations of the Hessenberg eigenvectors, the so-called Ritz vectors, would serve as corresponding approximate eigenvectors. Obviously, the matrix of residuals most probably will be rank deficient when IDR terminates. Nevertheless, the specially structured Hessenberg decomposition related to IDR will give good approximations to eigenvalues. To demonstrate the general approach, we first derive the generalized Hessenberg decomposition for the IDR residuals.

THEOREM 4.1 (the original Hessenberg/upper triangular IDR(s)ORES pencil). *The generalized Hessenberg decomposition for the* IDR(s)ORES *residuals is given by*

$$\mathbf{A}\mathbf{R}_n\mathbf{Y}_n\mathbf{D}_\omega^{(n)} = \mathbf{R}_{n+1}\underline{\mathbf{Y}}_n^\circ, \tag{4.8}$$

where $\mathbf{R}_{n+1} = \big(\mathbf{r}_0, \ldots, \mathbf{r}_n\big)$ *is the matrix of all residual vectors up to step* n. *For* $s < k \leqslant n$, *the* k*th columns of the upper triangular matrix* $\mathbf{Y}_n \in \mathbb{C}^{n\times n}$ *and of the extended unreduced Hessenberg matrix* $\underline{\mathbf{Y}}_n^\circ \in \mathbb{C}^{(n+1)\times n}$ *of* ORTHORES*-type are defined by*

$$\begin{aligned} \mathbf{Y}_n\mathbf{e}_k &:= \begin{pmatrix} \mathbf{o}_{k-(s+1)} \\ \mathbf{y}_k \\ \mathbf{o}_{n-k} \end{pmatrix}, \\ \underline{\mathbf{Y}}_n^\circ\mathbf{e}_k &:= \begin{pmatrix} \mathbf{o}_{k-(s+1)} \\ \mathbf{y}_k \\ -1 \\ \mathbf{o}_{n-k} \end{pmatrix}, \end{aligned} \qquad \text{where} \quad \mathbf{y}_k := \begin{pmatrix} \gamma_1^{(k)} \\ \gamma_2^{(k)} - \gamma_1^{(k)} \\ \vdots \\ \gamma_s^{(k)} - \gamma_{s-1}^{(k)} \\ 1 - \gamma_s^{(k)} \end{pmatrix} \in \mathbb{C}^{s+1}, \tag{4.9}$$

while the diagonal elements $\mathbf{e}_k^{\mathsf{T}}\mathbf{D}_\omega^{(n)}\mathbf{e}_k$ *of the diagonal matrix* $\mathbf{D}_\omega^{(n)}$ *are defined by*

$$\mathbf{e}_k^{\mathsf{T}}\mathbf{D}_\omega^{(n)}\mathbf{e}_k := \omega_j, \quad j = \lfloor k/(s+1)\rfloor. \tag{4.10}$$

The leading portions of the matrices $\mathbf{Y}_n$, $\underline{\mathbf{Y}}_n^\circ$, *and* $\mathbf{D}_\omega^{(n)}$ *are given by the Hessenberg decomposition of the starting procedure chosen.*

Proof. We sort terms in (4.7) according to the occurrence of the matrix $\mathbf{A}$ and obtain

$$\omega_j\mathbf{A}\mathbf{R}_{n-s-1:n-1}\mathbf{y}_n = \mathbf{R}_{n-s-1:n-1}\mathbf{y}_n - \mathbf{r}_n = \mathbf{R}_{n-s-1:n}\begin{pmatrix}\mathbf{y}_n \\ -1\end{pmatrix}, \tag{4.11}$$

which is the nth column

$$\omega_j\mathbf{A}\mathbf{R}_n\begin{pmatrix}\mathbf{o}_{n-(s+1)} \\ \mathbf{y}_n\end{pmatrix} = \mathbf{R}_{n+1}\begin{pmatrix}\mathbf{o}_{n-(s+1)} \\ \mathbf{y}_n \\ -1\end{pmatrix} \tag{4.12}$$

of the generalized Hessenberg decomposition (4.8). ∎

Remark 4.2. In the case of a Hessenberg decomposition

$$\mathbf{A}\mathbf{R}_s = \mathbf{R}_{s+1}\underline{\mathbf{C}}_s^\circ \tag{4.13}$$

of a starting procedure of ORTHORES-type ($\mathbf{e}^{\mathsf{T}}\underline{\mathbf{C}}_s^\circ = \mathbf{o}_s^{\mathsf{T}}$) we set $\mathbf{Y}_s := \mathbf{I}_s$, $\mathbf{D}_\omega^{(s)} := \mathbf{I}_s$, and $\underline{\mathbf{Y}}_s^\circ := \underline{\mathbf{C}}_s^\circ$.

Remark 4.3. In the original IDR(s)ORes variant a truncated OrthoRes (i.e., OrthoRes(1) $\approx$ ORes) was used for the first s steps. In this setting, we would define $\mathbf{Y}_s := \mathbf{I}_s$, $\mathbf{D}_\omega^{(s)} := \mathsf{diag}\,(\tilde{\omega}_1, \ldots, \tilde{\omega}_s)$, where $\tilde{\omega}_k$ is defined by the residual minimization in step k, and $\underline{\mathbf{Y}}_s^\circ := \underline{\mathbf{E}}_s^\circ$, where $\underline{\mathbf{E}}_s^\circ$ is defined in (1.7).

Definition 4.4. *We call the banded Hessenberg/banded upper triangular pencil* $(\mathbf{Y}_n^\circ, \mathbf{Y}_n\mathbf{D}_\omega^{(n)})$ *the Sonneveld pencil.*[6]

Remark 4.5. The Hessenberg/upper triangular Sonneveld pencil is banded with upper bandwidth s. This reflects that IDR(s)ORes is an $(s+1, s+1)$–step method [15]. Those elements of the upper triangular parts of $\mathbf{Y}_n^\circ$ and $\mathbf{Y}_n$ that are given by IDR(s)ORes are identical and defined in terms of differences of the elements of the vectors $\mathbf{c}_k$. The matrix $\mathbf{D}_\omega^{(n)}$ is uniquely defined by n, s, and ω_j, $1 \leqslant j \leqslant \lfloor n/(s+1) \rfloor$.

As we are going to see next, the generalized Hessenberg decomposition (4.8) can be turned into a Hessenberg decomposition.

Corollary 4.6. *Suppose that the Hessenberg decomposition of the starting procedure, i.e., of the first s steps, can be written as* $\mathbf{A}\mathbf{R}_s = \mathbf{R}_{s+1}\underline{\mathbf{C}}_s^\circ$. *Suppose further that* $\omega_j \neq 0$, $1 \leqslant j \leqslant \lfloor n/(s+1) \rfloor$, *and that* $\gamma_s^{(k)} \neq 1$, $s < k \leqslant n$. *Then the Hessenberg decomposition for the* IDR(s)ORes *residuals is given by*

$$\mathbf{A}\mathbf{R}_n = \mathbf{R}_{n+1}\underline{\mathbf{S}}_n^\circ, \tag{4.14}$$

where the Sonneveld matrix

$$\underline{\mathbf{S}}_n^\circ := \underline{\mathbf{Y}}_n^\circ (\mathbf{D}_\omega^{(n)})^{-1}\mathbf{Y}_n^{-1} \tag{4.15}$$

is an unreduced extended Hessenberg matrix of OrthoRes*-type.*

Proof. With this assumption on the starting Hessenberg decomposition, the diagonal matrix $\mathbf{D}_\omega^{(n)}$ is invertible if and only if $\omega_j \neq 0$, $1 \leqslant j \leqslant \lfloor n/(s+1) \rfloor$, and the upper triangular matrix $\mathbf{Y}_n$ is invertible if and only if $\gamma_s^{(k)} \neq 1$, $s+1 \leqslant k \leqslant n$. ∎

The Hessenberg matrix $\mathbf{S}_n^\circ$ can be used to compute eigenvalues and eigenvectors, which are prolonged by the residual matrix $\mathbf{R}_n$, to obtain Ritz values and Ritz vectors, respectively. The Ritz values, i.e., the eigenvalues of the Sonneveld matrix $\mathbf{S}_n^\circ$, are the roots of the residual polynomials $\mathcal{R}_n(z) := \mathsf{det}\,(\mathbf{I}_n - z(\mathbf{S}_n^\circ)^{-1})$. The Ritz vectors are the prolonged eigenvectors of the Sonneveld matrices; these will be discussed separately in section 5. The Sonneveld matrix, a typically full Hessenberg matrix, is not formed explicitly; instead we apply the QZ algorithm to the banded Hessenberg/upper triangular Sonneveld pencil $(\mathbf{Y}_n^\circ, \mathbf{Y}_n\mathbf{D}_\omega^{(n)})$.

However, this direct way of computing eigenvalue approximations is not always the method of choice in Lanczos-type product methods (LTPM). Indeed, we *know* some of the roots of the residual polynomials, namely, the roots $1/\omega_j$.

4.2. The purified IDR Hessenberg recurrence. The division of the residual polynomials $\mathcal{R}_n$ by the known linear factors, i.e., in step n by the residual polynomial factor

$$\Omega_j(z) = \prod_{i=1}^{j}(1 - \omega_i z), \quad j = \lfloor n/(s+1) \rfloor, \tag{4.16}$$

[6]This terminology honors Peter Sonneveld, who developed with the classical IDR variant [42] in 1979 not only the first Lanczos-type product method (LTPM), but soon after introduced with the conjugate gradient squared method another LTPM based on a different view, which led to a whole family of such methods.

results in polynomial recurrences which correspond to a generalized Hessenberg decomposition with an unreduced Hessenberg/upper triangular pencil that has only the unknown residual polynomial roots as eigenvalues and, additionally, some infinite eigenvalues.

To proceed, we rewrite the residual recurrence. We already have proven, see (4.7), that the recurrences for $s \geqslant 1$ and $n > s$ are mathematically equivalent to

$$\begin{aligned} \mathbf{v}_{n-1} &:= \mathbf{r}_{n-1} - \nabla \mathbf{R}_{n-s:n-1}\mathbf{c}_n = \mathbf{R}_{n-s-1:n-1}\mathbf{y}_n \\ &= (1-\gamma_s^{(n)})\,\mathbf{r}_{n-1} + \textstyle\sum_{\ell=1}^{s-1}(\gamma_{s-\ell+1}^{(n)} - \gamma_{s-\ell}^{(n)})\,\mathbf{r}_{n-\ell-1} + \gamma_1^{(n)}\,\mathbf{r}_{n-s-1}\,, \\ \mathbf{r}_n &:= (\mathbf{I} - \omega_j \mathbf{A})\,\mathbf{v}_{n-1}\,, \end{aligned} \tag{4.17}$$

where j is given by (4.16) and remains constant for every inner sweep of $s+1$ steps. We note that $k = n - j(s+1)$, which cycles between 0 and s, is the running index of the inner sweep.

