TWO-SIDED ARNOLDI AND NONSYMMETRIC LANCZOS ALGORITHMS

JANE CULLUM† AND TONG ZHANG‡

Abstract. We introduce new two-sided Arnoldi recursions and use them to define a model reduction procedure for large, linear, time-invariant, multi-input/multi-output differential algebraic systems. We prove that this procedure has desirable moment matching properties. We define a corresponding model reduction procedure which is based upon a band nonsymmetric Lanczos recursion and prove that if the deflation is exact and there are no breakdowns in the recursions, that these two model reduction procedures generate identical reduced-order systems. We prove similar equivalences for corresponding eigenelement procedures. We concentrate on the theoretical properties of the new algorithms.

Key words. two-sided Arnoldi, nonsymmetric Lanczos, equivalences, relationships, iterative methods, reduced-order systems, eigenvalues, model reduction

AMS subject classifications. 15A06, 15A18, 65F, 65L, 93C

1. Introduction. This paper arose out of our work on model reduction algorithms for large multi-input/multi-output (mi/mo), time-invariant, delay-differential-algebraic systems of equations which occur in modeling VLSI interconnects (wires, planes, conductors)[7] The inner loop of the outer/inner loop procedure developed in [7] repeatedly exercises an iterative model reduction procedure for time-invariant linear systems, [10]. This inner algorithm must be able to handle systems with arbitrary numbers of inputs and outputs.

Single-sided Arnoldi model reduction methods have been proposed which directly use the system block input matrix in the iterative method but do not directly use the system block output matrix. We define two-sided Arnoldi recursions with the capability of directly incorporating both of these matrices into an iterative model reduction procedure.

The two-sided nature of these recursions leads us to a comparison of a model reduction procedure which is based upon the new two-sided Arnoldi recursions and a corresponding procedure which is based upon the nonsymmetric band Lanczos recursion developed in [1]. We prove that these two procedures generate identical reduced-order models. We also prove that the iterates generated by corresponding eigenelement procedures are identical. We focus on the theoretical properties of these procedures.

In Section 2 we define a two-sided block Arnoldi recursion which consists of two independent applications of a corresponding one-sided block Arnoldi recursion and a computation which combines the quantities generated by these two applications. One application uses the system matrix with the system block input matrix. The second application uses the transpose of the system matrix with the system block output matrix. We illustrate some properties of this basic two-sided Arnoldi recursion, including the very interesting possibility of recovery from breakdown without any modifications

†MS B250, Los Alamos National Laboratory, Los Alamos, NM 87545. e-mail: cullumj@lanl.gov.
Partial support was provided by the Department of Energy under contract W-7405-ENG-36 and the Office of Science, MICS, Applied Mathematical Sciences Program grant KC-07-01-01.
‡IBM Research Center, Yorktown Heights, NY 10598. e-mail: tzhang@watson.ibm.com

of these recursions. This two-sided block Arnoldi recursion is an extension of work in [16].

In Section 3 we use the recursions in Section 2 to define a two-sided block Arnoldi model reduction procedure for generating approximations to transfer functions of large systems of time-invariant, differential-algebraic equations. We prove that these approximations possess desirable matrix moment matching properties [10]. The proof is interesting because it is self-contained and uses only general properties of two-sided Krylov recursions. For example, it does not use the orthogonality of the associated vectors.

In Section 4, we review briefly some of the properties of the band nonsymmetric Lanczos recursions defined in [1]. Assuming exact arithmetic and exact deflation, we derive common properties of the band Lanczos and of the two-sided block Arnoldi recursions. We exploit those properties within the context of corresponding model reduction and eigenelement methods to prove that corresponding methods generate identical approximations. These results complement the earlier work in [4], [3], comparing one-sided Arnoldi methods and nonsymmetric Lanczos methods for solving $Ax = b$ or $Ax = \lambda x$.

Assumptions: Unless it is stated explicitly otherwise, in any discussion of any Arnoldi or nonsymmetric Lanczos-based procedure, we will assume that all of the required quantities are well-defined; no breakdowns occur; the underlying recursions do not terminate prematurely; and the arithmetic is exact.

1.1. Notation.. We summarize the notation which is nonstandard.

si/so: single input, single output system
mi/mo: multiple input, multiple output system
r, l: subscripts (superscripts) to denote quantities associated with right and left vectors which were generated using A and A^T.
$B(:,[i:j])$ $[B([i:j],:)]$: columns[rows] i through j of matrix B
$I_{[i:j]}$: columns i through j of an identity matrix I_K where K is specified within the local context.
$I_{[xlast]}$, $I_{[xfirst]}$: denote corresponding $I_{[i:j]}$ where the block $[i:j]$ corresponds to the indices in the last(first) block column in an associated block structure. Setting $x = r, l(v, w)$, indicates right or left quantities for Arnoldi(Lanczos) recursions. .
$H([i:j],[k:l])$: submatrix of H consisting of the intersection of rows i through j with columns k through l.
$\left[Q_j, \widetilde{Q}_{j+1}\right]$: equals $\left[\widetilde{Q}_1, \ldots, \widetilde{Q}_j, \widetilde{Q}_{j+1}\right]$ for $Q_j \equiv \left[\widetilde{Q}_1, \ldots, \widetilde{Q}_j\right]$

2. Arnoldi Recursions. Blocks occur naturally in multi-input/multi-output mi/mo systems. System inputs and outputs are controlled by matrix blocks, and we measure the quality of our proposed reduced-order model by the number of rectangular block moments of the transfer function of the original system which are matched by corresponding moments of the transfer function of the reduced-order system. Therefore, initially we focus on block, two-sided Arnoldi recursions.

2.1. A Block Arnoldi Recursion. Given a matrix A, a consistent starting block of vectors X with d_r columns, and a deflation tolerance ϵ_d, we define a one-

sided, block Arnoldi recursion [17]. At each iteration of this recursion we are working with a block of vectors which was generated by applying A to a block of vectors which was generated at an earlier iteration and invoking orthogonalizations. It is possible and typical, that as the iterations proceed, that one or more vectors within the current block become dependent or nearly dependent upon vectors which have already been generated. In this situation, to preserve the integrity of the procedure, the (nearly) dependent vector(s) must be deflated from the process. Deflation is accomplished implicitly without any explicit permutations of vectors or modifications of the recursions. Example 2.1 provides a concrete illustration of the use of deflation in a block. For more details on deflation, see [5], [8], [1].

In the statement of Algorithm 2.1 we use Q_{j+1} to denote the Arnoldi vectors after j block steps of the recursion, d_j to denote the size of the j^{th} block, $\widetilde{Q}_j$, $s_m \equiv \sum_{j=1}^{m} d_j$ equals the number of columns in Q_m, and e_d is the deflation tolerance.

ALGORITHM 2.1. (*Block Arnoldi Recursion*)
Specify A, X, $\epsilon_d \geq 0$.
Decompose $X = \widetilde{Q}_1 S_r + \tilde{\Delta}_r$ *where* $\widetilde{Q}_1^T \widetilde{Q}_1 = I$
  and $\|\tilde{\Delta}_r\|_F \leq \sqrt{(d_r - d_1)}\epsilon_d$ *with* $d_1 = \mathrm{rank}\,(\widetilde{Q}_1)$.
Set $s = d = d_1$
$Y_1 = \widetilde{Q}_1$
$Q_1 = [\widetilde{Q}_1]$.
for $j = 1 : m$
  $Q(:, [s - d + 1 : s]) = Y_j$.
  $P_j = AY_j$.
  $H([1 : s], [s - d + 1 : s]) = Q_j^T P_j$.
  $P_j = P_j - Q_j H([1 : s], [s - d + 1 : s])$.
  if $j < m$
    $P_j = \widetilde{Q}_{j+1} S_j + \widetilde{\Delta}_j$ *where* $\widetilde{Q}_{j+1}^T \widetilde{Q}_{j+1} = I$
      and $\|\widetilde{\Delta}_j\|_F < \sqrt{d - d_{j+1}}\epsilon_d$ *with* $d_{j+1} = \mathrm{rank}\,(\widetilde{Q}_{j+1})$.
    $H([s + 1 : s + d_j], [s - d + 1 : s]) \equiv S_j$.
    $d = d_{j+1}$
    $s = s + d$
    $Q_{j+1} = [Q_j, \widetilde{Q}_{j+1}]$.
  else
    $R_m = P_m$
    $H_m \equiv H([1 : s_m], [1 : s_m])$.
  endg
end

H_m denotes the square block upper Hessenberg matrix generated with diagonal blocks of size d_j, $1 \leq j \leq m$. R_m denotes the final $n \times d_m$ block residual matrix. The matrix form for these recursions is

$$AQ_m = Q_m H_m + R_m I_{[last]}^T. \tag{2.1}$$

where $I_{[last]}$ denotes $I_{[s_m - d_m + 1 : s_m]}$. We will also use $I_{[first]}$ to denote $I_{[1:d_1]}$. If Algorithm 2.1 is applied to $\{A, X_r\}$ we use $\{X_r, S_r, Q_r, H_r, R_r, I_{[rlast]}, I_{[rfirst]}\}$. Similarly, for $\{A^T, X_l\}$ we use $\{X_l, S_l, Q_l, H_l, R_l, I_{[llast]}, I_{[lfirst]}\}$.

