Chapter 10
Computational Algorithms and Software

This chapter presents, in details, the main algorithms for descriptor systems, which underlie the computational methods used in the synthesis procedures considered in this book. The core computations in these algorithms involve several matrix decompositions and condensed forms, which are obtainable using orthogonal transformations and, therefore, are provably numerical stable. Important applications of the condensed form are in developing numerically stable computational algorithms for the solution of several generalized matrix equations (Lyapunov, Stein, Sylvester, Riccati), which are frequently encountered in addressing the solution of synthesis problems in the fields of control and fault detection. The use of condensed forms, obtainable using orthogonal transformations (instead of using the potentially highly sensitive Weierstrass, Kronecker, or Brunovsky canonical forms), is also instrumental in developing numerically reliable procedures for the solution of several basic computational problems for descriptor systems as well as in some, rather specialized, algorithms for proper descriptor systems. Although this chapter is primarily intended for numerical experts having interests in control-related numerical techniques, it also serves to highlight the complexity of the underlying computations, which are necessary to address the synthesis problems of fault detection and isolation filters in a numerically sound way. A collection of software tools implements the algorithms presented in this chapter and can be employed to reproduce all computational results presented in this book.

10.1 Matrix Decompositions and Condensed Forms

The condensed forms of matrices play an important role in solving many control-related computational problems. A widely used computational paradigm in solving many computational problems consists of three main steps: (1) transform the original problem into a simpler one by reducing the problem data to condensed forms; (2)

A. Varga, *Solving Fault Diagnosis Problems*, Studies in Systems,
Decision and Control 84, DOI 10.1007/978-3-319-51559-5_10

solve the transformed problem using specially devised methods for the respective condensed forms; and (3) recover the solution of the original problem using back transformation to the original form. In this section we present several basic matrix decompositions, obtainable using orthogonal transformations, which involve several condensed forms of matrices, pairs of matrices, or even triples of matrices.

The use of orthogonal transformations is a widely accepted approach to promote numerical reliability of computations with finite precision. These transformations are perfectly conditioned with respect to inversion and, therefore, have the very desirable property that they do not amplify the existing uncertainties in the data. This feature is very important, since uncertainties in problem data are ubiquitous, representing inherent inaccuracies in data (e.g., truncation or discretization errors), or roundoff errors occurred in previous computational steps, or both. When using orthogonal transformations to transform problem data, it is often possible to bound the roundoff errors resulted as an effect of performed transformations on the data and even to show that the computed results are the exact solution of a problem with slightly perturbed data. Numerical algorithms exhibiting such a property are called (backward) numerically stable and underlie many algorithms for basic linear algebra computations. The use of numerically stable algorithms guarantees that the computed solution is accurate, provided the computational problem is well conditioned.

In what follows, we present several matrix decompositions involving particular condensed forms, which can be obtained using exclusively orthogonal transformations. These decompositions are the basis for many numerically stable algorithms employed by the synthesis procedures presented in this book. We will not address detailed algorithms for the computation of these forms, because they are described in details in several numerical linear algebra textbooks. However, we will indicate the associated computational complexity, by giving an estimation of the number of performed *floating-point computations* (*flops*) by a typical algorithm. For each decomposition we mention several straightforward applications, which often represent the building blocks of more complex numerical algorithms.

10.1.1 Singular Value Decomposition

The *singular value decomposition* (SVD) is a fundamental matrix factorization, which plays an important conceptual and computational role in linear algebra. The computation of the SVD can be interpreted as the reduction of a given rectangular matrix to a "diagonal" form using pre- and post-multiplications with orthogonal matrices. The main theoretical result regarding the SVD is the following theorem.