We consider a general starting procedure (e.g., GMRes, GCR $\approx$ OrthoRes) that produces the s residuals $\mathbf{r}_1, \ldots, \mathbf{r}_s$. This information uniquely defines the first s residual polynomials we are interested in. Furthermore, we have that for all $j \geqslant 1$,

$$\begin{matrix} \mathbf{r}_{j(s+1)+0} \\ \mathbf{r}_{j(s+1)+1} \\ \vdots \\ \mathbf{r}_{j(s+1)+s} \end{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \begin{matrix} \text{is based on the construction of } \omega_j, \\ \\ \text{is constructed using } \omega_j. \end{matrix} \tag{4.18}$$

We define $\mathcal{S} := \mathcal{P}^\perp$ as the orthogonal complement of the range $\mathcal{P}$ of the columns of $\mathbf{P}$. We observe that every residual $\mathbf{r}_n$ with index n in the range $j(s+1)+k$, $0 \leqslant k \leqslant s$, is obtained from previous information by a multiplication of $\mathbf{v}_{n-1}$ with the linear mapping

$$(\mathbf{I} - \omega_j \mathbf{A}) : \mathcal{G}_{j-1} \cap \mathcal{S} \to \mathcal{G}_j, \tag{4.19}$$

that defines the jth Sonneveld space $\mathcal{G}_j$; see (1.4). Recall that $\mathcal{G}_0 = \mathcal{K}_N(\mathbf{A}, \mathbf{r}_0) \subset \mathbb{C}^N$ is the *full* Krylov subspace generated by $\mathbf{A}$ with starting vector $\mathbf{r}_0$. We use the fact pointed out in [34, p. 1046] that all residuals constructed are of the form

$$\mathbf{r}_n = \Omega_j(\mathbf{A})\mathbf{w}_n, \quad \mathbf{v}_{n-1} = \Omega_{j-1}(\mathbf{A})\mathbf{w}_n, \quad j = \lfloor n/(s+1) \rfloor\,, \tag{4.20}$$

where the vectors $\mathbf{w}_n \in \mathcal{K}_{n-j}(\mathbf{A}, \mathbf{r}_0)$ can be interpreted as residuals too, since $\Omega_j(0) = 1$. We call these vectors $\mathbf{w}_n$ the *purified residuals*. The residual polynomial $\mathcal{W}_n$ corresponding to $\mathbf{w}_n$, which is of degree $n-j$, satisfies $\mathcal{W}_n(z) = \mathcal{R}_n(z)/\Omega_j(z)$. Since $\Omega_0(z) \equiv 1$, the first $s+1$ residuals $\mathbf{r}_j$ and $\mathbf{w}_j$, $0 \leqslant j \leqslant s$, coincide; especially, $\mathbf{w}_0 = \mathbf{r}_0$. Utilizing these connections we now obtain recurrences for these residual polynomials and for the vectors $\mathbf{w}_n$. For $n = j(s+1)+k$, $0 \leqslant k \leqslant s$, we know that for two families of polynomials denoted by ρ_{sj+k}, $0 \leqslant k < s$ and $\widehat{\rho}_{(j+1)s}$, the latter corresponding to the missing case $k = s$, the following holds true:

$$\mathbf{w}_n = \mathcal{W}_n(\mathbf{A})\mathbf{r}_0 = \mathcal{W}_{j(s+1)+k}(\mathbf{A})\mathbf{r}_0 =: \begin{cases} \rho_{js+k}(\mathbf{A})\mathbf{w}_0, & 0 \leqslant k < s, \\ \widehat{\rho}_{(j+1)s}(\mathbf{A})\mathbf{w}_0, & k = s, \end{cases} \tag{4.21}$$

i.e.,

$$\mathbf{r}_n = \Omega_j(\mathbf{A})\mathbf{w}_n = \begin{cases} \Omega_j(\mathbf{A})\rho_{js+k}(\mathbf{A})\mathbf{w}_0, & 0 \leqslant k < s, \\ \Omega_j(\mathbf{A})\widehat{\rho}_{(j+1)s}(\mathbf{A})\mathbf{w}_0, & k = s, \end{cases}$$

(4.22)

$$\mathbf{v}_{n-1} = \Omega_{j-1}(\mathbf{A})\mathbf{w}_n = \begin{cases} \Omega_{j-1}(\mathbf{A})\rho_{js+k}(\mathbf{A})\mathbf{w}_0, & 0 \leqslant k < s, \\ \Omega_{j-1}(\mathbf{A})\widehat{\rho}_{(j+1)s}(\mathbf{A})\mathbf{w}_0, & k = s. \end{cases}$$

In the case $s = 1$ of section 3, the polynomials ρ_{sj+0} correspond to the BiCG polynomials ρ_j and the polynomials $\widehat{\rho}_{(j+1)s}$ correspond to the polynomials $\widehat{\rho}_{j+1}$. We remark that the two consecutive polynomials $\mathcal{W}_{(j+1)(s+1)-1} = \mathcal{W}_{j(s+1)+s} = \widehat{\rho}_{(j+1)s}$ and $\mathcal{W}_{(j+1)(s+1)} = \rho_{(j+1)s+0}$ have the same degree $(j+1)s$.

We insert the polynomial expressions (4.22) into the second line of the relations (4.17), so that

$$\mathbf{v}_{n-1} = (1-\gamma_s^{(n)})\,\mathbf{r}_{n-1} + \sum_{\ell=1}^{s-1}(\gamma_{s-\ell+1}^{(n)} - \gamma_{s-\ell}^{(n)})\,\mathbf{r}_{n-\ell-1} + \gamma_1^{(n)}\,\mathbf{r}_{n-s-1}\,, \tag{4.23}$$

and divide by $\Omega_{j-1}(\mathbf{A})$, which is a common factor of all polynomials involved, to obtain for n divisible by $s+1$, i.e., $n = j(s+1)$, and for $k = 1, \ldots, s$,

$$\mathbf{w}_{n+0} = \sum_{\ell=0}^{s} \mathbf{w}_{n-\ell-1} \cdot y_{s+1-\ell}^{(n+0)}\,, \tag{4.24a}$$

$$\mathbf{w}_{n+k} = (\mathbf{I} - \omega_j\mathbf{A}) \sum_{\ell=0}^{k-1} \mathbf{w}_{n+k-\ell-1} \cdot y_{s+1-\ell}^{(n+k)} + \sum_{\ell=k}^{s} \mathbf{w}_{n+k-\ell-1} \cdot y_{s+1-\ell}^{(n+k)}\,. \tag{4.24b}$$

Here, $y_1^{(n)}$ to $y_{s+1}^{(n)}$ denote the elements of $\mathbf{y}_n \in \mathbb{C}^{s+1}$.

The index $n = j(s+1)$ has been chosen such that all "purified" residuals $\mathbf{w}_{n-\ell-1}$, $0 \leqslant \ell \leqslant s$, have no additional linear factor $(\mathbf{I}-\omega_j\mathbf{A})$. The *purified residual polynomial* corresponding to $\mathbf{w}_n = \mathbf{w}_{j(s+1)}$ is ρ_{js}, which is in theory an ML(s)BiCG polynomial.

Let us consider again the degrees of the polynomials. In the first step (4.24a) of a new cycle, the polynomial of interest is obtained as a linear combination of the previously computed polynomials, which results in a polynomial of the same degree as the last one (provided that $\gamma_s^{(n)} \neq 1$). Every other step increases the degree of the polynomial by one (provided that all $\gamma_s^{(n+k)} \neq 1$, $1 \leqslant k \leqslant s$). This corresponds to a slightly generalized Hessenberg decomposition, where the Hessenberg pencil is "reduced" at every $(s+1)$th column. The precise meaning of "reduced" becomes obvious in the following theorem.

THEOREM 4.7 (the purified Hessenberg/upper triangular IDR(s)ORes pencil). *The generalized Hessenberg decomposition for the purified* IDR(s)ORes *residuals is given by*

$$\mathbf{A}\mathbf{W}_n\mathbf{U}_n\mathbf{D}_\omega^{(n)} = \mathbf{W}_{n+1}\underline{\mathbf{Y}}_n^\circ, \tag{4.25}$$

where

$$\mathbf{W}_{n+1} := \big(\mathbf{w}_0, \ldots, \mathbf{w}_n\big) \tag{4.26}$$

is the matrix of all purified residual vectors up to step n, $\underline{\mathbf{Y}}_n^\circ \in \mathbb{C}^{(n+1)\times n}$ *and* $\mathbf{D}_\omega^{(n)} \in \mathbb{C}^{n\times n}$ *are defined in Theorem* 4.1, *and the upper triangular* $\mathbf{U}_n \in \mathbb{C}^{n\times n}$ *is obtained*

from $\mathbf{Y}_n$ *by setting to zero all elements in the lower triangular parts of the square submatrices* $(y_{(\ell-1)(s+1)+1+k,\ell(s+1)-1+j}; k,j = 0,\dots,s), \ell = 1,2,\dots,$ *such that the resulting matrix is block diagonal with alternating* $s \times s$ *upper triangular blocks of elements from* $\mathbf{Y}_n$ *and* 1×1 *zero blocks; see Figure* A.1.

Proof. The general equation (4.24b) can be sorted according to terms involving $\omega_j \mathbf{A}$ and others. For all $0 \leqslant k \leqslant s$ we have

$$\omega_j \mathbf{A} \sum_{\ell=0}^{k-1} \mathbf{w}_{n+k-\ell-1} \cdot y_{s+1-\ell}^{(n+k)} = -\mathbf{w}_{n+k} + \sum_{\ell=0}^{s} \mathbf{w}_{n+k-\ell-1} \cdot y_{s+1-\ell}^{(n+k)}. \tag{4.27}$$

For $n = j(s+1)$, the $(n+k)$th column of the resulting generalized Hessenberg decomposition is

$$\omega_j \mathbf{A} \mathbf{W}_{n+k} \begin{pmatrix} \mathbf{o}_n \\ \mathbf{y}_{n+k}^{s+2-k:s+1} \end{pmatrix} = \mathbf{W}_{n+k+1} \begin{pmatrix} \mathbf{o}_{n+k-s-1} \\ \mathbf{y}_{n+k} \\ -1 \end{pmatrix}, \quad 0 \leqslant k \leqslant s. \tag{4.28}$$

Together with the Hessenberg decomposition defining the starting block of residuals scaled to be of OrthoRes-type,

$$\mathbf{A}\mathbf{R}_s = \mathbf{R}_{s+1} \underline{\mathbf{C}}_s^\circ = \mathbf{R}_s \mathbf{C}_s^\circ - \mathbf{r}_s c_{s+1,s}^\circ \mathbf{e}_s^\mathsf{T}, \tag{4.29}$$

and the observation that $\mathbf{W}_{s+1} = \mathbf{R}_{s+1}$, we obtain (4.25). ∎

Remark 4.8. The generalized Hessenberg decomposition (4.25) corresponds to a *singular* projection, since $\mathbf{U}_n$ is singular. In fact, the rows and columns of index $(s+1)j$ contain only zeros.