If for some j, $d_j - d_{j+1} > 0$, then deflation has occurred in that block P_j. Checks for deflation are accomplished by applying modified Gram-Schmidt orthogonalization within each block and deflating any vector with norm smaller or equal to ϵ_d. Since

every right(left) Arnoldi vector generated is explicitly orthogonalized w.r.t all existing right(left) vectors, modifications in the recursion formulas are not required when deflation occurs.

Example 2.1 illustrates some possible deflation scenarios. In the implementations the vectors $p_j^{(k)}$ are ignored. Vectors are considered in their natural order and in place. For simplicity, in this example we assume that $\|p_1\| > \epsilon_d$.

EXAMPLE 2.1. *Let $P = [p_1, p_2, p_3]$ be a block of 3 vectors and consider the vectors in order. We have the following possible steps. The γ_{ij} denote the Gram-Schmidt orthogonalization coefficients.*

1. $q_1 \equiv p_1/\|p_1\|$; $\quad p_2^{(2)} \equiv p_2 - \gamma_{21} q_1$; $\quad p_3^{(2)} \equiv p_3 - \gamma_{31} q_1$; $\quad$ *go to step 2.*
2. *There are three cases.*
   a. *If $\|p_2^{(2)}\| \leq \epsilon_d$ and $\|p_3^{(2)}\| > \epsilon_d$, set $q_2 \equiv p_3^{(2)}/\|p_3^{(2)}\|$ and terminate.*
   b. *If $\|p_2^{(2)}\| > \epsilon_d$, set $q_2 \equiv p_2^{(2)}/\|p_2^{(2)}\|$; $p_3^{(3)} = p_3^{(2)} - \gamma_{32} q_2$; go to step 3.*
   c. *If $\max(\|p_2^{(2)}\|, \|p_3^{(2)}\|) \leq \epsilon_d$, terminate.*
3. *There are two cases.*
   a. *If $\|p_3^{(3)}\| \leq \epsilon_d$, terminate.*
   b. *If $\|p_3^{(3)}\| > \epsilon_d$, set $q_3 \equiv p_3^{(3)}/\|p_3^{(3)}\|$ and terminate.*

The following combinations of steps and matrix block relationships are possible.

A. $\{1, 2a\} \Rightarrow [p_1, p_2, p_3] = [q_1, q_2] \begin{bmatrix} \|p_1\| & \gamma_{21} & \gamma_{31} \\ 0 & 0 & \|p_3^{(2)}\| \end{bmatrix} + [0, p_2^{(2)}, 0]$

B. $\{1, 2b, 3a\} \Rightarrow [p_1, p_2, p_3] = [q_1, q_2] \begin{bmatrix} \|p_1\| & \gamma_{21} & \gamma_{31} \\ 0 & \|p_2^{(2)}\| & \gamma_{32} \end{bmatrix} + [0, 0, p_3^{(3)}]$

C. $\{1, 2b, 3b\} \Rightarrow [p_1, p_2, p_3] = [q_1, q_2, q_3] \begin{bmatrix} \|p_1\| & \gamma_{21} & \gamma_{31} \\ 0 & \|p_2^{(2)}\| & \gamma_{32} \\ 0 & 0 & \|p_3^{(3)}\| \end{bmatrix}$

D. $\{1, 2c\} \Rightarrow [p_1, p_2, p_3] = [q_1] \begin{bmatrix} \|p_1\| & \gamma_{21} & \gamma_{31} \end{bmatrix} + [0, p_2^{(2)} p_3^{(2)}]$

Since the vectors within a candidate block P_j are considered in the natural order, and modified Gram Schmidt orthogonalization is applied successively to each of these vectors, the first d_{j+1} columns of the subblock below each j^{th} diagonal block in H_m form an upper triangular matrix. Therefore, we can truncate this matrix interior to a $(j+1)^{st}$ block and the truncated matrix $\widehat{H}$ retains the block upper Hessenberg structure. The corresponding block residual matrix $\widehat{R}$ for $\widehat{H}$ will have the same number of columns as the residual corresponding to the H matrix with $(j+1)$ complete blocks. The indices of the columns corresponding to $\widehat{R}$ are obtained by shifting the column indices of the $(j+1)^{st}$ block left by the number of columns truncated from that block. We will exercise this ability to truncate and retain structure in Section 4 in our comparisons of methods which are based upon a two-sided block Arnoldi recursion with corresponding methods which are based upon the nonsymmetric band Lanczos recursion in [1].

2.2. A Two-Sided Block Arnoldi Recursion. We construct a two-sided block Arnoldi recursion by combining two independent applications of Algorithm 2.1, to $\{A, X_r\}$ and to $\{A^T, X_l\}$, with an appropriate *vector merge* of the resulting left and right Arnoldi vectors, $\{Q_l, Q_r\}$. The merge creates a modified right(left) residual matrix that is biorthogonal to the left(right) Arnoldi vectors. To maintain the equalities, the modification to the residual matrix must also be applied to the corresponding H matrix.

ALGORITHM 2.2. (*Two-Sided Block Arnoldi Recursion*)

Specify A, X_r, X_l, ϵ_d, m_r, m_l.
Apply Algorithm 2.1 to $\{A, X_r\}$ *for* m_r *block steps to generate* Q_r.
Apply Algorithm 2.1 to $\{A^T, X_l\}$ *for* m_l *block steps to generate* Q_l.
Compute $\overline{Q} = Q_l^T Q_r$.
If $\text{rank}\,(Q_l) \leq \text{rank}\,(Q_r)$
  Solve $\overline{Q} Z = Q_l^T R_r$.
  Set $\overline{H} \equiv [\overline{H}]_r = H_r + Z I_{[rlast]}^T$.
else
  Solve $\overline{Q}^T Z = Q_r^T R_l$.
  Set $\overline{H} \equiv [\overline{H}]_l = (H_l + Z I_{[llast]}^T)^T$.
end

2.3. Properties and Breakdown. A two-sided Arnoldi recursion possesses several important properties. It is easy to implement. Each one-sided block recursion is well-defined for blocks of any size, and deflation can occur independently in either or both single-sided recursions at any point in the computations. There is no requirement that $\text{rank}\,(Q_r) = \text{rank}\,(Q_l)$ and typically they are not equal.

The vector merge computations require $\overline{Q} \equiv Q_l^T Q_r$ to have full rank. The column block Z which is generated in a merge is incorporated into the appropriate H matrix. If the right recursions are used, this modified matrix equals $\overline{H}$. If the left recursions are used, the transpose of this modified matrix equals $\overline{H}$. $\overline{H}$ is a representation of A and these matrices can be used to define two-sided iterative methods involving A.

If $\overline{Q} \equiv Q_l^T Q_r$ does not have full rank, then the merge operation cannot be accomplished and $\overline{H}$ cannot be generated. Breakdown in Algorithm 2.2 occurs. A near-breakdown will exhibit itself as a nearly rank deficient $\overline{Q}$. However, as illustrated in Example 2.2, even if breakdown occurs, there is the possibility of recovery from breakdown without modifications of the recursions. The recursions can simply be continued until the corresponding $\overline{Q}$ has full rank.