Theorem 10.1 *For any matrix $A \in \mathbb{R}^{m\times n}$, there exist orthogonal matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ such that*

$$A = U\Sigma V^T,$$

where $\Sigma = \mathrm{diag}(\Sigma_r, 0)$ *with* $\Sigma_r = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$ *and* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

The value of r defines obviously the *rank* of A. If we partition $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ and $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$ column-wise compatible with the row and column partitions of Σ, respectively, then

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 \Sigma_r V_1^T, \qquad (10.1)$$

which can be interpreted as a full rank factorization of A. We denote with $\sigma_i(A)$, $i = 1, \ldots, p$, the $p := \min(m, n)$ *singular values* of A, which are formed of the r nonzero singular values $\sigma_1, \ldots, \sigma_r$ together with $p - r$ zero singular values. The largest singular value $\overline{\sigma}(A) := \sigma_1$ is equal to $\|A\|_2$, the 2-norm of matrix A. For a square invertible matrix of order n, the 2-norm condition number with respect to inversion can be computed as $\kappa_2(A) := \|A\|_2 \|A^{-1}\|_2 = \sigma_1/\sigma_n$. The *Moore–Penrose pseudo-inverse* of A can be computed as $A^\dagger = V_1 \Sigma_r^{-1} U_1^T$. The minimum norm solution of the linear least-squares problem $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$ is simply $x = A^\dagger b = V_1 \Sigma_r^{-1} U_1^T b$.

Remark 10.1 The SVD is considered the primary tool to reliably determine the rank of a matrix. However, by applying any of the available numerically stable algorithms to compute the SVD, there will be almost always p nonzero singular values because of the incurred roundoff errors. If the original matrix A has the "mathematical rank" equal to r, then we can expect that $p - r$ of the numerically computed singular values to be "small." Thus, to determine the rank of A correctly, we need to choose a tolerance $\varepsilon > 0$ and define the "numerical rank" of A as r if the r-th and $r + 1$-th computed singular values satisfy

$$\sigma_r > \varepsilon \geq \sigma_{r+1} \,. \qquad (10.2)$$

Such a rank decision can be seen "reliable" if the gap $\sigma_r - \sigma_{r-1}$ is "large." It is important to note that the significance of the terms "small" and "large" is always in direct relation with the actual magnitudes of the matrix elements. The choice of the tolerance ε should be consistent with both the machine precision (i.e., $\varepsilon \geq \mathbf{u}\overline{\sigma}(A)$, where $\mathbf{u} = 2^{-52} \approx 2.22 \cdot 10^{-16}$ is the unit roundoff for the IEEE double precision floating-point representation), but also with the relative errors in the data (i.e., $\varepsilon \geq 10^{-k}\overline{\sigma}(A)$, where k is the number of correct decimal digits in the entries of A). We call the rank r determined such as (10.2) holds the *ε-rank* of A. □

A typical numerical algorithm for the computation of the full SVD (i.e., Σ, U and V) requires, for $m \geq n$, about $4m^2n + 8mn^2 + 9m^3$ *flops*, but only $4mn^2 - 4n^3/4$ *flops* for rank determination (i.e., computation of only Σ). For a properly implemented SVD algorithm, it can be shown that the computed diagonal matrix $\overline{\Sigma}$ is exact for a slightly perturbed A, in the following sense:

$$U^T(A + E)V = \overline{\Sigma},$$

where $U^T U = I$, $V^T V = I$, $\|E\|_2 = \mathcal{O}(\mathbf{u}\|A\|_2)$ and the computed $\overline{U}$ and $\overline{V}$ are almost orthogonal satisfying $\left\|U - \overline{U}\right\|_2 = \mathcal{O}(\mathbf{u})$ and $\left\|V - \overline{V}\right\|_2 = \mathcal{O}(\mathbf{u})$.

In the rest of this section we discuss some straightforward applications of the SVD. We assume the SVD of A has the partitioned form in (10.1), where r represents the ε-rank for a given tolerance ε satisfying (10.2) (i.e., all singular values of A satisfying $\sigma_i(A) \leq \varepsilon$ are considered equal to zero). The partitioned SVD (10.1) can be used to define orthogonal bases for the range and kernel of the matrix A as

$$\mathcal{R}(A) = \mathcal{R}(U_1), \qquad \mathcal{N}(A) = \mathcal{R}(V_2)\,,$$

as well as for its transpose A^T as

$$\mathcal{R}(A^T) = \mathcal{R}(V_1), \qquad \mathcal{N}(A^T) = \mathcal{R}(U_2)\,.$$

The orthogonal projections on the respective subspaces can be computed as

$$\begin{aligned} P_{\mathcal{R}(A)} &= U_1 U_1^T, \quad P_{\mathcal{N}(A)} = V_2 V_2^T, \\ P_{\mathcal{R}(A^T)} &= V_1 V_1^T, \quad P_{\mathcal{N}(A^T)} = U_2 U_2^T, \end{aligned}$$

where $P_{\mathcal{X}}$ denotes the orthogonal projection on a subspace $\mathcal{X}$.