Remark 4.9. As the columns of $\underline{\mathbf{Y}}_n^\circ$ sum to zero, (4.25) defines again a method of OrthoRes-type. In the part of $\underline{\mathbf{Y}}_n^\circ$ that does not belong to the starting procedure, the lower codiagonal consists of minus ones.

Remark 4.10. The transition from the generalized Hessenberg decomposition (4.8) for $\mathbf{R}_n$ to the generalized Hessenberg decomposition (4.25) for $\mathbf{W}_n$ is computationally rather simple: To obtain the only new matrix $\mathbf{U}_n$ we just blanked certain lower triangles in $\mathbf{Y}_n$ and thus cut this upper triangular band matrix into $s \times s$ upper triangular matrices alternating with zeros; compare the first two matrix pairs in Figure A.1. In fact, the linear combination of $s+1$ terms implicit on the left-hand side of (4.12) has been replaced by a linear combination of k terms, where $0 \leqslant k \leqslant s$, implicit on the left-hand side of (4.27) and (4.28). At this stage no additional round-off errors occur.

We have shown in section 2 that the residual polynomials are up to the known nonzero factors $\mathsf{det}\,(-\mathbf{Y}_k^\circ) = y_{1:k}^\circ$, the leading subdeterminants of the regular[7] Hessenberg pencil

$${}^z\mathbf{H}_n := z\mathbf{U}_n \mathbf{D}_\omega^{(n)} - \mathbf{Y}_n^\circ = \left(\mathbf{I}_n - z\mathbf{U}_n \mathbf{D}_\omega^{(n)} (\mathbf{Y}_n^\circ)^{-1}\right)(-\mathbf{Y}_n^\circ). \tag{4.30}$$

The elements

$$h_{j+1,j}(z) = -y_{j+1,j}^\circ = \begin{cases} \text{defined by starting procedure} & \text{if} \quad j \leqslant s, \\ 1 & \text{if} \quad j > s, \end{cases} \tag{4.31}$$

of the lower codiagonal of ${}^z\mathbf{H}$ are all nonzero and constant and thus the scaling by products of elements from the lower codiagonal can be neglected as we are only interested in the *roots* of the residual polynomials.

[7] The pencil ${}^z\mathbf{H}_n$ is regular, as the determinant of $-\mathbf{Y}_n^\circ$ is given by the product of the nonzero subdiagonal elements of $\mathbf{Y}_n^\circ$; see (1.8) and (4.31).

The pencil ${}^z\mathbf{H}_n$ has exactly $j = \lfloor n/(s+1) \rfloor$ infinite eigenvalues, as every $(s+1)$th diagonal element of the upper triangular matrix $\mathbf{U}_n$ is zero. The resulting determinant, the characteristic polynomial $\chi_n(z) := \mathsf{det}\,({}^z\mathbf{H}_n)$, is of degree $n-j$. The degree of the polynomial remains constant in every step where in $\mathbf{U}_n$ a new zero is introduced in the diagonal; otherwise the degree increases by one, which is in accordance with our observation for the purified residual polynomials.

Remark 4.11. Every $s+1$ steps we have two consecutive purified residual polynomials with the same degree $n-j = js$, $j \geqslant 1$, namely, $\mathcal{W}_{j(s+1)-1} = \mathcal{W}_{(j-1)(s+1)+s} = \widehat{\rho}_{js}$ and $\mathcal{W}_{j(s+1)} = \rho_{js}$. This is reflected by the positions of the zeros in the diagonal of $\mathbf{U}_n$.

This shows that we can compute the roots of the purified residual polynomials by computing the eigenvalues of the pencil ${}^z\mathbf{H}_n$, i.e., by application of the QZ algorithm to the banded unreduced upper Hessenberg/upper triangular pencil $(\mathbf{Y}_n^\circ, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$, e.g., by invoking $\mathtt{eig}(\mathbf{Y}_n^\circ, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$ in MATLAB.

In summary, the computation of the eigenvalues of the purified pencil including the infinite eigenvalues is viable.

4.3. The deflated IDR Hessenberg recurrence. It is possible to deflate the purified pencil in order to obtain a smaller pencil having only the finite eigenvalues. This also enhances our understanding of IDR(s)ORes, as this results in the reconstruction of the underlying BiORes$(s,1)$ process. The deflation of the purified pencil is based on the Schur complement [29, pp. 216–217], i.e., block Gaussian elimination. To be more precise, with the submatrix

$$
{}^z\mathbf{H}_{(j-1)(s+1)+1:(j+1)(s+1)-1} =: \begin{pmatrix} {}^z\mathbf{H}^\star & \mathbf{h}_c & \mathbf{L}^\star \\ \mathbf{e}_s^\mathsf{T} & \gamma^\star - 1 & \mathbf{h}_r^\mathsf{T} \\ \mathbf{O} & \mathbf{e}_1 & {}^z\mathbf{H}_\star \end{pmatrix} \in \mathbb{C}^{(2s+1)\times(2s+1)} \tag{4.32}
$$

of the purified pencil ${}^z\mathbf{H}_n = z\mathbf{U}_n\mathbf{D}_\omega^{(n)} - \mathbf{Y}_n^\circ$, for some matching sweep index j, $1 \leqslant j < \lfloor n/(s+1) \rfloor$, we let ${}^z\mathbf{H}^\star := {}^z\mathbf{H}_{(j-1)(s+1)+1:j(s+1)-1}$ and ${}^z\mathbf{H}_\star := {}^z\mathbf{H}_{j(s+1)+1:(j+1)(s+1)-1}$ denote the leading and trailing Hessenberg matrix of size $s \times s$ of this submatrix, respectively. These are the only matrices that depend on the variable z in the purified pencil; compare with the zero structure of $\mathbf{U}_n\mathbf{D}_\omega^{(n)}$ in the purified pencil depicted in Figure A.1. The element in the diagonal in position $j(s+1)$ is denoted by $\gamma^\star - 1$, where $\gamma^\star := \gamma_s^{(j(s+1))}$. The possibly nonzero elements in the $j(s+1)$th column and row of the upper triangular part of ${}^z\mathbf{H}_n$, i.e., of $-\mathbf{Y}_n^\circ$ are collected in the vectors $\mathbf{h}_c \in \mathbb{C}^s$ and $\mathbf{h}_r \in \mathbb{C}^s$, respectively. The matrix $\mathbf{L}^\star \in \mathbb{C}^{s\times s}$ is a strictly lower triangular matrix whose elements are independent of z.

We now use the following variant of the so-called *Schur determinant formula* [48, eq. (0.3.2), p. 5],

$$
\begin{aligned}
&\begin{pmatrix} {}^z\mathbf{H}^\star & \mathbf{h}_c & \mathbf{L}^\star \\ \mathbf{e}_s^\mathsf{T} & \gamma^\star - 1 & \mathbf{h}_r^\mathsf{T} \\ \mathbf{O} & \mathbf{e}_1 & {}^z\mathbf{H}_\star \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{o} & \mathbf{O} \\ -\mathbf{e}_s^\mathsf{T}/(\gamma^\star - 1) & 1/(\gamma^\star - 1) & -\mathbf{h}_r^\mathsf{T}/(\gamma^\star - 1) \\ \mathbf{O} & \mathbf{o} & \mathbf{I} \end{pmatrix} \\
&= \begin{pmatrix} {}^z\mathbf{H}^\star - \mathbf{h}_c\mathbf{e}_s^\mathsf{T}/(\gamma^\star - 1) & \mathbf{h}_c^\mathsf{T}/(\gamma^\star - 1) & \mathbf{L}^\star - \mathbf{h}_c\mathbf{h}_r^\mathsf{T}/(\gamma^\star - 1) \\ \mathbf{o}^\mathsf{T} & 1 & \mathbf{o}^\mathsf{T} \\ -\mathbf{e}_1\mathbf{e}_s^\mathsf{T}/(\gamma^\star - 1) & \mathbf{e}_1^\mathsf{T}/(\gamma^\star - 1) & {}^z\mathbf{H}_\star - \mathbf{e}_1\mathbf{h}_r^\mathsf{T}/(\gamma^\star - 1) \end{pmatrix} \\
&= \begin{pmatrix} {}^z\mathbf{H}^\star & \mathbf{h}_c^\mathsf{T}/(\gamma^\star - 1) & \mathbf{L}^\star \\ \mathbf{o}^\mathsf{T} & 1 & \mathbf{o}^\mathsf{T} \\ \mathbf{O} & \mathbf{e}_1^\mathsf{T}/(\gamma^\star - 1) & {}^z\mathbf{H}_\star \end{pmatrix} - \begin{pmatrix} \mathbf{h}_c\mathbf{e}_s^\mathsf{T}/(\gamma^\star - 1) & \mathbf{o} & \mathbf{h}_c\mathbf{h}_r^\mathsf{T}/(\gamma^\star - 1) \\ \mathbf{o}^\mathsf{T} & 0 & \mathbf{o}^\mathsf{T} \\ \mathbf{e}_1\mathbf{e}_s^\mathsf{T}/(\gamma^\star - 1) & \mathbf{o} & \mathbf{e}_1\mathbf{h}_r^\mathsf{T}/(\gamma^\star - 1) \end{pmatrix}.
\end{aligned} \tag{4.33}
$$