EXAMPLE 2.2. *Apply Algorithm 2.2 to*

$$A = \begin{bmatrix} 5 & 12 & 38 & -21 \\ 3 & 8 & 24 & -13 \\ -2 & -6 & -19 & 12 \\ -1 & -4 & -12 & 8 \end{bmatrix}, \quad l \equiv X_l = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \quad r \equiv X_r = \begin{bmatrix} 7 \\ 4 \\ -3 \\ -2 \end{bmatrix}$$

After two steps of the recursion, we obtain the left and right vectors:

$$Q_l \equiv \begin{bmatrix} l/\sqrt{3}, & [2,1,6,-1]^T/\sqrt{42}, & [-4,19,2,23]^T/\sqrt{910} \end{bmatrix}$$
$$Q_r \equiv \begin{bmatrix} r/\sqrt{78}, & [-4,7,-2,3]^T/\sqrt{78}, & [1,1,3,1]^T/\sqrt{12} \end{bmatrix}.$$

If we stop after one step of the recursion, and attempt a merge, we encounter breakdown. However, if we ignore the breakdown and continue the recursions one more step, then $\overline{Q}$ *has full rank.*

The merge operation in Algorithm 2.2 can accept any number of left and right vectors. As indicated earlier the H matrices generated by a one-sided band Arnoldi recursion can be truncated and still retain the block upper Hessenberg structure in the truncated H-matrix and the block structure for the corresponding residual. Therefore, we can restate Algorithm 2.2 to allow for such truncation. We will use this flexibility in Section 4 where we prove the equivalence of methods based upon the nonsymmetric band Lanczos recursion in [1] and Algorithm 2.3 which is a truncated version of Algorithm 2.2.

ALGORITHM 2.3. (*Truncated Two-Sided Block Arnoldi Recursion*)
Specify A, X_r, X_l, ϵ_d, s.
Apply Algorithm 2.1 to $\{A, X_r\}$ *to generate* Q_r *with* $\text{rank}\,(Q_r) \geq s$.
Apply Algorithm 2.1 to $\{A^T, X_l\}$ *to generate* Q_l *with* $\text{rank}\,(Q_l) \geq s$.
Set $\widehat{Q}_r = Q_r(:, [1:s])$, $\widehat{Q}_l = Q_l(:, [1:s])$, $\widehat{H}_r = H_r([1:s], [1:s])$, $\widehat{H}_l = H_l([1:s], [1:s])$.
Define $\widehat{R}_l, \widehat{R}_r$ *as the residual blocks needed to preserve equality in the recursions.*
Compute $\overline{\widehat{Q}} = \widehat{Q}_l^T \widehat{Q}_r$.
Solve $\overline{\widehat{Q}} \widehat{Z}_r = \widehat{Q}_l^T \widehat{R}_r$.
Solve $\overline{\widehat{Q}}^T \widehat{Z}_l = \widehat{Q}_r^T \widehat{R}_l$.
Set $[\overline{\widehat{H}}]_r = \widehat{H}_r + \widehat{Z}_r I_{[rlast]}^T$.
Set $[\overline{\widehat{H}}]_l = (\widehat{H}_l + \widehat{Z}_l I_{[llast]}^T)^T$.
end

If we exercise Algorithm 2.3, Lemma 2.1 states that the resulting $\overline{\widehat{H}}$ matrices are Petrov-Galerkin projections of A [18]. There is the implicit assumption that $\overline{\widehat{Q}}$ is nonsingular.

LEMMA 2.1. *Let* $\widehat{Q}_l$, $\widehat{Q}_r$, $[\overline{\widehat{H}}]_r$ *and* $[\overline{\widehat{H}}]_l$ *be generated by Algorithm 2.3, then*

$$
\begin{aligned}
[\overline{\widehat{H}}]_r &\equiv \overline{\widehat{Q}}^{-1} \widehat{Q}_l^T A \widehat{Q}_r \\
[\overline{\widehat{H}}]_l &\equiv \widehat{Q}_l^T A \widehat{Q}_r \overline{\widehat{Q}}^{-1} \\
[\overline{\widehat{H}}]_l &= \overline{\widehat{Q}} [\overline{\widehat{H}}]_r \overline{\widehat{Q}}^{-1}.
\end{aligned} \tag{2.2}
$$

$[\overline{\widehat{H}}]_r$ *and* $[\overline{\widehat{H}}]_l$ *are oblique projections of* A *onto* $\text{span}\{\widehat{Q}_r\}$ *and along the* $\text{span}\{\widehat{Q}_l\}$.

Proof. The proof is a direct consequence of the fact that $\widehat{Q}_l^T \overline{\widehat{R}}_r = 0$ and $\widehat{Q}_r^T \overline{\widehat{R}}_l = 0$. □

2.4. Related Work. Ruhe [16] introduced the two-sided Arnoldi method specifically as a method for computing approximations to left eigenvectors of a matrix A. In [16] a one-sided Arnoldi method is applied to A to obtain converged approximations $\{\theta, x_r\}$ to an eigenvalue and right eigenvector of A. The quantities generated are used to compute an approximation, x_l, to a corresponding left eigenvector of A. A second application of the one-sided Arnoldi method is applied to $\{A^T, x_l\}$ to generate a better approximation $\tilde{x}_l$ to the left eigenvector. The two applications of the Arnoldi method produce different eigenvalue approximations. To obtain a consistent triplet for approximations to an eigenvalue and to the corresponding right and left eigenvector of A, additional computations are introduced which correspond to the merge operations.

The proposed two-sided Arnoldi recursions are generalizations of the algorithm in Ruhe [16]. The recursions in Algorithm 2.2 can handle starting blocks with any numbers of vectors. The two applications of a one-sided Arnoldi method are exercised independently. The merge computations can handle any numbers of left and right vectors, as long as the corresponding matrix $\overline{Q}$ has full rank, and the resulting $\overline{H}$ matrices can be used to define a variety of iterative methods including model reduction and eigenelement methods.

3. Two-Sided Block Arnoldi Model Reduction. We are interested in iterative methods for computing reduced-order models of large linear systems of time

invariant, differential-algebraic equations,

$$C\dot{x} = Gx + Bu, \qquad y = E^T x. \tag{3.1}$$

C and G are $n \times n$ matrices where n is the order of the system. The block input matrix B is $n \times q$ where q is the number of input variables. The block output matrix E^T is $o \times n$ where o is the number of output variables of the system. The behavior of such a system is encapsulated in the system transfer function function $\mathcal{T}(s)$ which maps the Laplace transform of the input functions u to the transform of the output functions y, [15].

$$y(s) = \widetilde{T}(s)u(s) \equiv E^T(sC - G)^{-1}Bu(s). \tag{3.2}$$

If $q = o = 1$, then the system is (si/so) and $\widetilde{T}(s)$ is a rational function of s. In general, a system is (mi/mo) and $\widetilde{T}(s)$ is an $o \times q$ matrix of rational functions. Each entry in $\widetilde{T}(s)$ is a si/so transfer function for one of the possible input/output combinations.

Typically C is not invertible, and the matrix $(Cs - G)^{-1}B$ is replaced by a matrix $(I + \sigma F)^{-1}R =$ where $F \equiv (Cs_0 - G)^{-1}C$, $R \equiv (Cs_0 - G)^{-1}B$, s_0 is some well-chosen expansion point and $\sigma = s - s_0$. An iterative model reduction method can then be applied to the system

$$F\dot{x} = -x + Ru, \quad y = E^T x, \tag{3.3}$$

defined by $\{F, R, E\}$ to obtain smaller systems $\{\overline{F}, \overline{R}, \overline{E}\}$. The original and the reduced systems have the same number of inputs and outputs. The approximation of the smaller system to the larger system is expressed as relationships between the transfer function of the original system and the transfer function of the smaller system. We have used R as the input block for a system and also as the residual matrix in a recursion. The reader should be able to deduce which use is intended from the local context.