Several row and column compressions can be easily obtained in terms of the elements of the SVD (10.1). Let Π_c and Π_r be permutation matrices defined as

$$\Pi_c = \begin{bmatrix} 0 & I_r \\ I_{n-r} & 0 \end{bmatrix}, \qquad \Pi_r = \begin{bmatrix} 0 & I_{m-r} \\ I_r & 0 \end{bmatrix}. \tag{10.3}$$

Then

$$U^T A = \begin{bmatrix} \Sigma_r V_1^T \\ 0 \end{bmatrix}, \qquad \Pi_r U^T A = \begin{bmatrix} 0 \\ \Sigma_r V_1^T \end{bmatrix},$$

represent two widely used row compressions of A to full row rank matrices via orthogonal transformations. Similarly,

$$AV = \begin{bmatrix} U_1 \Sigma_r & 0 \end{bmatrix}, \qquad AV\Pi_c = \begin{bmatrix} 0 & U_1 \Sigma_r \end{bmatrix}$$

are column compressions of A to full column rank matrices via orthogonal transformations.

10.1.2 QR Decomposition

The QR decomposition of a rectangular matrix in a product of an orthogonal matrix and an upper triangular matrix has many applications, which are similar to those of the SVD. Since the associated computational burden for the determination of the QR decomposition is significantly smaller than for the computation of the SVD, it is almost always advantageous to employ the QR decomposition instead the SVD, whenever this is possible. We cautiously remark that this gain of efficiency may sometime involve a certain loss of reliability in problems involving rank determinations. Fortunately, this may only occur for some rather "exotic" matrices and, therefore, QR factorization-based techniques are generally preferred to SVD-based methods in many control-oriented algorithms.

The main result on the QR decomposition is the following one.

Theorem 10.2 *For any matrix $A \in \mathbb{R}^{m\times n}$, there exists an orthogonal matrix $Q \in \mathbb{R}^{m\times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m\times n}$ such that*

$$A = QR\,.$$

Specifically, if $m > n$, then R has the form $R = \begin{bmatrix} R_{11} \\ 0 \end{bmatrix}$ with $m-n$ trailing zero rows, while if $n \geq m$ then $R = \begin{bmatrix} R_{11} & R_{12} \end{bmatrix}$. In both cases, R_{11} is a $p \times p$ upper triangular matrix, with $p = \min(m, n)$.

The QR decomposition (some authors prefer the term QR factorization) is the basic tool to solve the linear least-squares problem $\min_{x\in\mathbb{R}^n} \|Ax - b\|_2$, in the case when A is a full column rank matrix. The least-squares solution is simply $x = \begin{bmatrix} R_{11}^{-1} & 0 \end{bmatrix} Q^T b$. Furthermore, if R_{11} is chosen with positive diagonal elements, then R_{11} is the upper triangular factor of the *Cholesky factorization* of A^TA as $A^TA = R_{11}^T R_{11}$. Another application in the case $m > n$ is the computation of the SVD using bidiagonalization-based methods. These techniques can exploit the upper triangular shape of R_{11} to improve the overall computational efficiency.

We have a similar result for the so-called *RQ decomposition*, which is mainly relevant for the case $m \leq n$.