As the determinant of the block-Gauss eliminator is given by the nonzero value

$$\begin{vmatrix} \mathbf{I} & \mathbf{o} & \mathbf{O} \\ -\mathbf{e}_s^{\mathsf{T}}/(\gamma^\star-1) & 1/(\gamma^\star-1) & -\mathbf{h}_r^{\mathsf{T}}/(\gamma^\star-1) \\ \mathbf{O} & \mathbf{o} & \mathbf{I} \end{vmatrix} = \frac{1}{\gamma^\star-1}\,, \tag{4.34}$$

we have proven that

$$\begin{aligned} \frac{\chi_{(j-1)(s+1)+1:(j+1)(s+1)-1}(z)}{\gamma^\star-1} &= \begin{vmatrix} {}^z\mathbf{H}^\star - \mathbf{h}_c\mathbf{e}_s^{\mathsf{T}}/(\gamma^\star-1) & \mathbf{h}_c^{\mathsf{T}}/(\gamma^\star-1) & \mathbf{L}^\star - \mathbf{h}_c\mathbf{h}_r^{\mathsf{T}}/(\gamma^\star-1) \\ \mathbf{o}^{\mathsf{T}} & 1 & \mathbf{o}^{\mathsf{T}} \\ -\mathbf{e}_1\mathbf{e}_s^{\mathsf{T}}/(\gamma^\star-1) & \mathbf{e}_1^{\mathsf{T}}/(\gamma^\star-1) & {}^z\mathbf{H}_\star - \mathbf{e}_1\mathbf{h}_r^{\mathsf{T}}/(\gamma^\star-1) \end{vmatrix} \\ &= \begin{vmatrix} {}^z\mathbf{H}^\star - \mathbf{h}_c\mathbf{e}_s^{\mathsf{T}}/(\gamma^\star-1) & \mathbf{L}^\star - \mathbf{h}_c\mathbf{h}_r^{\mathsf{T}}/(\gamma^\star-1) \\ -\mathbf{e}_1\mathbf{e}_s^{\mathsf{T}}/(\gamma^\star-1) & {}^z\mathbf{H}_\star - \mathbf{e}_1\mathbf{h}_r^{\mathsf{T}}/(\gamma^\star-1) \end{vmatrix}. \end{aligned} \tag{4.35}$$

The multiplication of the pencil ${}^z\mathbf{H}_n$ by an enlarged block-Gauss eliminator $\mathbf{G}_n^{(j)}$ does not affect the roots of the residual polynomial, i.e., the eigenvalues of the pencil, as this is a scaling by the constant $(\gamma^\star-1)$. The resulting rank-one update (4.33) of the pencil only affects the elements of the submatrix given in (4.32) without changing the Hessenberg structure of the leading and trailing block. Looking closely we realize that the resulting Hessenberg/upper triangular pencil is of course still regular and has lost one infinite eigenvalue by construction but, moreover, the Hessenberg part is of OrthoRes-type. This follows from

$$\mathbf{e}^{\mathsf{T}}\begin{pmatrix} \mathbf{h}_c \\ \gamma^\star-1 \\ 1 \end{pmatrix} = 0 \quad \Rightarrow \quad -\mathbf{e}^{\mathsf{T}}\begin{pmatrix} \mathbf{h}_c/(\gamma^\star-1) \\ 1/(\gamma^\star-1) \end{pmatrix} = 1, \tag{4.36}$$

and the observation that only the last column of the pencil ${}^z\mathbf{H}^\star$ is modified, where the element 1, which is needed for column sum zero, is missing. The column sums of the second half of the Hessenberg part are zero, since the omitted row elements have been dispersed across the rows by the rank-one update.

The pattern of the original banded Hessenberg/upper triangular pencil is only modified in the upper triangle of the Hessenberg matrix. The constant lower triangular matrices $\mathbf{L}^\star \in \mathbb{C}^{s\times s}$ become full, as the rank-one update $\mathbf{L}^\star - \mathbf{h}_c\mathbf{h}_r^{\mathsf{T}}/(\gamma^\star-1)$ usually gives a full $s\times s$ matrix. Apart from altering the lower triangular matrices $\mathbf{L}^\star$ to full matrices by a rank-one update, the nonzero structure is only changed in the last column before and the first row after a diagonal zero element in $\mathbf{U}_n$. We remark that the change in the last column of the Hessenberg blocks is just a variation of the correction in [28, Theorem 3.1, Corollary 3.3, pp. 286–287].

Computing the block-Gauss eliminators $\mathbf{G}_n^{(j)}$, $j=1,\ldots$, where a little care has to be taken for the last eliminator which, depending on the relation of n to $s+1$, may work on a smaller trailing Hessenberg matrix, and applying them to the pencil before computing the next eliminator is equivalent to computing *all* block-Gauss eliminators at once and thus their product $\mathbf{G}_n := \mathbf{G}_n^{(1)}\mathbf{G}_n^{(2)}\cdots$, as the eliminators solely depend on quantities which are not altered in previous steps. The basic eliminators $\mathbf{G}_n^{(j)}$

commute, thus the order of multiplication is not important. Due to the zero structure of the pencil, we compute in our implementation the Schur complement of the matrix indexed by all diagonal elements equal to zero in $\mathbf{U}_n$ to enable the use of the BLAS Level 3. The elimination, i.e., multiplication of the pencil $(\mathbf{Y}_n^\circ, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$ by $\mathbf{G}_n$ from the right results in the modified pencil $(\mathbf{Y}_n^\circ\mathbf{G}_n, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$, as the elimination does not change the upper triangular matrix $\mathbf{U}_n\mathbf{D}_\omega^{(n)}$ due to the zero structure; see the second and third pencil in Figure A.1. After elimination, every $(s+1)$th row of the pencil $z\mathbf{U}_n\mathbf{D}_\omega^{(n)} - \mathbf{Y}_n^\circ\mathbf{G}_n$ is independent of the variable z and has its only nonzero element in the diagonal position; see the third pencil in Figure A.1.

In the next step we remove every $(s+1)$th row and column from the pencil, which results in the smaller, fourth pencil depicted in Figure A.1. The infinite eigenvalues have been deflated, and the transformed but not yet deflated pencil clearly reveals that the corresponding eigenvectors are the standard unit vectors $\mathbf{e}_{j(s+1)}$, $1 \leqslant j \leqslant \lfloor n/(s+1)\rfloor$. Turning our attention to the underlying generalized Hessenberg decompositions, we observe that the deflation removes every $(s+1)$th vector from $\mathbf{W}_n$. Because we enumerate the purified residuals starting with index 0, the deflated vectors are $\mathbf{w}_{(j+1)(s+1)-1} = \mathbf{w}_{j(s+1)+s} = \widehat{\rho}_{(j+1)s}(\mathbf{A})\mathbf{w}_0$, $0 \leqslant j < \lfloor n/(s+1)\rfloor$. Let us denote by D the deflation operator that removes every $(s+1)$th row and column up to the nth of an $n \times n$ or $(n+1) \times n$ matrix, and by $\widetilde{D}$ the deflation operator that removes every $(s+1)$th column up to the nth of an $N \times n$ matrix. Then we have proven the following theorem.

THEOREM 4.12 (the deflated Hessenberg/upper triangular IDR(s)ORES pencil). *The generalized Hessenberg decomposition for the deflated* IDR(s)ORES *residuals is given by*

$$\mathbf{A}\widetilde{\mathbf{Q}}_m D(\mathbf{U}_n\mathbf{D}_\omega^{(n)}) = \widetilde{\mathbf{Q}}_{m+1} D(\underline{\mathbf{Y}}_n^\circ\mathbf{G}_n), \quad m := n - \lfloor n/(s+1)\rfloor\,, \tag{4.37}$$

where $\widetilde{\mathbf{Q}}_m := \widetilde{D}(\mathbf{W}_n)$, $\widetilde{\mathbf{Q}}_{m+1} := (\widetilde{\mathbf{Q}}_m, \mathbf{w}_n)$, *and* $\mathbf{G}_n$ *is the block-Gauss eliminator defined above.*

As the deflated triangular matrix $D(\mathbf{U}_n\mathbf{D}_\omega^{(n)})$ is now an invertible block-diagonal matrix with $s \times s$ upper triangular blocks, we can construct an algebraic eigenvalue problem by inversion. To retain the zero column sums we multiply the inverse from the right to the Hessenberg matrix. We obtain a banded Hessenberg matrix, which has triangles poking out of the band:

COROLLARY 4.13. *Suppose that all* $\omega_j \neq 0$ *and that all* $\gamma_s^j \neq 1$. *Then the Hessenberg decomposition for the deflated* IDR(s)ORES *residuals is given by*

$$\mathbf{A}\widetilde{\mathbf{Q}}_m = \widetilde{\mathbf{Q}}_{m+1}\underline{\widetilde{\mathbf{L}}}_m^\circ, \quad m = n - \lfloor n/(s+1)\rfloor\,, \tag{4.38}$$

where

$$\underline{\widetilde{\mathbf{L}}}_m^\circ := D(\underline{\mathbf{Y}}_n^\circ\mathbf{G}_n)\,(D(\mathbf{U}_n\mathbf{D}_\omega^{(n)}))^{-1} \in \mathbb{C}^{(m+1)\times m}. \tag{4.39}$$

The last column of the matrix $\underline{\widetilde{\mathbf{L}}}_m^\circ$ *is altered in the next step of* IDR(s)ORES *if and only if* $n = j(s+1) - 1$ *for some* $j \geqslant 1$. *This matrix is an unreduced extended Hessenberg matrix of* ORTHORES*-type and, at the same time, block tridiagonal with blocks of dimension* s.

In the transition from step $n = j(s+1) - 1$ to step $n = j(s+1)$ the index m and thus the sizes of $\underline{\widetilde{\mathbf{L}}}_m^\circ$ and $\widetilde{\mathbf{Q}}_m$ remain unchanged, but the last columns of these two matrices are changed in the deflation process. So, for these particular values

The BiORes(3, 1) matrix $\mathbf{L}_9^\circ$ obtained with 3 full sweeps:

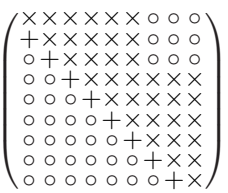


FIG. 4.1. *The* BiORes(3, 1) *matrix* $\mathbf{L}_9^\circ$ *corresponding to the pictorial example given in Figure* A.1.

of m there exist two different pairs of matrices $\underline{\widetilde{\mathbf{L}}}_m^\circ$, $\widetilde{\mathbf{Q}}_m$. In the next s steps, new columns are added to $\underline{\widetilde{\mathbf{L}}}_m^\circ$ and $\widetilde{\mathbf{Q}}_m$, but all other elements of the matrices from the previous steps remain unchanged. We propose here to introduce a subset of these two matrix sequences obtained by deleting those matrices that are generated in steps where $n = j(s+1)-1$ for some j. For distinction, we rename the remaining matrices $\underline{\mathbf{L}}_m^\circ$ and $\mathbf{Q}_m$. They correspond to the leading unchanged parts of the decomposition (4.38). Interestingly, we can interpret these leading unchanged parts as those of the BiORes$(s, 1)$ process underlying IDR(s)ORes.