We use Algorithm 2.2 to define a model reduction algorithm, Algorithm 3.1, for Equations(3.3). Formally, we can expand the transfer function $\widetilde{T}(s)$ of the original system in terms of the *moments*, $E^T F^j R$.

$$\widetilde{T}(s) \equiv E^T(I + \sigma F)^{-1}R = \sum_{j=0}^{\infty} (-1)^j E^T F^j R \sigma^j. \tag{3.4}$$

The performance of a model reduction procedure is typically measured by the number of moments of the transfer function of the reduced-order system which match the corresponding moments of the transfer function of the original system. See for example, [9, 10]. For some $0 \leq k \leq M$, the moments of the reduced-order system $\{\overline{F}, \overline{R}, \overline{E}\}$ satisfy

$$\overline{E}^T \overline{F}^k \overline{R} = E^T F^k R. \tag{3.5}$$

ALGORITHM 3.1. (*Two-Sided Block Arnoldi Model Reduction*)
Specify F, R, E, $\epsilon_d \geq 0$.
Apply Algorithm 2.2 to $\{F, R, E\}$ *for* m_r, m_l *steps to generate* $Q_r, Q_l, \overline{H}$.
Set $\overline{F} = \overline{H}, \qquad \overline{Q} = Q_l^T Q_r$.

If $\text{rank}\,(Q_l) \leq \text{rank}\,(Q_r)$
$\overline{R} = I_{[rfirst]} S_r, \quad \overline{E} = \overline{Q}^T I_{[lfirst]} S_l.$
else
$\overline{R} = \overline{Q} I_{[rfirst]} S_r, \quad \overline{E} = I_{[lfirst]} S_l.$
end

There is an implicit assumption in Algorithm 3.1 that the corresponding $\overline{Q}$ has full rank. The following theorem states that if the deflation tolerance $\epsilon_d = 0$, then the transfer function of a reduced order system, $\{\overline{F}, \overline{R}, \overline{E}\}$, obtained using Algorithm 3.1 achieves the maximum number of block moment matches to the transfer function of the original system $\{F, R, E\}$.

THEOREM 3.1. *Apply Algorithm 3.1 with* $\epsilon_d = 0$ *to the system* $\{F, R, E\}$ *to generate a reduced-order system* $\{\overline{F}, \overline{R}, \overline{E}\}$. *Then the first* $0 \leq k \leq m_l + m_r - 1$ *block moments of the reduced system match the corresponding block moments of the original system:*

$$\overline{E}^T \overline{F}^k \overline{R} = E^T F^k R \quad \textit{for} \quad 0 \leq k \leq m_l + m_r - 1. \tag{3.6}$$

The proof of Theorem 3.1 uses only the basic form of the recursions, for example, $AQ_r = Q_r \overline{H} + \overline{R}_r I_{[rlast]}^T$, the relationship between the modified residual matrix in one recursion and the vectors generated in the other, $Q_l^T \overline{R}_r = 0$, and the block upper Hessenberg shape of the projection matrices $\overline{H}$. The proof invokes Lemmas 3.2, 3.3, and 3.6.

Lemma 3.2 states that for small $\epsilon_d > 0$, deflation introduces correspondingly small perturbations in the one-sided Arnoldi recursions. This lemma is a direct consequence of the constructions in Algorithm 2.1. The proof of Equation(3.7) is by induction. Equation(3.8) is an immediate consequence of the fact that for any block upper Hessenberg matrix H with diagonal block sizes $d_1, \ldots, d_m$ that for any $0 \leq k < m - 1$, $\text{span}\{H_m^k I_{[first]}\}$ is contained in $\text{span}\{e_1, \ldots e_K\}$ where $K = \sum_{j=1}^{k+1} d_j$ and e_l denotes the l^{th} coordinate vector.

LEMMA 3.2. *After* m *steps of Algorithm 2.1,*

$$\begin{aligned} AQ_m &= Q_m H_m + R_m I_{[last]}^T + \Delta_m, \\ X &= Q_m I_{[first]} S + \widetilde{\Delta}, \end{aligned} \tag{3.7}$$

where $\|\widetilde{\Delta}\|_F \leq \sqrt{d_r - d_1}\epsilon_d$, *and* $\|\widetilde{\Delta}_m\|_F \leq \sqrt{d_1 - d_m}\epsilon_d$. *For* $0 \leq k < m - 1$.

$$I_{[last]}^T H_m^k I_{[first]} = 0. \tag{3.8}$$

Lemma 3.3 relates the action of powers of a matrix A as applied to a starting block X to powers of a corresponding reduced-order matrix H operating on the corresponding reduced starting block. The proof is by induction and uses Lemma 3.2. This relationship will be used to prove that block moments of the transfer functions of the reduced-order systems approximate block moments of the transfer function of the original system.

LEMMA 3.3. *Let* $\{Q_m, H_m, R_m, S\}$ *be generated by applying* m *steps of Algorithm 2.1 to* $\{A, X\}$. *For* $0 \leq k < m$,

$$\begin{aligned} &A^k X = Q_m H_m^k I_{[first]} S + (A^k \widetilde{\Delta} + \textstyle\sum_{\ell=0}^{k-1} A^\ell \Delta_m H_m^{k-1-\ell} I_{[first]} S), \quad \textit{and} \\ &A^m X = Q_m H_m^m I_{[first]} S + R_m I_{[last]}^T H_m^{m-1} I_{[first]} S + \\ &\qquad (A^m \widetilde{\Delta} + \textstyle\sum_{\ell=0}^{m-1} A^\ell \Delta_m H_m^{m-1-\ell} I_{[first]} S). \end{aligned} \tag{3.9}$$

COROLLARY 3.4. *Apply Algorithm 2.2 to* $\{A, X_r, A^T, X_l\}$. *Assume* $\operatorname{rank}(Q_r) \geq \operatorname{rank}(Q_l)$. *By construction* $\overline{H}$ *is block upper Hessenberg. If* $\epsilon_d = 0$, *then for* $0 \leq k < m_r$,

$$\begin{aligned} &A^k X_r = Q_r \overline{H}^k I_{[rfirst]} S_r \qquad \textit{and} \\ &A^{m_r} X_r = Q_r \overline{H}^{m_r} I_{[rfirst]} S + \overline{R}_r I_{[rlast]}^T \overline{H}^{m_r - 1} I_{[rfirst]} S_r. \end{aligned} \tag{3.10}$$

Corollary 3.4 follows directly from Lemma 3.3 and the fact that $\overline{H}$ retains the block upper Hessenberg form of H_r. We now consider the two-sided block Arnoldi recursion, Algorithm 2.2 with $\epsilon_d = 0$. $[\overline{H}]_r$ and $[\overline{H}]_l$ denote, respectively, $\overline{H}$ matrices which correspond to $Q_l^T \overline{R}_r = 0$ and $Q_r^T \overline{R}_l = 0$.

LEMMA 3.5. *Assume that* $Q_r, [\overline{H}]_r, \overline{R}_r, Q_l, H_l, R_l$ *satisfy*

$$\begin{aligned} &AQ_r = Q_r [\overline{H}]_r + \overline{R}_r I_{[rlast]}^T \\ &A^T Q_l = Q_l H_l + R_l I_{[llast]}^T \\ &Q_l^T \overline{R}_r = 0. \end{aligned}$$

For any Y_l *such that* $I_{[llast]}^T Y_l = 0$,

$$Y_l^T H_l^T Q_l^T Q_r = Y_l^T Q_l^T Q_r [\overline{H}]_r. \tag{3.11}$$

A similar statement is valid if the roles of the right and the left Arnoldi vectors are reversed. For any Y_r *such that* $I_{[rlast]}^T Y_r = 0$,

$$Y_r^T H_r^T Q_r^T Q_l = Y_r^T Q_r^T Q_l [\overline{H}]_l. \tag{3.12}$$

Proof. We prove Equation(3.11).

$$\begin{aligned} Y_l^T H_l^T Q_l^T Q_r &= Y_l^T (Q_l H_l)^T Q_r \\ &= Y_l^T (A^T Q_l - R_l I_{[llast]}^T)^T Q_r \\ &= Y_l^T (A^T Q_l)^T Q_r \\ &= Y_l^T Q_l^T (A Q_r) \\ &= Y_l^T Q_l^T Q_r [\overline{H}]_r. \end{aligned}$$