Theorem 10.3 *For any matrix $A \in \mathbb{R}^{m\times n}$, with $m \leq n$, there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ and an upper triangular matrix $R \in \mathbb{R}^{m\times m}$ such that*

$$A = \begin{bmatrix} 0 & R \end{bmatrix} Q\,.$$

If $r = \operatorname{rank} A < \min(m, p)$, the rank information cannot be usually read out from the resulting upper triangular factor R of the QR decomposition. An alternative rank-revealing factorization can be used which allows the determination of rank. The *QR factorization with column pivoting* has the form

$$A = Q\begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}\Pi =: Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix}\Pi, \tag{10.4}$$

where Q is orthogonal, $R_{11} \in \mathbb{R}^{r\times r}$ is upper triangular and invertible and Π is a permutation matrix. Obviously $r = \operatorname{rank} A$. The role of the column permutations is to enforce the invertibility of the leading block R_{11}. The term *column pivoting* indicates a specific column permutation strategy which tries to additionally enforce that R_{11} is well conditioned (with respect to inversion).

Remark 10.2 The rank determination using the QR factorization with column pivoting can be performed during the computation of this factorization. The factorization procedure iteratively constructs the upper triangular matrix R_{11} in the leading position. After $r = \operatorname{rank} A$ iterations, we have the partial decomposition

$$A = \widehat{Q}\begin{bmatrix} \widehat{R}_{11} & \widehat{R}_{12} \\ 0 & \widehat{R}_{22} \end{bmatrix}\widehat{\Pi},$$

where we expect that $\widehat{R}_{22}$ has a suitably small norm. A typical termination criterion might be

$$\|\widehat{R}_{22}\|_2 \le \varepsilon, \tag{10.5}$$

where $\varepsilon = \varepsilon_1\|A\|_2$ for some small parameter ε_1 depending on the machine roundoff unit $\mathbf{u}$ and the relative errors in the elements of A. If the above condition is fulfilled, then the matrix has "numerical rank" r (also called ε-rank). Surprisingly, there exist some artificially constructed examples (e.g., the Kahan matrices), for which the nearly rank deficiency cannot be detected in this way. Nevertheless, in practice, the QR factorization with column pivoting is almost as reliable as the SVD in determining matrix ranks. Therefore, it is widely used in many algorithms which involve repeated rank determinations (see, for example, the staircase algorithms in Sect. 10.3.1). Here, the repeated use of the full SVD would increase tremendously the computational complexity, due to the need to explicitly compute the involved orthogonal transformation matrices at each reduction step. □

A typical numerical algorithm for the computation of the QR factorization with column pivoting is based on the Householder QR factorization technique combined with column permutations, and requires about $4mnr - 2r^2(m+n) + 4r^3/3$ *flops*. Therefore, this algorithm is much more efficient than the algorithms for the computation of the SVD. Using the Householder reduction, the orthogonal transformation matrix Q is determined in a factored form $Q = H_1H_2\cdots H_r$, where H_i for $i = 1, \ldots, r$, are elementary orthogonal Householder transformation matrices (also known as Householder reflectors). Therefore, it is possible to avoid the explicit building of Q when computing products as Q^TB or CQ, where B and C are arbitrary matrices of compatible dimensions. For the Householder QR algorithm without pivoting, it can be shown that the computed $\overline{R}$ is exact for a nearby A in the sense

$$Q^T(A+E) = \overline{R},$$

where $Q^TQ = I$ and $\|E\|_2 = \mathcal{O}(\mathbf{u}\|A\|_2)$. The computed $\overline{Q}$ is almost orthogonal in the sense that $\left\|Q - \overline{Q}\right\|_2 = \mathcal{O}(\mathbf{u})$. A similar statement is obviously valid for the QR factorization with column pivoting.