DEFINITION 4.14. *We define for* $m = js+k \geqslant 1$, $j = \lfloor m/s \rfloor$, $0 \leqslant k < s$, *the* m*th extended* BiORes$(s, 1)$ *Lanczos matrix* $\underline{\mathbf{L}}_m^\circ$ *and the* m*th matrix of the* BiORes$(s, 1)$ *residuals* $\mathbf{Q}_m$ *by*

$$\underline{\mathbf{L}}_m^\circ := D(\underline{\mathbf{Y}}_n^\circ \mathbf{G}_n)\,(D(\mathbf{U}_n \mathbf{D}_\omega^{(n)}))^{-1} \in \mathbb{C}^{(m+1)\times m} \tag{4.40a}$$

and

$$\mathbf{Q}_m := \widetilde{D}(\mathbf{W}_n), \tag{4.40b}$$

respectively, where the index n *is defined by*

$$n := j(s+1)+k, \quad 0 \leqslant k < s. \tag{4.41}$$

We remark that by construction the index n *omits all values of the form* $j(s+1)+s$.

For the reader's convenience, the block-tridiagonal unreduced upper Hessenberg matrix $\mathbf{L}_9^\circ$ corresponding to Figure A.1 with $s = 3$ for 3 sweeps is depicted in Figure 4.1.

For $k = 0, \ldots, s-1$ and $j = 0, 1, \ldots$, the reduced residuals $\mathbf{q}_{js+k}$ are defined by

$$\Omega_j(\mathbf{A})\mathbf{q}_{js+k} := \mathbf{r}_{j(s+1)+k} = (\mathbf{I} - \omega_j \mathbf{A})\mathbf{v}_{j(s+1)+k-1}.$$

Every $\mathbf{v}_{j(s+1)+k-1}$ is orthogonal to $\mathcal{P}$, the range of the columns of $\mathbf{P}$. Thus, $\mathbf{q}_{js+k} \perp \Omega_{j-1}(\mathbf{A}^{\mathsf{H}})\mathcal{P}$. Using induction [32] one can prove that $\mathbf{q}_{js+k} \perp \mathcal{K}_j(\mathbf{A}^{\mathsf{H}}, \mathbf{P})$; thus, the Hessenberg/block-tridiagonal matrices $\mathbf{L}_m^\circ$ can be identified as those of a two-sided Lanczos process with s left and one right starting vectors.

THEOREM 4.15 (the Hessenberg/block-tridiagonal BiORes$(s, 1)$ matrix). *The Hessenberg decomposition of the* BiORes$(s, 1)$ *residuals is given by*

$$\mathbf{A}\mathbf{Q}_m = \mathbf{Q}_{m+1}\underline{\mathbf{L}}_m^\circ, \tag{4.42}$$

where, for $k = 0, \ldots, s-1$ *and* $j = 0, 1, \ldots$, *the* BiORes$(s, 1)$ *residuals satisfy*

$$\mathbf{q}_{js+k} \in \mathbf{q}_0 + \mathbf{A}\mathcal{K}_{js+k}(\mathbf{A}, \mathbf{q}_0), \quad \mathbf{q}_{js+k} \perp \mathcal{K}_j(\mathbf{A}^{\mathsf{H}}, \mathbf{P}). \tag{4.43}$$

This follows easily from [32, Theorem 4.2] and [31, Theorem 4.1].

5. Ritz vectors and accuracy. Ritz vectors are obtained from a given *generalized* Hessenberg decomposition (1.13), $\mathbf{A}\mathbf{Q}_n\mathbf{U}_n = \mathbf{Q}_{n+1}\underline{\mathbf{H}}_n$, by computing the eigenvalues θ_j and eigenvectors $\mathbf{s}_j$ of the pencil $(\mathbf{H}_n, \mathbf{U}_n)$, followed by a prolongation $\mathbf{z}_j := \mathbf{Q}_n\mathbf{U}_n\mathbf{s}_j$ of the eigenvectors to obtain the Ritz vectors $\mathbf{z}_j$; see also [26, p. 595]. We only consider eigenvectors; the transition to handle principal vectors is straightforward and thus omitted. The residual of the unscaled Ritz pair $(\theta_j, \mathbf{z}_j)$ is then given by

$$\mathbf{A}\mathbf{z}_j - \mathbf{z}_j\theta_j = \mathbf{Q}_n(\mathbf{H}_n\mathbf{s}_j - \mathbf{U}_n\mathbf{s}_j\theta_j) + \mathbf{q}_{n+1}h_{n+1,n}\mathbf{e}_n^{\mathsf{T}}\mathbf{s}_j = \mathbf{q}_{n+1}h_{n+1,n}\mathbf{e}_n^{\mathsf{T}}\mathbf{s}_j. \tag{5.1}$$

As is the case in standard Hessenberg decompositions, the size of the last element $\mathbf{e}_n^{\mathsf{T}}\mathbf{s}_j$ of the eigenvector $\mathbf{s}_j$ of the pencil dominates the accuracy,

$$\|\mathbf{A}\mathbf{z}_j - \mathbf{z}_j\theta_j\|_2 = \|\mathbf{q}_{n+1}\|_2|h_{n+1,n}\mathbf{e}_n^{\mathsf{T}}\mathbf{s}_j|. \tag{5.2}$$

In contrast to the standard case, the eigenvector $\mathbf{s}_j$ is not only prolonged with the matrix $\mathbf{Q}_n$ spanning the basis, but beforehand multiplied with the upper triangular, possibly singular, $\mathbf{U}_n$ to obtain the Ritz vector $\mathbf{z}_j$. This affects the 2-norm backward error

$$\begin{aligned}\eta_{(\theta_j,\mathbf{z}_j)} &:= \min\{\eta \mid \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2 \leqslant \eta, \widetilde{\mathbf{A}}\mathbf{z}_j = \mathbf{z}_j\theta_j\} \\ &= \frac{\|\mathbf{A}\mathbf{z}_j - \mathbf{z}_j\theta_j\|_2}{\|\mathbf{z}_j\|_2} = \frac{\|\mathbf{q}_{n+1}\|_2|h_{n+1,n}\mathbf{e}_n^{\mathsf{T}}\mathbf{s}_j|}{\|\mathbf{z}_j\|_2}\end{aligned} \tag{5.3}$$

of the Ritz pair $(\theta_j, \mathbf{z}_j)$ whenever the length of $\mathbf{z}_j$ differs significantly from unity.

Using the Sonneveld pencil $(\mathbf{Y}_n^\circ, \mathbf{Y}_n\mathbf{D}_\omega^{(n)})$ (or, equivalently, the Sonneveld matrix $\mathbf{S}_n^\circ = \mathbf{Y}_n^\circ(\mathbf{D}_\omega^{(n)})^{-1}\mathbf{Y}_n^{-1}$) we can easily prolong the eigenvectors $\mathbf{s}_j^{\text{Sonneveld}}$, $1 \leqslant j \leqslant n$, to obtain Ritz vectors $\mathbf{z}_j^{\text{Sonneveld}} := \mathbf{R}_n\mathbf{Y}_n\mathbf{D}_\omega^{(n)}\mathbf{s}_j^{\text{Sonneveld}}$, $1 \leqslant j \leqslant n$, as we know the basis vectors, i.e., the residuals $\mathbf{r}_0, \ldots, \mathbf{r}_{n-1}$. Of course we do not store them, but similar to the course of action in Lanczos' method we can recompute them one by one if we store all the scalars (which saves some work). Thus, one chooses some Ritz values and computes afterwards the Ritz vectors. This choice usually is based on an interest in eigenvalues closest to some target or in some given region and some estimation of the accuracy obtained, mostly a measure or guess on the backward error. This can be done for the Sonneveld pencil or matrix representation, since we can compute the size of the residuals of the unscaled eigenpairs. Unfortunately, in this case, this is rather a *guess* of the accuracy of the eigenvalue, as the expression $\|\mathbf{r}_n\||s_{n+1,n}^\circ s_{nj}^{\text{Sonneveld}}|$ is an error bound for the residual of an *unnormalized* approximation to an eigenpair, where $\mathbf{r}_n$ and $s_{n+1,n}^\circ$ are the last column of the matrix $\mathbf{R}_{n+1}$ and the only nontrivial element in the last row of the Sonneveld matrix $\underline{\mathbf{S}}_n^\circ$ in (4.14), respectively.

To make use of the purified pencil $(\mathbf{Y}_n^\circ, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$, the deflated pencil $(D(\mathbf{Y}_n^\circ\mathbf{G}_n)$, $D(\mathbf{U}_n\mathbf{D}_\omega^{(n)}))$, or even the BiOres$(s,1)$ matrix $\mathbf{L}_m^\circ$, we have to prolong using the unknown vectors $\mathbf{w}_\ell$ or the unknown $\mathbf{q}_\ell$, i.e., a subset of the $\mathbf{w}_\ell$. These could be computed using the recurrences (4.42). The question arises as to how much the process of computing the vectors $\mathbf{w}_\ell$ in finite precision deteriorates; see also [35]. Naively implemented, this turns out to be a bad idea in finite precision, as the errors are amplified far beyond norm one and thus the computed vectors $\mathbf{w}_\ell$ are useless. Numerical experiments support the assumption that the computed $\mathbf{w}_\ell$ are accurate as long as the true and updated residuals in IDR(s)Ores are close; see (4.5) and

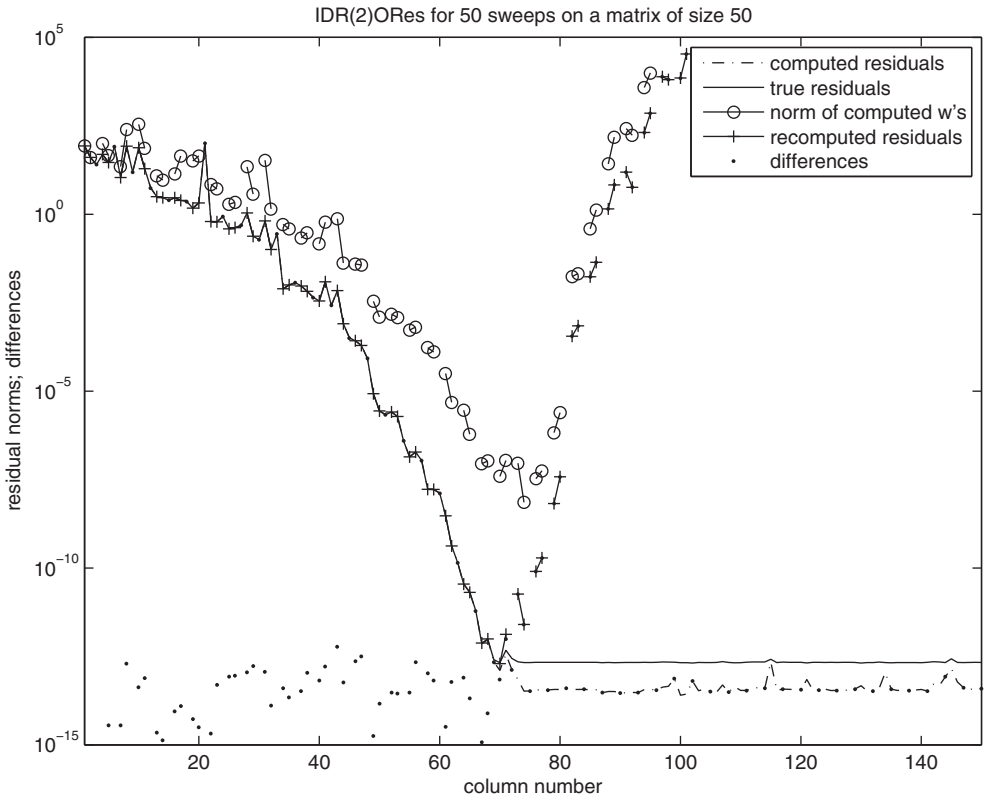


FIG. 5.1. *Quality of computed* $\mathbf{W}_n$ *measured by column norms. We compute the vectors* $\mathbf{w}$ *with the aid of the computed Lanczos-*BIORES$(s,1)$ *matrix* $\underline{\mathbf{L}}^\circ$*. To quickly check the quality of the obtained vectors we multiply them with the polynomials* $\Omega_j(\mathbf{A})$ *and compare the norms of the resulting recomputed residuals with the norms of the actually computed residuals* $\mathbf{r}$*.*