□

LEMMA 3.6. *Let the hypotheses of Lemma 3.5 corresponding to* $[\overline{H}]_r$ *be satisfied. Assume that for some* X_k *that*

$$A^k Q_r I_{[rfirst]} = Q_r [\overline{H}]_r^k I_{[rfirst]} + \overline{R}_r X_k \qquad (0 \leq k \leq m_r) \tag{3.13}$$

$$(A^T)^k Q_l I_{[lfirst]} = Q_l H_l^k I_{[lfirst]} \qquad (0 \leq k < m_l). \tag{3.14}$$

and that $I_{[lfirst]}^T (H_l^T)^k I_{[llast]} = 0$ *for* $0 \leq k < m_l - 1$. *Then for* $0 \leq k < m_l + m_r$,

$$(Q_l I_{[lfirst]})^T A^k (Q_r I_{[rfirst]}) = I_{[lfirst]}^T Q_l^T Q_r [\overline{H}]_r^k I_{[rfirst]}. \tag{3.15}$$

A similar statement is valid if the roles of the right and the left Arnoldi vectors are reversed. Assume that for some X_k that

$$(A^T)^k Q_l I_{[lfirst]} = Q_l([\overline{H}]_l^T)^k I_{[lfirst]} + \overline{R}_l X_k \qquad (0 \le k \le m_l) \tag{3.16}$$

$$A^k Q_r I_{[rfirst]} = Q_r H_r^k I_{[rfirst]} \qquad (0 \le k < m_r). \tag{3.17}$$

and that $I_{[rfirst]}^T (H_r^T)^k I_{[rlast]} = 0$ for $0 \le k < m_r - 1$, then for $0 \le k < m_l + m_r$,

$$(Q_l I_{[lfirst]})^T A^k (Q_r I_{[rfirst]}) = I_{[lfirst]}^T [\overline{H}]_l^k Q_l^T Q_r I_{[rfirst]}. \tag{3.18}$$

Proof. We prove Equation(3.15). By construction for $0 \le k < m_r$,

$$\begin{aligned}(Q_l I_{[lfirst]})^T A^k (Q_r I_{[rfirst]}) &= (Q_l I_{[lfirst]})^T (Q_r [\overline{H}]_r^k I_{[rfirst]} + \overline{R}_r X_k) \\ &= (Q_r^T Q_l I_{[lfirst]})^T [\overline{H}]_r^k I_{[rfirst]}.\end{aligned} \tag{3.19}$$

For $m_r \le k < m_l + m_r$, let $k = m_r + k_l$. Clearly, $0 \le k_l < m_l$ and

$$\begin{aligned}(Q_l I_{[lfirst]})^T A^k (Q_r I_{[rfirst]}) &= ((A^T)^{k_l} Q_l I_{[lfirst]})^T (A^{m_r} Q_r I_{[rfirst]}) \\ &= (Q_l H_l^{k_l} I_{[lfirst]})^T (Q_r [\overline{H}]_r^{m_r} I_{[rfirst]} + R_r X_{m_r}) \\ &= (Q_l H_l^{k_l} I_{[lfirst]})^T (Q_r [\overline{H}]_r^{m_r} I_{[rfirst]}) \\ &= I_{[lfirst]}^T (H_l^T)^{k_l} (Q_l^T Q_r) [\overline{H}]_r^{m_r} I_{[rfirst]}.\end{aligned}$$

For any $p < k_l$, define $Y_l^p \equiv H_l^p I_{[lfirst]}$. By assumption, $I_{[llast]}^T Y_l^p = 0$. If we apply Lemma 3.5 recursively for $p = 0, \ldots, k_l - 1$, we obtain $I_{[lfirst]}^T (H_l^T)^{k_l} (Q_l^T Q_r) = I_{[lfirst]}^T (Q_l^T Q_r) [\overline{H}]_r^{k_l}$. Therefore,

$$\begin{aligned}(Q_l I_{[lfirst]})^T A^k (Q_r I_{[rfirst]}) &= I_{[lfirst]}^T (H_l^T)^{k_l} (Q_l^T Q_r) [\overline{H}]_r^{m_r} I_{[rfirst]} \\ &= I_{[lfirst]}^T (Q_l^T Q_r) [\overline{H}]_r^{k_l} [\overline{H}]_r^{m_r} I_{[rfirst]} \\ &= (Q_r^T Q_l I_{[lfirst]})^T [\overline{H}]_r^k I_{[rfirst]}.\end{aligned}$$

□

Proof. (Theorem 3.1). We consider the case $\overline{H} = [\overline{H}]_r$. An analogous proof applies when $\overline{H} = [\overline{H}]_l$. By construction, $FQ_r = Q_r \overline{H} + \overline{R}_r I_{[rlast]}^T$, and $Q_l^T \overline{R}_r = 0$. Therefore, from Lemma 3.3,

$$\begin{aligned}F^k Q_r I_{[rfirst]} &= Q_r \overline{H}^k I_{[rfirst]} + \overline{R_r} X_k \qquad (0 \le k \le m_r). \\ (F^T)^\ell Q_l I_{[lfirst]} &= Q_l H_l^\ell I_{[lfirst]}, \qquad (0 \le \ell < m_l),\end{aligned}$$

for some consistent X_k. From Lemma 3.2, $I_{[lfirst]}^T (H_l^T)^k I_{[llast]} = 0$ for $0 \le k < m_l - 1$. From Lemma 3.6, for $0 \le k \le m_l + m_r - 1$,

$$(Q_l I_{[lfirst]})^T F^k (Q_r I_{[rfirst]}) = (Q_r^T Q_l I_{[lfirst]})^T \overline{H}^k I_{[rfirst]}. \tag{3.20}$$

But, $E = Q_l I_{[lfirst]} S_l$, $R = Q_r I_{[rfirst]} S_r$, $\overline{E} = Q_r^T Q_l I_{[lfirst]} S_l$, $\overline{R} = I_{[rfirst]} S_r$ and $\overline{F} = \overline{H}$. Therefore, for $0 \leq k \leq m_l + m_r - 1$,

$$E^T F^k R = (Q_l I_{[lfirst]} S_l)^T F^k (Q_r I_{[rfirst]} S_r) = (Q_r^T Q_l I_{[lfirst]} S_l)^T \overline{H}^k I_{[rfirst]} S_r = \overline{E}^T \overline{F}^k \overline{R}. \tag{3.21}$$

□

Reference [12] also uses a two-sided Arnoldi method to obtain approximations to transfer functions of control systems. The focus in [12] is on *si/so* systems, and the emphasis is on approximating Lyapunov functions [15]. Moment matching connections, as presented in this Section are not discussed. Connections with nonsymmetric Lanczos methods are mentioned but not developed. The statement is made that the results extend to block methods but deflation is not discussed and the infeasibility of a nonsymmetric block Lanczos recursion is not acknowledged.

4. Lanczos Recursions. Reference [4] focuses on relationships between nonsymmetric Lanczos and one-sided Arnoldi methods for solving $Ax = b$. In [4] it is proved that any residual norm behavior resulting from the application of the Lanczos-based biconjugate gradient method (BiCG) [18] to $Ax = b$ can be replicated by the application of the one-sided Arnoldi-based, full orthogonal method (FOM), but to a different problem $Cy = d$. The applications $\{BiCG, Ax = b\}$ and $\{FOM, Cy = d\}$ generate identical residual norms. Reference [3] focuses on relationships between corresponding eigenelement methods.

The two-sided nature of Algorithm 2.2 leads us to ask whether or not we can prove much stronger relationships between iterative methods which are based upon it and corresponding methods which are based upon the nonsymmetric band Lanczos recursion in [1]. In this section we explore that question. We prove that corresponding two-sided iterative methods generate identical iterates. Therefore, they are simply different implementations of the same iterative methods.

The nonsymmetric band Lanczos recursions in [1] generate sets of right vectors, V_s, and left vectors, W_s, which are biorthogonal. For each s, $W_s^T V_s = D_s$ with D_s a diagonal matrix. The vectors V_s and W_s are bases for corresponding right and left subspaces spanned by sets of Krylov vectors. Typically, as the recursion accumulates information about the original problem, global biorthogonality is lost [2].

Procedures based upon nonsymmetric Lanczos recursions have been used successfully in a variety of applications. See for example, [9, 6, 13]. However, the basic nonsymmetric Lanczos recursions may *breakdown*. If there is no mismatch in the left and the right starting Lanczos vectors [14], breakdown can be circumvented by invoking *look-ahead* ideas [14, 11]. Incorporating lookahead requires modifications in the basic Lanczos recursions.