In the rest of this section we discuss some straightforward applications of the QR decomposition, which parallel those of the SVD. We assume the QR decomposition with column pivoting of A has the partitioned form in (10.4), where r represents the ε-rank for a given tolerance ε satisfying (10.5) (i.e., the trailing $m - r$ rows of $Q^T A$ are considered equal to zero). Assume the orthogonal matrix Q in (10.4) is partitioned as $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$, where $Q_1 \in \mathbb{R}^{m\times r}$ and $Q_2 \in \mathbb{R}^{m\times(m-r)}$. We can determine orthogonal bases for the range of matrix A and the kernel of the matrix A^T (which is also its orthogonal complement) as

$$\mathcal{R}(A) = \mathcal{R}(Q_1), \qquad \mathcal{N}(A^T) = \mathcal{R}(A)^{\perp} = \mathcal{R}(Q_2)\,.$$

The orthogonal projections on these subspaces can be computed as $P_{\mathcal{R}(A)} = Q_1 Q_1^T$ and $P_{\mathcal{N}(A^T)} = Q_2 Q_2^T$, respectively. Obviously, orthogonal bases for $\mathcal{R}(A^T)$ and $\mathcal{N}(A)$ can be determined in terms of the QR decomposition with column pivoting of the transposed matrix A^T.

The row and column compressions can be obtained similarly as for the SVD. Let Π_r be the permutation matrix defined in (10.3). The row compressions of A to full row rank matrices, via orthogonal transformations, can be obtained in one of the following forms:

$$Q^T A = \begin{bmatrix} R_1 \Pi \\ 0 \end{bmatrix}, \qquad \Pi_r Q^T A = \begin{bmatrix} 0 \\ R_1 \Pi \end{bmatrix}.$$

Column compressions can be computed from the row compressions of the transposed matrix A^T, or, in the case of full row rank matrices, using directly the RQ decomposition (see Theorem 10.3).

10.1.3 Real Schur Decomposition

The *real Schur decomposition* (RSD) of a square real matrix A is a basic matrix decomposition which reveals the eigenvalues of A, by determining its *real Schur form* (RSF) (an upper quasi-triangular form) using orthogonal similarity transformations. The following theorem is the main theoretical result regarding the RSD.

Theorem 10.4 *For any $A \in \mathbb{R}^{n\times n}$ there exists an orthogonal $Q \in \mathbb{R}^{n\times n}$ such that $S = Q^T A Q$ is upper quasi-triangular of the form*

$$S = Q^T A Q = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1k} \\ 0 & S_{22} & \cdots & S_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_{kk} \end{bmatrix}, \qquad (10.6)$$

where each S_{ii} for $i = 1, \ldots, k$ is either a 1×1 or a 2×2 matrix having complex conjugate eigenvalues.

From the RSF (10.6), the eigenvalues of A result simply as

$$\Lambda(A) = \bigcup_{i=1}^{k} \Lambda(S_{ii}) \,.$$

The RSF also plays an important role in solving various linear matrix equations (Lyapunov, Stein, Sylvester), while the associated transformation matrix Q can be used to compute orthogonal bases of invariant subspaces (see below), which are useful in solving quadratic matrix Riccati equations.

An important property of the RSF is that the order of eigenvalues (and thus of the associated diagonal blocks) is arbitrary. The reordering of diagonal blocks (thus also of corresponding eigenvalues) can be simply done by interchanging two adjacent diagonal blocks of the RSF. For the swapping of such two blocks orthogonal similarity transformations can be used. Thus, any arbitrary reordering of blocks (and thus of the corresponding eigenvalues) can be achieved in this way. An important application of this fact is the computation of orthogonal bases for the invariant subspaces of A corresponding to a particular eigenvalue or a particular set of eigenvalues.

Consider a disjunct partition of the complex plane as $\mathbb{C} = \mathbb{C}_g \cup \mathbb{C}_b$, $\mathbb{C}_g \cap \mathbb{C}_b = \emptyset$, where $\mathbb{C}_g$ and $\mathbb{C}_b$ denote the "good" and "bad" regions of $\mathbb{C}$ for the location of eigenvalues of A, respectively. The ordered RSF is frequently employed in computational algorithms to exhibit a separation of eigenvalues into two sets, namely, all eigenvalues located in $\mathbb{C}_g$ gathered in the leading diagonal block of the RSF and all eigenvalues located in $\mathbb{C}_b$ gathered in the trailing diagonal block of the RSF. Overall we can achieve the orthogonal reduction of A to an ordered RSF matrix S in the form

$$S = Q^T A Q = \begin{bmatrix} A_g & A_{gb} \\ 0 & A_b \end{bmatrix},$$

where $\Lambda(A_g) \subset \mathbb{C}_g$ and $\Lambda(A_b) \subset \mathbb{C}_b$. If we partition Q as $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$ compatibly with the structure of the above S, then we can write

$$AQ_1 = Q_1 A_g \,.$$

It follows that

$$A\mathcal{R}(Q_1) \subset \mathcal{R}(Q_1)$$

and thus $\mathcal{R}(Q_1)$ is an *invariant subspace* corresponding to the eigenvalues of A lying in $\mathbb{C}_g$.