the remark thereafter. Unfortunately, there exist examples where this no longer holds true. A picture that compares the norms of the $\mathbf{w}_\ell$, computed using the matrix $\underline{\mathbf{L}}_n^\circ$ and $\mathbf{w}_0 = \mathbf{r}_0$ as starting vector, the differences between the recomputed $\mathbf{r}_\ell = \Omega_{\lfloor \ell/(s+1) \rfloor}(\mathbf{A})\mathbf{w}_\ell$, and the computed $\mathbf{r}_\ell$ is given in Figure 5.1. Of course, only those $\mathbf{w}_\ell$ are computed, which are needed for the prolongation, that is, omitting the columns $\ell = j(s+1)$ with the indices $\ell - 1$.

6. A numerical example. We used MATLAB Version 7.10.0.499 (R2010a) with machine precision $2^{-53} \approx 1.1102 \cdot 10^{-16}$ and the matrix `e05r0500` of dimension 236×236 with the corresponding right-hand side `e05r0500_rhs1` from the set "Driven Cavity" from Matrix Market. This matrix is a hard test case for Krylov subspace methods. As the starting guess we used $\mathbf{x}_0 = \mathbf{o}_{236}$. The parameters in IDR were $s = 4$ for 50 sweeps in IDRORES and orthonormalized random vectors constructed by `orth(randn(236,4))` in MATLAB. The convergence curves in Figure 6.1 show that IDRORES has not yet started to converge to the solution and that the true and updated residuals are close to each other.

The eigenvalue approximations obtained using the three different possible approaches are depicted in Figure 6.2. We used the eigenvalue approximations obtained using `eig(A)` in MATLAB as reference values. We remark that we obtain $200 = 4 \cdot 50$

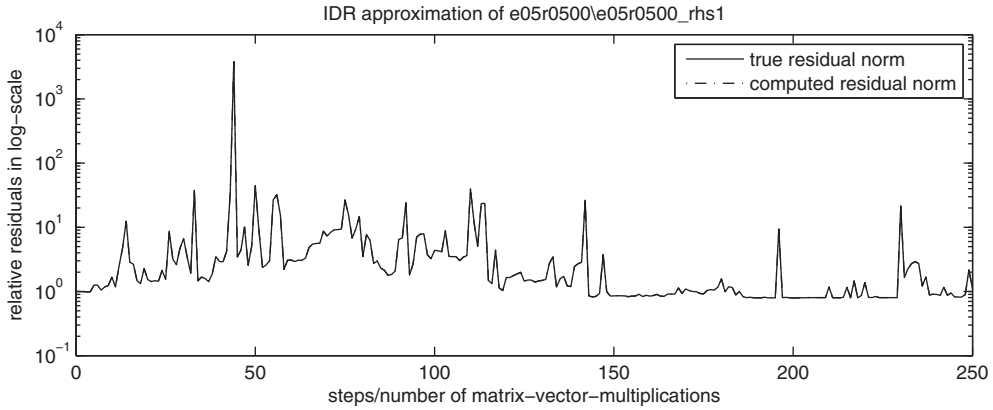


FIG. 6.1. *Convergence of* IDR(4)ORES *for the matrix* `e05r0500` *with right-hand side* `e05r0500_rhs1` *for* 50 *sweeps. The true and updated residuals coincide.*

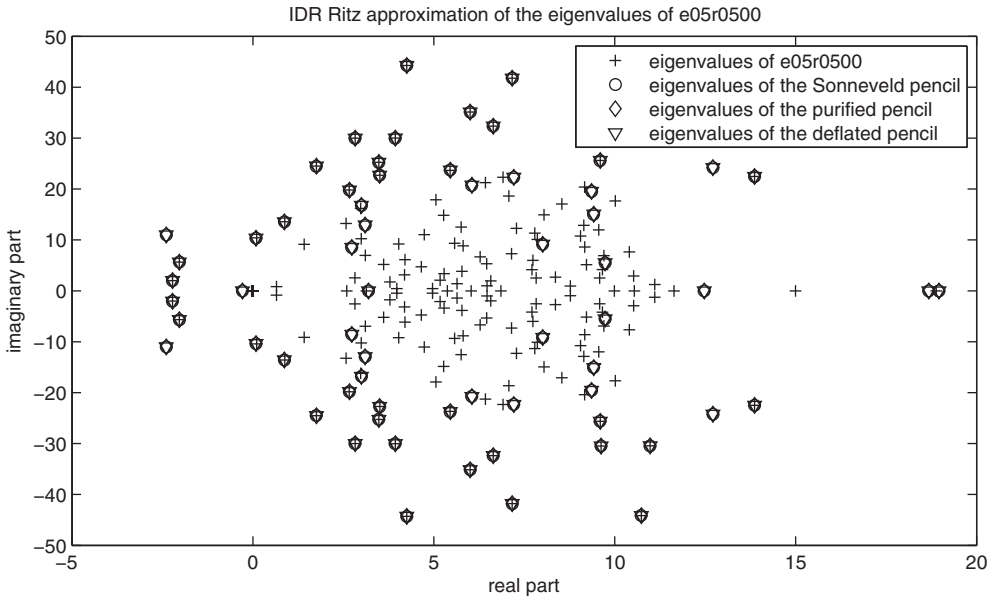


FIG. 6.2. *Eigenvalue approximations of* IDR(4)ORES *for the matrix* `e05r0500` *with right-hand side* `e05r0500_rhs1` *for* 50 *sweeps. Plusses depict the eigenvalues, circles depict the Ritz values of the Sonneveld pencil, diamonds depict the Ritz values of the purified pencil, and triangles depict the Ritz values of the deflated pencil. The outliers are approximated first.*

eigenvalue approximations with $250 = (4+1)\cdot 50$ matrix-vector multiplications. Obviously, many eigenvalue approximations are not visible in Figure 6.2. Our experiments showed that many ghost values occur at the positions of the roots $1/\omega_j$. This can be observed in Figure 6.3, where we depict the area of the roots $1/\omega_j$ together with the eigenvalue approximations.

To rate the quality of approximation, we computed the convergence history of all three approaches to the outlier $\lambda \approx 4.2505 - 44.2719i$ of `e05r0500`; see Figure 6.4. The deflated pencil is of smaller size. To be fair in comparison of matrix-vector multiplications we inserted the value `NaN` every $s+1$ steps starting with index s.

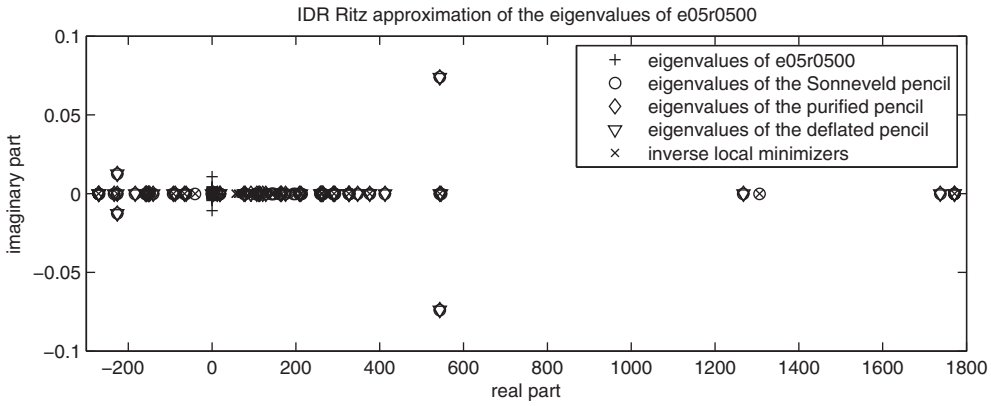


FIG. 6.3. *Eigenvalue approximations of* IDR(4)ORES *for the matrix* `e05r0500` *with right-hand side* `e05r0500_rhs1` *for* 50 *sweeps. Plusses depict the eigenvalues, circles depict the Ritz values of the Sonneveld pencil, diamonds depict the Ritz values of the purified pencil, triangles depict the Ritz values of the deflated pencil, crosses depict the inverse local minimizers. Many ghost eigenvalue approximations occur close to the inverse local minimizers.*

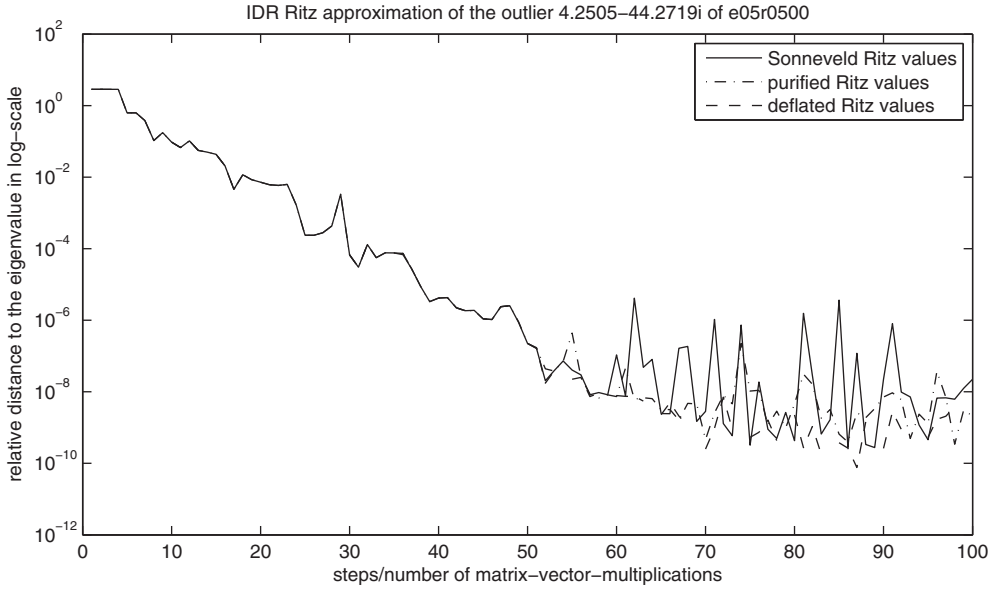


FIG. 6.4. *Eigenvalue approximations of* IDR(4)ORES *for the outlier at* $4.2505 - 44.2719\,\mathrm{i}$ *of the matrix* `e05r0500` *with right-hand side* `e05r0500_rhs1` *for* 100 *steps. The solid line depicts the convergence of the Sonneveld Ritz values, the dot-dashed line the convergence of the purified Ritz values, and the dashed line the convergence of the deflated Ritz values.*

All eigenvalue approximations converge to roughly the same relative accuracy. The approximations using the Sonneveld pencil behave a little more erratically; we suggest always using the purified pencil, as it is easily obtained from the Sonneveld pencil and does not share some of its unpleasant properties.