Attempts have been made to construct nonsymmetric block Lanczos algorithms. However, it is now recognized that it is not feasible to construct nonsymmetric Lanczos recursions which are based upon explicit blocks of vectors. This difficulty is a consequence of the facts that the left and the right Lanczos vectors are biorthogonal (not independently orthogonal) so at each stage must be generated in pairs, and that as the recursions proceed, vectors within a w-block or a v-block can become dependent upon vectors generated earlier and must be deflated. Deflation does not, however, have to occur in (v, w) pairs. There can be deflations in the left(right) vector block without similar deflations in the right(left) block. If at some point in the recursions, the sizes of the left and of the right blocks are not equal, then the corresponding equations

which determine the biorthogonalization coefficients are overdetermined, and the recursions cannot be continued. This problem does not occur in the two-sided block Arnoldi recursion because the left and the right vectors are generated independently.

In the band nonsymmetric Lanczos recursion in [1], this problem is resolved by generating individual pairs of (v, w) vectors, one vector at a time by alternating back and forth between the generation of a $v-$vector and the generation of a $w-$*vector*. One complete iteration corresponds to the generation of one $v-$vector and one $w-$vector. Therefore, at the completion of each iteration there are equal numbers of left and right Lanczos vectors.

At each iteration possible candidates for the next $v-$vector are drawn from an *implicit block* of vectors. The first implicit $v-$block is $\widetilde{V}_1$ obtained from the starting block $X_v = \widetilde{V}_1 S_v + \widetilde{\Delta}_v$ and similarly, for $\widetilde{W}_1$ from X_w, where the S_v, S_w are constructed so that $\widetilde{V}_1$ and $\widetilde{W}_1$ are biorthogonal. The second implicit $v-$block consists of those $v-$vectors which were generated from the $v-$candidate vectors obtained by applying A to $\widetilde{V}_1$ and invoking appropriate Gram-Schmidt biorthogonalizations with respect to $w-$vectors. The j^{th} implicit block consists of those $v-$vectors which were generated from the $v-$candidate vectors which were obtained by applying A to $\tilde{V}_{j-1}$ and invoking biorthogonalizations. If deflation occurs during the construction of some implicit block $\tilde{V}_j$ then $\mathrm{rank}\,(\tilde{V}_j) < \mathrm{rank}\,(\tilde{V}_{j-1})$. If no suitable candidates are found for some such block then the recursions terminate. Analogous statements hold for the $w-$vector implicit blocks with A replaced by A^T.

We use $d_j^v (d_j^w)$ to denote the number of vectors in the current implicit $\tilde{V}_j, \tilde{W}_j$ blocks. Since deflation occurs independently in the right and the left vectors, d_j^v need not equal d_j^w. If the deflation tolerance, $\epsilon_d > 0$, and deflation occurs, then the recursions must be modified to include explicit biorthogonalization of each new $v-$vector ($w-$vector) with respect to the parents of deflated $w-$candidates($v-$candidates). For example, if a $v-$candidate vector which was generated from some Av_j is deflated, then the left recursions must be modified to include explicit biorthogonalization of each new $w-$vector with respect to the parent of this candidate, v_j. If $\epsilon_d = 0$, the equalities are unaffected by any deflation and no modifications are needed.

Thus, the nonsymmetric band Lanczos algorithm is analogous to a corresponding block algorithm where the vectors within a given block are constructed one by one and this one by one construction alternates between the construction of a $v-$vector and a $w-$vector. The alternation is required to maintain the feasibility of the biorthogonalization.

In our comparisons of Lanczos-based and Arnoldi-based methods, we will assume that the deflation is exact ($\epsilon_d = 0$), that no breakdown occurs, and that the ranks of the biorthogonal left and right starting blocks are equal. If these ranks differ, then the initial phase of the band Lanczos recursions has to be modified to generate enough right or left vectors to make the number of left and right vectors equal. See [1] for details.

The banded nonsymmetric Lanczos recursion has the following matrix form.

$$\begin{aligned} AV_s &= V_s T_v + R_v I_{[vlast]}^T \\ A^T W_s &= W_s T_w + R_w I_{([wlast])}^T \end{aligned} \tag{4.1}$$

vlast $= [s - d_v + 1 : s]$ and wlast $= [s - d_w + 1 : s]$ and $d_v (d_w)$ denote the number of columns in the final implicit $v-$block($w-$block). R_v and R_w are respectively, residual blocks of vectors with d_v and d_w columns. A merge operation is not necessary since the

block residuals generated satisfy $W_s^T R_v = 0$ and $V_s^T R_w = 0$. The Lanczos matrices T_v (T_w) are $s \times s$ banded matrices with maximum upper bandwidth of $d_1^w (d_1^v)$ and maximum lower bandwidth of $d_1^v (d_1^w)$. Typically, the bandwidths decrease as the iterations proceed.

We develop relationships between methods based upon the band nonsymmetric Lanczos recursion, and methods based upon the (truncated) two-sided block Arnoldi recursion, Algorithm 2.3.

EXAMPLE 4.1. *Apply the nonsymmetric Lanczos recursion to the triplet* $\{A, r, l\}$ *defined in Example 2.2. The first two* $\{v, w\}$ *Lanczos pairs are*

$$\begin{array}{ll} v_1 = r/\sqrt{78} & w_1 = l/\sqrt{3} \\ v_2 = [4, 3, -2, -1]^T/\sqrt{30} & w_2 = [0, 1, 2, -1]/\sqrt{6}. \end{array}$$

Since $w_2^T v_2 = 0$, *breakdown occurs at step* 2. *This coincides with the observed breakdown in Example 2.2 in the two-sided Arnoldi recursion.*

For the two-sided block Arnoldi recursion, breakdown did not result in modifications in the Arnoldi recursions. Those recursions were simply continued until the merge matrix $Q_l^T Q_r$ had full rank. For the Lanczos recursions, however, breakdown does necessitate modifications in the recursions. See [1]. Breakdown is a function of the starting blocks and the associated Krylov subspaces. In Lemma 4.3 we prove that if breakdown occurs in either the Lanczos or the two-sided block Arnoldi recursions, it must occur at corresponding points in these recursions.

Lemma 4.1 states that, until deflation occurs, the band nonsymmetric Lanczos recursion and the two-sided block Arnoldi recursion are generating bases for the same subspaces. Lemma 4.1 can be proved using mathematical induction with the fact that within each block of the one-sided Arnoldi recursions in the two-sided block Arnoldi recursion, candidate vectors are considered in order and one vector at a time. As defined by Algorithm 2.3, the Arnoldi recursions can be truncated at any intermediate vectors.

LEMMA 4.1. *For some* s, *apply the band nonsymmetric Lanczos recursion and the truncated two-sided block Arnoldi recursion, Algorithm 2.3, to* $\{A, X_r, X_l\}$ *to generate* $V_s, W_s, \widehat{Q}_l, \widehat{Q}_r$. *Assume no breakdown, no deflation, and exact arithmetic. Then* $\mathrm{span}(\widehat{Q}_r) = \mathrm{span}(V_s)$ *and* $\mathrm{span}(\widehat{Q}_l) = \mathrm{span}(W_s)$.

Deflation occurs only if a candidate vector is dependent upon previously-generated vectors. Since at each stage, each recursion is generating vectors which span the same subspaces, if the deflation is exact, $\epsilon_d = 0$, then any deflation must occur simultaneously in both recursions.

LEMMA 4.2. *Under the hypotheses of Lemma 4.1 allow exact deflation,* $\epsilon_d = 0$. *Assume no breakdown. If deflation of some right(left) candidate vector corresponding to some* $Av_i (A^T w_i)$ *occurs in the band Lanczos recursion, then the corresponding right(left) candidate vector corresponding to* $Aq_{ri} (A^T q_{li})$ *in the right(left) one-sided block Arnoldi recursion must also be deflated and vice-versa.*

Thus, with exact deflation, the corresponding subspaces generated using either the band nonsymmetric Lanczos recursion or the two-sided block Arnoldi recursion are identical. Therefore, breakdown, if it occurs, must occur simultaneously in both recursions.