For the computation of the RSD the so-called Francis QR algorithm (or one of its modern variants) is usually used. This algorithm requires about $25n^3$ *flops* if both Q and S are computed. If the eigenvalue reordering is necessary, for example, to move p eigenvalues in the leading diagonal block of the RSF, then additionally at most $12n(n-p)p$ *flops* are necessary (e.g., $3n^3$ flops if $p = n/2$). If only the eigenvalues are desired, then $10n^3$ *flops* are necessary. The roundoff properties of the

QR algorithm are what one would expect of any orthogonal matrix technique. The computed RSF $\overline{S}$ is orthogonally similar to a matrix near to A, that is,

$$Q^T(A+E)Q = \overline{S},$$

where $Q^TQ = I$ and $\|E\|_2 = \mathcal{O}(\mathbf{u}\|A\|_2)$. The computed $\overline{Q}$ is almost orthogonal, in the sense that $\left\|I - \overline{Q}^T\overline{Q}\right\|_2 = \mathcal{O}(\mathbf{u})$. These relations are valid also in the case of employing eigenvalue reordering.

If A is a symmetric real matrix, then all eigenvalues of A are real and the symmetric real Schur form S is the diagonal form formed from the (real) eigenvalues. If additionally A is positive semi-definite, then all eigenvalues of A are non-negative and we have the following simple formula for the square root of A:

$$A^{\frac{1}{2}} = QS^{\frac{1}{2}}Q^T,$$

where $S^{\frac{1}{2}}$ is the diagonal matrix formed from the square roots of the eigenvalues. We can even compute the factor R of a Cholesky-like decomposition $A = R^TR$ as

$$R = S^{\frac{1}{2}}Q^T\,.$$

Such a factor is sometimes (improperly) called the square root of A.

10.1.4 Generalized Real Schur Decomposition

The eigenvalue structure of a regular pencil $A - \lambda E$ is completely described by the Weierstrass canonical form (see Lemma 9.8). However, the computation of this canonical form involves the use of (potentially ill-conditioned) general invertible transformations, and therefore numerical reliability cannot be guaranteed. Fortunately, the computation of Weierstrass canonical form can be avoided in almost all computations, and alternative "less"-condensed forms can be employed instead, which can be computed by employing exclusively orthogonal similarity transformations. The *generalized real Schur decomposition* (GRSD) of a matrix pair (A, E) reveals the eigenvalues of the regular pencil $A - \lambda E$, by determining the *generalized real Schur form* (GRSF) of the pair (A, E) (a quasi-triangular–triangular form) using orthogonal similarity transformations on the pencil $A - \lambda E$. The main theoretical result regarding the GRSD is the following theorem.

Theorem 10.5 *Let $A - \lambda E$ be an $n \times n$ regular pencil, with A and E real matrices. Then, there exist orthogonal transformation matrices Q and Z such that*

$$S - \lambda T := Q^T(A - \lambda E)Z = \begin{bmatrix} S_{11} & \cdots & S_{1k} \\ & \ddots & \vdots \\ 0 & & S_{kk} \end{bmatrix} - \lambda \begin{bmatrix} T_{11} & \cdots & T_{1k} \\ & \ddots & \vdots \\ 0 & & T_{kk} \end{bmatrix}, \qquad (10.7)$$

where each diagonal subpencil $S_{ii} - \lambda T_{ii}$, for $i = 1, \ldots, k$, is either of dimension 1×1 in the case of a finite real or infinite eigenvalue of the pencil $A - \lambda E$ or of dimension 2×2, with T_{ii} upper triangular, in the case of a pair of finite complex conjugate eigenvalues of $A - \lambda E$.