In this example all three eigenvalue approximations are more or less comparable. In the underlying report [14] we collected some more numerical examples that support

the use of the purified pencil. Since the report [14] has been published in 2010, the eigenvalue computations of more recent IDR variants have been analyzed based on it; [24] treats eigenvalue computations based on IDRBiO [41], whereas [25] treats eigenvalue computations based on IDRStab [38, 33]. It turns out that these newer approaches are also more stable in the case of eigenvalue computations. A genuine eigenvalue solver based on IDR still has to be developed; this is work in progress.

7. Conclusions. We showcased how to obtain eigenvalue approximations using IDR-based algorithms. The focus is on the development of the necessary new theory. Therefore much space is devoted to the so-called *generalized Hessenberg decomposition* and the relation of IDR(s) to BiORes($s,1$) and Lanczos($s,1$). The latter provide us with "correct" eigenvalue approximations, while the pencil of the generalized Hessenberg decomposition directly produced by IDR(s) also contains eigenvalues ω_k^{-1} that just represent the chosen smoothing parameters. The proposed spectral reduction procedure can also be applied to classical BiCGStab, but it is not needed there because the recurrence coefficients of BiCGStab yield directly those of the classical nonsymmetric Lanczos algorithm, in our notation Lanczos$(1,1)$—at least in exact arithmetic. When applied to BiCGStab, the spectral reduction gives *exactly* the same tridiagonal matrix comprising the recurrence coefficients, even in finite precision. In our numerical experiments the prototype IDR(s) [34] turned out to be less stable than the other more recent variants [41, 33]. For some interesting observations on the numerical behavior of IDRORes we refer the reader to our report [14]. Experiments with eigenvalue computations based on other IDR(s) variants indicated that these methods deteriorate in finite precision; in contrast to other Krylov subspace methods based on short recurrences, we observe no multiple Ritz values, but ghost Ritz values at the roots of the stabilizing polynomials. This still has to be analyzed. We omitted the computation of eigenvector approximations; these are discussed to some extent in the report [14]. As the topic of eigen*vector* computations based on IDR is even more technical, we decided to focus on eigen*value* computations in this article.

Appendix A. Hessenberg pencils and residual polynomials. In this appendix we compile for the reader's convenience various quantities that play a role in this paper. As stressed before, the transition from IDR(s)ORes to BiORes($s,1$) proceeds in two steps, a *purification step* and a *deflation step*, of which the second can be split up into two substeps. Table A.1 lists the various reduction steps and the associated residuals together with the corresponding extended Hessenberg pencils and extended Hessenberg matrices. Likewise, Table A.2 lists the residual polynomials together with the corresponding residuals and some relations between them.

Table A.1
Sets of residual matrices, extended Hessenberg pencils, and extended Hessenberg matrices.

IDR(s)	Residuals	Hessenberg pencil	Hessenberg matrix
Original	$\mathbf{R}_{n+1}$	$(\underline{\mathbf{Y}}_n^\circ, \mathbf{Y}_n\mathbf{D}_\omega^{(n)})$	$\underline{\mathbf{S}}_n^\circ := \underline{\mathbf{Y}}_n^\circ(\mathbf{D}_\omega^{(n)})^{-1}\mathbf{Y}_n^{-1}$
Purified	$\mathbf{W}_{n+1}$	$(\underline{\mathbf{Y}}_n^\circ, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$	does not exist
Elimin'd	$\mathbf{W}_{n+1}$	$(\underline{\mathbf{Y}}_n^\circ\mathbf{G}_n, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$	does not exist
Deflated	$\mathbf{Q}_{n+1}$	$(D(\underline{\mathbf{Y}}_n^\circ\mathbf{G}_n), D(\mathbf{U}_n\mathbf{D}_\omega^{(n)}))$	$\underline{\mathbf{L}}_n^\circ := D(\underline{\mathbf{Y}}_n^\circ\mathbf{G}_n)(D(\mathbf{U}_n\mathbf{D}_\omega^{(n)}))^{-1}$

TABLE A.2
Sets of residual matrices with residual vectors and corresponding residual polynomials.

Residual vectors	Residual polynomials
$\mathbf{R}_{n+1} = (\mathbf{r}_0, \ldots, \mathbf{r}_n)$	$\mathcal{R}_n(z) := \det(\mathbf{I}_n - z\,(\mathbf{S}_n^\circ)^{-1})$
never defined; not needed	$\Omega_j(z) := \prod_{k=1}^{j}(1-\omega_k z)$, $\Omega_0(z) \equiv 1$
$\mathbf{W}_{n+1} = (\mathbf{w}_0, \ldots, \mathbf{w}_n)$	$\mathcal{W}_n(z) := \det(\mathbf{I}_n - z\mathbf{U}_n\mathbf{D}_\omega^{(n)}\,(\mathbf{Y}_n^\circ)^{-1})$
$\mathbf{Q}_{n+1} = (\mathbf{q}_0, \ldots, \mathbf{q}_n)$	$\rho_n(z) := \det(\mathbf{I}_n - z\,(\mathbf{L}_n^\circ)^{-1})$
$= (\mathbf{w}_0, \ldots, \mathbf{w}_{s-1}, \mathbf{w}_{s+1}, \ldots)$	
Residual polynomial relations (including the BIORES$(s,1)$ polynomials)	
$\mathcal{R}_{j(s+1)+k}/\Omega_j = \mathcal{W}_{j(s+1)+k} = \rho_{js+k}, \quad 0 \leqslant k < s$ $\mathcal{R}_{j(s+1)+s}/\Omega_j = \mathcal{W}_{j(s+1)+s} = \widehat{\rho}_{(j+1)s}, \quad 0 \leqslant j < \lfloor (n+1)/(s+1) \rfloor$	

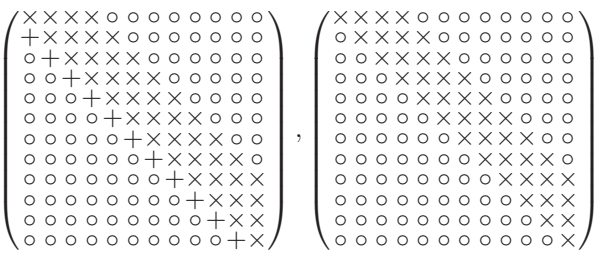
By abuse of notation the matrix $\mathbf{Q}_{n+1}$ denotes either a generic matrix of column vectors which form a basis of a (rational) Krylov subspace or the special matrix of deflated residuals. In the case of the purified residuals $\mathbf{w}_n$, there exists no corresponding unreduced Hessenberg decomposition, and we use two different denominations of tantamount importance for the residual polynomials. One denomination, $\mathcal{W}_n$, better reflects the construction of the basis vectors and uniformly uses the column indices; the other, involving both ρ_k and $\widehat{\rho}_k$, is closer to the polynomial point of view: k is the degree of the polynomial and the "hat" indicates that there are two different types of polynomials in use. We remark that the polynomials $\mathcal{W}_n := \mathcal{R}_n/\Omega_{\lfloor n/(s+1)\rfloor}$ are obviously residual polynomials, i.e., $\mathcal{W}_n(0) = 1$, since $\Omega_{\lfloor n/(s+1)\rfloor}(0) = 1$ and $\mathcal{R}_n(0) = 1$. In the last row of Table A.1, D denotes the deflation operator that removes every $(s+1)$th row and column.

Figure A.1 illustrates with a small example where $n = 12$ and $s = 3$ how the various transitions affect the pencils. The deflation operator D is responsible for the reduction of the size in the last step. The third pencil $(\mathbf{Y}_n^\circ\mathbf{G}_n, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$ is obtained from the second one, the purified pencil $(\mathbf{Y}_n^\circ, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$, by a multiplication from the right by the block-Gauss eliminator $\mathbf{G}_n$. For this reason we refer to it as the *eliminated pencil*. The deflated pencil consists of two block matrices with $s \times s$ blocks.[8] The matrix $D(\mathbf{U}_n\mathbf{D}_\omega^{(n)})$ is block diagonal, the upper triangular diagonal blocks are defined directly in terms of (differences of) IDR(s)ORES quantities. We stress the fact that the matrix $D(\mathbf{Y}_n^\circ\mathbf{G}_n)$ is simultaneously *unreduced Hessenberg* and *block tridiagonal*. The diagonal blocks of $D(\mathbf{Y}_n^\circ\mathbf{G}_n)$ are unreduced Hessenberg matrices, and only the first rows and last columns are altered in the process of purification and deflation.

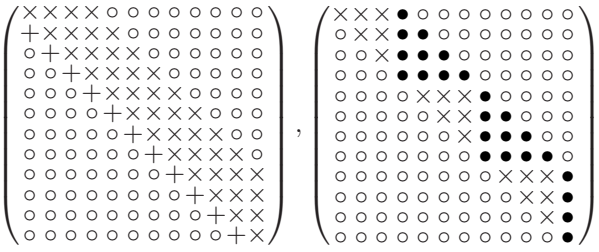
Acknowledgments. The authors would like to thank two anonymous referees for their critical remarks that helped to improve the presentation of the paper. We also would like to thank Martin B. van Gijzen, Olaf Rendel, Anisa Rizvanolli, and Gerard L.G. Sleijpen for their helpful comments on the report [14]; the second author is indebted to Olaf Rendel and Anisa Rizvanolli for their interest in applying the theory developed in [14] to other variants of IDR(s).