LEMMA 4.3. *Under the hypotheses of Lemma 4.2, if breakdown occurs at step* $s+1$ *in the nonsymmetric band Lanczos recursion,* $w_{s+1}^T v_{s+1} = 0$, *and we extend the truncated two-sided block Arnoldi recursion to* $s+1$ *vectors, the corresponding* $\widehat{Q}_l^T \widehat{Q}_r$ *is singular. Similarly, if we extend the truncated two-sided block Arnoldi recursion to*

$s+1$ vectors and the corresponding $\widehat{Q}_l^T\widehat{Q}_r$ is singular, then extending the nonsymmetric band Lanczos recursion yields $w_{s+1}^T v_{s+1} = 0$.

Proof. If $w_{s+1}^T v_{s+1} = 0$, then the diagonal matrix $W_{s+1}^T V_{s+1}$ is singular and the Lanczos recursions cannot be continued. By Lemmas 4.1 and 4.2, the two recursions generate bases for the same subspaces. Therefore, there exist nonsingular matrices B and C such that $W_{s+1} = \widehat{Q}_l C$ and $V_{s+1} = \widehat{Q}_r B$, and $\widehat{Q}_l^T\widehat{Q}_r = C^{-T} W_{s+1}^T V_{s+1} B$ must be singular. The argument is easily reversed. □

We can use either recursion, nonsymmetric band Lanczos or the truncated two-sided block Arnoldi, to construct oblique projections of the matrix A. Lemma 4.4 relates the oblique projection matrices generated by these two recursions. The corresponding matrix recursions are Equations(4.1) and the Arnoldi recursions,

$$
\begin{aligned}
A\widehat{Q}_r &= \widehat{Q}_r[\overline{\widetilde{H}}]_r + \overline{\widehat{R}}_r I^T_{\widehat{[rlast]}} \\
A^T\widehat{Q}_l &= \widehat{Q}_l[\overline{\widetilde{H}}]_l + \overline{\widehat{R}}_l I^T_{\widehat{[llast]}}.
\end{aligned} \tag{4.2}
$$

$\widehat{[rlast]}$ and $\widehat{[llast]}$ denote the indices of the columns which contain the residual matrix corresponding to the truncated right and left block Arnoldi recursions. By construction, $\widehat{Q}_l^T\overline{\widehat{R}}_r = 0$, $\widehat{Q}_r^T\overline{\widehat{R}}_l = 0$, $W_s^T R_v = 0$, and $V_s^T R_w = 0$.

LEMMA 4.4. *Apply the nonsymmetric band Lanczos recursion and the truncated two-sided block Arnoldi recursion to $\{A, X_r, X_l\}$ using exact deflation. Define corresponding projection matrices T_v, T_w and $\{[\overline{\widehat{H}}]_r, [\overline{\widetilde{H}}]_l\}$ as defined in Equations(4.2). Then there exists nonsingular matrices B, C such that $T_v = B^{-1}[\overline{\widetilde{H}}]_r B$ and $T_w{}^T = C^T[\overline{\widetilde{H}}]_l C^{-T}$.*

Iterative methods based upon these recursions compute approximations to quantities associated with the original problem by solving reduced-order problems associated with these projection matrices. We use Lemma 4.4 to prove that eigenelement and model reduction methods defined using these recursions generate identical iterates. Therefore, they are different implementations of the same methods.

4.1. Computing Eigenvalues/Eigenvectors:. We will use $[\theta^L, z_r^L, z_l^L]$ and $[\theta^A, z_r^A, z_l^A]$ to denote approximations to eigenvalues, and to corresponding right and left eigenvectors of A generated by a Lanczos or an Arnoldi procedure. The two-sided block Arnoldi methods can be defined using either Algorithm 2.2 or Algorithm 2.3. In the comparisons we need to work with the same numbers of left and right vectors in the Lanczos and in the Arnoldi methods so we use Algorithm 2.3.

ALGORITHM 4.1. (*Two-sided Block Arnoldi Eigenelement Algorithm*)

Specify A, X_r, X_l, ϵ_d, s.

Apply Algorithm 2.3 to $\{A, X_r, X_l\}$.

Compute $[\overline{\widehat{H}}]_r u_r = \theta u_r$ and $[\overline{\widehat{H}}]_l^T u_l = \theta u_l$.

Compute $z_r \equiv \widehat{Q}_r u_r$, $z_l \equiv \widehat{Q}_l u_l$.

Compute error estimates $\epsilon_r = \overline{\widehat{R}}_r u_r(\widehat{[rlast]})$. and $\epsilon_l = \overline{\widehat{R}}_l u_l(\widehat{[rlast]})$

Lemma 4.4 tells us that the eigenvalues computed using $[\overline{\widetilde{H}}]_r$ and $[\overline{\widetilde{H}}]_l^T$ are identical. We define a corresponding band Lanczos eigenelement algorithm.

ALGORITHM 4.2. (*Two-sided Band Lanczos Eigenelement Algorithm*)

Specify A, X_r, X_l, ϵ_d, s.

Apply the nonsymmetric band Lanczos recursions to $\{A, X_r, X_l\}$.

Compute $T_v u_r = \theta u_r$ *and* $T_w u_l = \theta u_l$.
Compute $z_r \equiv V u_r$, $z_l \equiv W u_l$.
Compute unnormalized error estimates $\epsilon_r = R_v u_r([vlast])$, $\epsilon_l = R_w u_l([wlast])$.

4.2. Model Reduction. Similarly, we can define methods for model reduction of linear systems. See Equations(3.1). Algorithm 3.1 specifies a two sided block Arnoldi model reduction method. This definition maps directly onto a corresponding model reduction method which is based upon Algorithm 2.3.

ALGORITHM 4.3. (*Band Nonsymmetric Lanczos Model Reduction*)
Specify F, R, E, $\epsilon_d \geq 0$, s.
Apply the nonsymmetric band Lanczos recursions to generate V_s, W_s, T_v, T_w.
Set $\overline{F} = T_v$, $\quad \overline{R} = I_{[vfirst]} S_v^L$, $\quad \overline{E} = (V_s^T W_s) I_{[wfirst]} S_w^L$.

4.3. Equivalences between Two-Sided Arnoldi and Lanczos Methods.
THEOREM 4.5. *Set* $\epsilon_d = 0$. *Apply Algorithm 4.2 to* $\{A, X_r, X_l\}$ *to generate eigenelement approximations* $\{\theta_j^L, z_{rj}^L, z_{lj}^L\}$. *Apply Algorithm 4.1 to* $\{A, X_r, X_l\}$ *to generate eigenelement approximations* $\{\theta_j^A, z_{rj}^A, z_{lj}^A\}$. *Then the eigenvalue approximations and the corresponding left and right Ritz vectors generated by these two algorithms are identical.*

Proof. By Lemma 4.1 there exist nonsingular B, C such that $V_s = \widehat{Q}_r B$ and $W_s = \widehat{Q}_l C$. By Lemma 4.4, $T_v = B^{-1}[\widehat{H}]_r B$ and $T_w = C^{-1}[\widehat{\widetilde{H}}]_l^T C$. Furthermore, $T_w^T = (W_s^T V_s) T_v (W_s^T V_s)^{-1}$. Therefore, these two procedures generate identical eigenvalue approximations. Moreover, each $u_{rj}^A = B u_{rj}^L$. Therefore, $z_{rj}^A = \widehat{Q}_r u_{rj}^A = \widehat{Q}_r B u_{rj}^L = V_s u_{rj}^L = z_{rj}^L$, and similarly, for z_{lj}^A, z_{lj}^L. ∎

Theorem 4.6 states that corresponding Lanczos and Arnoldi model reduction algorithms generate identical approximations to the transfer function of the original system.