The pair (S, T) in (10.7) is in a GRSF and the eigenvalues of $A - \lambda E$ (or the generalized eigenvalues of the pair (A, E)) are given by

$$\Lambda(A - \lambda E) = \bigcup_{i=1}^{k} \Lambda(S_{ii} - \lambda T_{ii}) .$$

If $E = I$, then we can always choose $Q = Z$, $T = I$ and S is the RSF of A.

Similar to the RSF, the order of eigenvalues (and thus of the associated pairs of diagonal blocks) of the reduced pencil $S - \lambda T$ is arbitrary. The reordering of the pairs of diagonal blocks (thus also of corresponding eigenvalues) can be done by interchanging two adjacent pairs of diagonal blocks of the GRSF. For the swapping of such two pairs of blocks orthogonal similarity transformations can be used. Thus, any arbitrary reordering of pairs of blocks (and thus of the corresponding eigenvalues) can be achieved in this way. An important application of this fact is the computation of orthogonal bases for the deflating subspaces of the pencil $A - \lambda E$ corresponding to a particular eigenvalue or a particular set of eigenvalues.

For the computation of the GRSD the so-called QZ algorithm is usually used. This algorithm requires about $66n^3$ *flops* if all matrices S, T, Q and Z are computed. If the eigenvalue reordering is necessary, for example, to move p eigenvalues in the leading diagonal blocks of the GRSF, then additionally at most $24n(n-p)p$ *flops* are necessary (e.g., $6n^3$ flops if $p = n/2$). If only the eigenvalues are desired, then $30n^3$ *flops* are necessary. The roundoff properties of the QZ algorithm are what one would expect of any orthogonal matrix technique. The computed pair $(\overline{S}, \overline{T})$, in GRSF, is orthogonally similar to a matrix pair near to (A, E) and satisfies

$$Q^T(A + F)Z = \overline{S}, \qquad Q^T(E + G)Z = \overline{T} ,$$

where $Q^TQ = I$, $Z^TZ = I$, $\|F\|_2 = \mathcal{O}(\mathbf{u}\|A\|_2)$ and $\|G\|_2 = \mathcal{O}(\mathbf{u}\|E\|_2)$. The computed $\overline{Q}$ and $\overline{Z}$ are almost orthogonal, in the sense that $\left\|I - \overline{Q}^T\overline{Q}\right\|_2 = \mathcal{O}(\mathbf{u})$ and $\left\|I - \overline{Z}^T\overline{Z}\right\|_2 = \mathcal{O}(\mathbf{u})$. These relations are valid also in the case of employing eigenvalue reordering.

Consider a disjunct partition of the complex plane as $\mathbb{C} = \mathbb{C}_g \cup \mathbb{C}_b$, $\mathbb{C}_g \cap \mathbb{C}_b = \emptyset$, where $\mathbb{C}_g$ and $\mathbb{C}_b$ denote the "good" and "bad" regions of $\mathbb{C}$, respectively. We assume that $\mathbb{C}_g$, and therefore also $\mathbb{C}_b$, are symmetric with respect to the real axis. Then, it is possible to determine the orthogonal transformation matrices Q and Z such that

$$Q^T(A - \lambda E)Z = \begin{bmatrix} A_g - \lambda E_g & A_{gb} - \lambda E_{gb} \\ 0 & A_b - \lambda E_b \end{bmatrix} \tag{10.8}$$

is in a GRSF, where $\Lambda(A_g - \lambda E_g) \subset \mathbb{C}_g$ and $\Lambda(A_b - \lambda E_b) \subset \mathbb{C}_b$. Frequently used eigenvalue splittings are the stable–unstable splitting (i.e., $\mathbb{C}_g = \mathbb{C}_s$ and $\mathbb{C}_b = \mathbb{C} \setminus \mathbb{C}_s$) or the finite-infinite splitting (i.e., $\mathbb{C}_g = \mathbb{C} \setminus \{\infty\}$ and $\mathbb{C}_b = \{\infty\}$). More complicated splittings are possible by combining two or more partitions (see below).