[8] The last block column has fewer columns than s, if n or $n+1$ is not divisible by $s+1$. This is not reflected in this educational example.

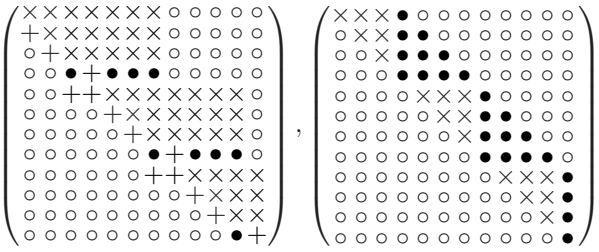
Original IDR(s)ORes pencil, the Sonneveld pencil $(\mathbf{Y}_n^\circ, \mathbf{Y}_n\mathbf{D}_\omega^{(n)})$:



Purified IDR(s)ORes pencil $(\mathbf{Y}_n^\circ, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$:



Eliminated IDR(s)ORes pencil $(\mathbf{Y}_n^\circ\mathbf{G}_n, \mathbf{U}_n\mathbf{D}_\omega^{(n)})$:



Deflated IDR(s)ORes pencil $(D(\mathbf{Y}_n^\circ\mathbf{G}_n), D(\mathbf{U}_n\mathbf{D}_\omega^{(n)}))$:

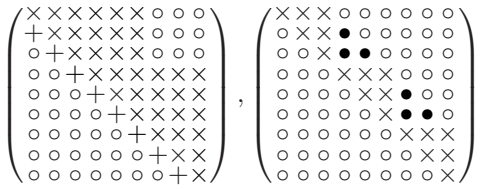


FIG. A.1. *The transition from* IDR(s)ORes *to* BiORes(s, 1)*, here depicted pencilwise for* $s = 3$*. Circles and bullets mark zeros, crosses mark entries defined in terms of* IDR(s)ORes *quantities, and plusses mark nonzero values. Black bullet zero elements are those introduced in the purification step (right-hand side) and deflation step (left-hand side).*

REFERENCES

[1] A. M. Abdel-Rehim, R. B. Morgan, D. A. Nicely, and W. Wilcox, *Deflated and restarted symmetric Lanczos methods for eigenvalues and linear equations with multiple right-hand sides*, SIAM J. Sci. Comput., 32 (2010), pp. 129–149.

[2] J. I. Aliaga, D. L. Boley, R. W. Freund, and V. Hernández, *A Lanczos-type method for multiple starting vectors*, Math. Comp., 69 (2000), pp. 1577–1601.

[3] W. E. Arnoldi, *The principle of minimized iterations in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.
[4] S. F. Ashby, T. A. Manteuffel, and P. E. Saylor, *A taxonomy for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568.
[5] Z. Bai, D. Day, and Q. Ye, *ABLE: An adaptive block Lanczos method for non-Hermitian eigenvalue problems*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 1060–1082.
[6] T. Ericsson, *On the Eigenvalues and Eigenvectors of Hessenberg Matrices*, Numerical Analysis Group, Göteborg, Report 10, Chalmers University of Technology and the University of Göteborg, Department of Computer Sciences, Göteborg, Sweden, 1990. Available online from http://www.cs.chalmers.se/pub/num_analysis/reports/.
[7] D. K. Faddeev, *Some properties of a matrix that is the inverse of a Hessenberg matrix*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI), 111 (1981), pp. 177–179, 238 (in Russian).
[8] D. K. Faddeev, *Properties of a matrix, inverse to a Hessenberg matrix*, Journal of Soviet Mathematics, 24 (1984), pp. 118–120.
[9] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Numerical Analysis, G. A. Watson, ed., Lecture Notes in Math. 506, Springer, Berlin, 1976, pp. 73–89.
[10] J. N. Franklin, *Matrix Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1968.
[11] R. W. Freund and M. Malhotra, *A block QMR algorithm for non-Hermitian linear systems with multiple right-hand sides*, Linear Algebra Appl., 254 (1997), pp. 119–157.
[12] R. Freund, *Hermitian eigenvalue problems: Band Lanczos method*, in Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide, Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, eds., SIAM, Philadelphia, section 4.6, 2000, pp. 80–88.
[13] R. Freund, *Non-Hermitian eigenvalue problems: Band Lanczos method*, in Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide, Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, eds., SIAM, Philadelphia, section 7.10, 2000, pp. 205–216.
[14] M. H. Gutknecht and J.-P. M. Zemke, *Eigenvalue computations based on IDR*, Bericht 145, TUHH, Institute of Numerical Simulation, 2010; also available online from http://doku.b.tu-harburg.de/volltexte/2010/875/.
[15] M. H. Gutknecht, *Stationary and almost stationary iterative (k,l)-step methods for linear and nonlinear systems of equations*, Numer. Math., 56 (1989), pp. 179–213.
[16] M. H. Gutknecht, *The unsymmetric Lanczos algorithms and their relations to Padé approximation, continued fractions, and the qd algorithm*, in Preliminary Proceedings of the Copper Mountain Conference on Iterative Methods, Copper Mountain, CO, 1990; available online from http://www.sam.math.ethz.ch/~mhg/pub/CopperMtn90.pdf and http://www.sam.math.ethz.ch/~mhg/pub/CopperMtn90-7.pdf.
[17] M. H. Gutknecht, *Lanczos-type solvers for nonsymmetric linear systems of equations*, Acta Numer., 6 (1997), pp. 271–397.
[18] M. H. Gutknecht, *IDR explained*, Electron. Trans. Numer. Anal., 36 (2010), pp. 126–148.
[19] K. Hessenberg, *Behandlung linearer Eigenwertaufgaben mit Hilfe der Hamilton-Cayleyschen Gleichung*, Numerische Verfahren, Bericht 1, Institut für Praktische Mathematik (IPM), Technische Hochschule Darmstadt, July 1940. Scanned report and biographical sketch of Karl Hessenberg's life available online from http://www.hessenberg.de/karl1.html.
[20] A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, Toronto, London, 1964.
[21] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Nat. Bur. Standards, 45 (1950), pp. 255–282.
[22] C. Lanczos, *Solution of systems of linear equations by minimized iterations*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 33–53.
[23] D. Loher, *Reliable Nonsymmetric Block Lanczos Algorithms*, Ph.D. thesis, ETH Zurich, Zurich, Switzerland, 2006.
[24] O. Rendel, *Aspects of Eigenvalue Computations using Induced Dimension Reduction (IDR)*, Bachelor's thesis, Technische Universität Hamburg-Harburg, Institut für Numerische Simulation, 2010.
[25] A. Rizvanolli, *Eigenwertberechnung mittels IDRStab*, Studienarbeit, Technische Universität Hamburg-Harburg, Institut für Numerische Simulation, 2011.
[26] A. Ruhe and D. Skoogh, *Rational Krylov algorithms for eigenvalue computation and model reduction*, in Applied Parallel Computing, Lecture Notes in Comput. Sci. 1541, Springer, Berlin, 1998, pp. 491–502.
[27] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[28] Y. Saad, *Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices*, Linear Algebra Appl., 34 (1980), pp. 269–295.
[29] J. Schur, *Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind*, J. Reine Angew. Math., 147 (1917), pp. 205–232.
[30] F. Schweins, *Theorie der Differenzen und Differentiale*, Verlag der Universitäts-Buchhandlung von C. F. Winter, Heidelberg, 1825; also available online from http://books.google.de/books?id=dntNAAAAMAAJ.
[31] V. Simoncini and D. B. Szyld, *Interpreting IDR as a Petrov-Galerkin method*, SIAM J. Sci. Comput., 32 (2010), pp. 1898–1912.
[32] G. L. Sleijpen, P. Sonneveld, and M. B. van Gijzen, *Bi-CGSTAB as an induced dimension reduction method*, Appl. Numer. Math., 60 (2010), pp. 1100–1114.
[33] G. L. G. Sleijpen and M. B. van Gijzen, *Exploiting BiCGstab(ℓ) strategies to induce dimension reduction*, SIAM J. Sci. Comput., 32 (2010), pp. 2687–2709.
[34] 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 (2009), pp. 1035–1062.
[35] P. Sonneveld, *On the convergence behavior of IDR(s) and related methods*, SIAM J. Sci. Comput., 34 (2012), pp. A2576–A2598.
[36] E. Stiefel, *Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme*, Comment. Math. Helv., 29 (1955), pp. 157–179.
[37] M. Tanio and M. Sugihara, *GIDR(s,L): Generalized IDR(s)*, in Proceedings of the 2008 Annual Conference of the Japan Society for Industrial and Applied Mathematics, Chiba, Japan, September 2008, pp. 411–412 (in Japanese).
[38] M. Tanio and M. Sugihara, *GBi-CGSTAB(s, L): IDR(s) with higher-order stabilization polynomials*, J. Comput. Appl. Math., 235 (2010), pp. 765–784.
[39] H. A. van der Vorst and P. Sonneveld, *CGSTAB, A More Smoothly Converging Variant of CG-S*, Report 90-50, Department of Mathematics and Informatics, Delft University of Technology, 1990.
[40] H. A. van der Vorst, *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631–644.
[41] M. B. van Gijzen and P. Sonneveld, *Algorithm* 913: *An elegant IDR(s) variant that efficiently exploits biorthogonality properties*, ACM Trans. Math. Software, 38 (2011), 5.
[42] P. Wesseling and P. Sonneveld, *Numerical experiments with a multiple grid and a preconditioned Lanczos type method*, in Approximation Methods for Navier-Stokes Problems, Lecture Notes in Math. 771, R. Rantmann, ed., Springer-Verlag, Berlin, Heidelberg, New York, 1980, pp. 543–562.
[43] M.-C. Yeung and T. F. Chan, *ML(k)BiCGSTAB: A BiCGSTAB variant based on multiple Lanczos starting vectors*, SIAM J. Sci. Comput., 21 (1999), pp. 1263–1290.
[44] D. M. Young and K. C. Jea, *Generalized conjugate-gradient acceleration of nonsymmetrizable iterative methods*, Linear Algebra Appl., 34 (1980), pp. 159–194.
[45] J.-P. M. Zemke, *Hessenberg eigenvalue–eigenmatrix relations*, Linear Algebra Appl., 414 (2006), pp. 589–606.
[46] J.-P. M. Zemke, *Abstract perturbed Krylov methods*, Linear Algebra Appl., 424 (2007), pp. 405–434.
[47] J.-P. M. Zemke, *IDR(s) and IDR(s)Eig in parallel computing*, Supercomput. News, 12 (2010), pp. 31–48.
[48] F. Zhang, ed., *The Schur complement and its applications*, vol. 4, Numerical Methods and Algorithms, Springer, New York, 2005.