THEOREM 4.6. *Let* $\{F, R, E\}$ *be a mi/mo system defined by Equations(3.3). Apply the band nonsymmetric Lanczos model reduction procedure, Algorithm 4.3, to* $\{F, R, E\}$ *to obtain the reduced-order system* $\{\overline{F}^L, \overline{R}^L, \overline{E}^L\}$ *where* $\overline{F}^L \equiv \overline{T}^L \equiv T_v$. *Apply the truncated two-sided block Arnoldi model reduction procedure to* $\{F, R, E\}$ *to obtain the reduced-order system* $\{\overline{F}^A, \overline{R}^A, \overline{E}^A\}$ *of the same size where* $\overline{F}^A \equiv \overline{\widehat{H}}^A \equiv [\overline{\widehat{H}}]_r$. *The transfer functions of the Arnoldi and of the Lanczos reduced-order systems are equal to*

$$\widetilde{\mathcal{T}}^A(\sigma) \equiv [\overline{E}^A]^T (I + \sigma \overline{\widehat{H}}^A)^{-1} \overline{R}^A \tag{4.3}$$

$$\widetilde{\mathcal{T}}^L(\sigma) \equiv [\overline{E}^L]^T (I + \sigma \overline{T}^L)^{-1} \overline{R}^L. \tag{4.4}$$

For all complex σ,

$$\widetilde{\mathcal{T}}^A(\sigma) = \widetilde{\mathcal{T}}^L(\sigma). \tag{4.5}$$

Proof. By construction,

$$\widetilde{\mathcal{T}}^A(\sigma) \equiv (S_l^A)^T I_{[\widehat{lfirst}]}^T \overline{\widehat{Q}} (I + \sigma \overline{\widehat{H}}^A)^{-1} I_{[\widehat{rfirst}]} S_r^A. \tag{4.6}$$

From Lemmas 4.1 and 4.4, there exist nonsingular, upper triangular matrices B, C such that

$$V_s = \widehat{Q}_r B, \quad W_s = \widehat{Q}_l C, \quad \overline{\widehat{H}}^A = B\overline{T}^L B^{-1}. \tag{4.7}$$

Let $C_1 = I^T_{[\widehat{lfirst}]} C I_{[\widehat{lfirst}]}$ and $B_1 = I^T_{[\widehat{rfirst}]} B I_{[\widehat{rfirst}]}$ Since C and B are upper triangular, $C_1^{-1} = (C^{-1})_1$ and $B_1^{-1} = (B^{-1})_1$. Therefore, by Equations(4.7),

$$\begin{aligned} I^T_{[\widehat{lfirst}]} \overline{\widehat{Q}} &= (I^T_{[\widehat{lfirst}]} \widehat{Q}_l^T)\widehat{Q}_r = I^T_{[\widehat{lfirst}]} (W_s C^{-1})^T V_s B^{-1} = \\ &(W_s I_{[\widehat{lfirst}]} C_1^{-1})^T V_s B^{-1} = C_1^{-T} I^T_{[\widehat{lfirst}]} W_s^T V_s B^{-1}. \end{aligned} \tag{4.8}$$

By construction,

$$\begin{aligned} X_r &= \widehat{Q}_r I_{[\widehat{rfirst}]} S_r^A = V_s I_{[vfirst]} S_v^L \\ X_l &= \widehat{Q}_l I_{[\widehat{lfirst}]} S_l^A = W_s I_{[wfirst]} S_w^L \end{aligned} \tag{4.9}$$

Therefore,

$$\begin{aligned} S_r^A &= I_{[\widehat{rfirst}]} Q_r^T V_s I_{[vfirst]} S_v^L = I_{[\widehat{rfirst}]} B I_{[vfirst]} S_v^L = B_1 S_v^L \\ S_l^A &= I_{[\widehat{lfirst}]} Q_l^T W_s I_{[wfirst]} S_w^L = C_1 S_w^L \end{aligned} \tag{4.10}$$

Using Equations(4.8,4.10) in Equation(4.6), we obtain

$$\begin{aligned} \widetilde{\mathcal{T}}_A(\sigma) = (S_l^A)^T C_1^{-T} I^T_{[\widehat{lfirst}]} W_s^T V_s B^{-1} (I + \sigma B \overline{T}^L B^{-1})^{-1} I_{[\widehat{rfirst}]} S_r^A = \\ (S_w^L)^T I^T_{[\widehat{lfirst}]} W_s^T V_s (I + \sigma \overline{T}^L)^{-1} I_{[\widehat{rfirst}]} S_v^L = \widetilde{\mathcal{T}}^L(\sigma). \end{aligned} \tag{4.11}$$

□

5. Summary. We have proposed new two-sided block Arnoldi recursions which are extensions of the work in [16] for use in iterative methods. Iterative methods which are based upon these recursions have the advantage that any breakdown is centered in a vector merge matrix, and that breakdown can be handled without requiring modifications to the recursions. We used these two-sided block Arnoldi recursions to define a model reduction procedure which was proved to have maximum block moment matching properties. In comparisons of eigenelement and model reduction algorithms based upon these two-sided Arnoldi recursions and the band nonsymmetric Lanczos recursion in [1], we proved that the corresponding methods produce identical iterates, Therefore, they are different implementations of the same method.

6. Acknowledgements. The authors would like to thank the referees and the editor for their helpful comments.

REFERENCES

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

[2] Z. Bai, *Error analysis of the Lanczos algorithm for nonsymmetric eigenvalue problem*, Math. Comp., 62 (1994), pp. 209–226.

[3] J. Cullum, *Arnoldi versus nonsymmetric Lanczos algorithms for solving matrix eigenvalue problems*, BIT, 36 (1996), pp. 470–493.

[4] ———, *Iterative methods for solving $Ax = b$, GMRES/FOM versus QMR/BiCG*, Adv. Comp. Math., 6 (1996), pp. 1–24.

[5] J. CULLUM AND W. E. DONATH, *A block Lanczos algorithm for computing the q algebraically-largest eigenvalues and a corresponding eigenspace for large, sparse symmetric matrices*, in Proc. 1974 IEEE Conference on Decision and Control, IEEE Press, New York, 1974, pp. 505–509.

[6] J. CULLUM, W. KERNER, AND R. WILLOUGHBY, *A generalized nonsymmetric Lanczos procedure*, Comput. Phys. Comm., 53 (1989), pp. 19–48.

[7] J. CULLUM, A. RUEHLI, AND T. ZHANG, *A method for reduced-order modeling and simulation of large interconnect circuits and its application to PEEC models with retardation*, IEEE Trans. on Circ. and Syst. II, 47 (2000), pp. 261–273.

[8] J. CULLUM AND R. WILLOUGHBY, *Large Symmetric Eigenvalue Computations. Vol. 1, Theory, Vol. 2, Programs*, Birkhauser, Basel, Switzerland, 1985.

[9] P. FELDMAN AND R. W. FREUND, *Efficient linear circuit analysis by Padé approximation via the Lanczos process*, IEEE Trans. on Computer-Aided Design, 14 (1995), pp. 639–649.

[10] R. W. FREUND, *Computation of matrix Padé approximations of transfer functions via a Lanczos-type process*, in Approximation Theory VIII, Vol.1, Approximation and Interpolation, C. K. Chui and L. L. Schumaker, eds., World Scientific Publishing Co, 1995, pp. 215–222.

[11] R. W. FREUND, M. H. GUTKNECHT, AND N. M. NACHTIGAL, *An implementation of the look-ahead Lanczos algorithm for non-hermitian matrices*, SIAM J. Sci. Comp., 14 (1993), pp. 137–158.

[12] I. M. JAIMOUKHA AND E. M. KASENALLY, *Oblique projection methods for large scale model reduction*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 602–627.

[13] W. KERNER, *Large-scale complex eigenvalue problems*, J. Comp. Phys., 85 (1989), pp. 1–85.

[14] B. N. PARLETT, D. R. TAYLOR, AND Z. LIU, *A look-ahead Lanczos algorithm for unsymmetric matrices*, Math. Comp., 44 (1985), pp. 105–124.

[15] R. V. PATEL, A. J. LAUB, AND P. M. VAN DOOREN, *Numerical Linear Algebra Techniques for Systems and Control*, IEEE Press, Piscataway, N.J, 1994.

[16] A. RUHE, *The two-sided Arnoldi algorithm for nonsymmetric eigenvalue problems*, in Matrix Pencils, Lecture Notes in Mathematics, B. Kågström and A. Ruhe, eds., vol. 973, 1983, pp. 104–120.

[17] Y. SAAD, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, Manchester, UK, 1992.

[18] ———, *Iterative Methods for Sparse Linear System*, PWS Publishing Co., Boston, Massachusetts, 1996.