The eigenvalue splitting achieved in the ordered GRSF (10.8) is the main tool for determining deflating subspaces corresponding to the eigenvalues of the pencil $A - \lambda E$. The subspaces $\mathcal{X}$ and $\mathcal{Y}$ form a *deflating pair* for the eigenvalues of $A - \lambda E$ if

$$\dim \mathcal{X} = \dim \mathcal{Y}$$

and

$$A\mathcal{X} \subset \mathcal{Y}, \qquad E\mathcal{X} \subset \mathcal{Y},$$

where $\dim \mathcal{S}$ denotes the dimension of the subspace $\mathcal{S}$. If we partition Q and Z compatibly with the structure of the GRSF (10.8) as $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$ and, respectively, $Z = \begin{bmatrix} Z_1 & Z_2 \end{bmatrix}$, then we can write

$$AZ_1 = Q_1 A_g, \qquad EZ_1 = Q_1 E_g \,.$$

It follows that $\dim \mathcal{R}(Q_1) = \dim \mathcal{R}(Z_1)$ and

$$A\mathcal{R}(Z_1) \subset \mathcal{R}(Q_1), \qquad E\mathcal{R}(Z_1) \subset \mathcal{R}(Q_1)\,.$$

Thus, $\mathcal{R}(Q_1)$ and $\mathcal{R}(Z_1)$ form a pair of (left and right) deflating subspaces associated to the eigenvalues of $A_g - \lambda E_g$. Deflating subspaces generalize the notion of invariant subspaces. If E is invertible, then the (right) deflating subspace $\mathcal{R}(Z_1)$ is an invariant subspace of $E^{-1}A$ corresponding to the eigenvalues of $E^{-1}A$ lying in $\mathbb{C}_g$. An important application of deflating subspaces is the solution of generalized Riccati equations, which can be equivalently formulated as the problem of determining orthogonal bases of the right deflating subspace corresponding to the stable eigenvalues of suitably defined regular pencils (see Sect. 10.2.2).

We describe now a special splitting of eigenvalues, which is instrumental for the computation of the proper and stable coprime factorizations using the methods described in Sect. 10.3.5. Assume $\mathbb{C}_g$ is finite region of $\mathbb{C}$, symmetric with respect to the real axis and $\mathbb{C}_b$ is its complement including also the point at infinity. The eigenvalue splitting in question involves the reduction of $A - \lambda E$ to the form

$$\widetilde{A} - \lambda \widetilde{E} = Q^T(A - \lambda E)Z = \begin{bmatrix} A_\infty & * & * \\ 0 & A_g - \lambda E_g & * \\ 0 & 0 & A_b - \lambda E_b \end{bmatrix}, \qquad (10.9)$$

where A_∞ is an $(n-r) \times (n-r)$ invertible (upper triangular) matrix, with $r = \operatorname{rank} E$, $\Lambda(A_g - \lambda E_g) \subset \mathbb{C}_g$ and $\Lambda(A_b - \lambda E_b) \subset \mathbb{C}_b$. The leading pair $(A_\infty, 0)$ contains all infinite eigenvalues of $A - \lambda E$ corresponding to first-order eigenvectors, while the rest of infinite eigenvalues are included in $A_b - \lambda E_b$.

The **Procedure GSORSF**, presented in what follows, computes the specially ordered GRSF in (10.9). The same procedure can be also used to obtain a reverse ordering of the diagonal blocks of $Q^T(A-\lambda E)Z$ in (10.9). For this, we apply the procedure to the transposed pencil $A^T-\lambda E^T$ to obtain Q_1 and Z_1 such that $Q_1^T(A^T-\lambda E^T)Z_1$ is in a form, as in the right side of (10.9). Let P be the permutation matrix

$$P=\begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}. \tag{10.10}$$

Then, with $Q=Z_1P$ and $Z=Q_1P$ we obtain $Q^T(A-\lambda E)Z$ in the form

$$Q^T(A-\lambda E)Z=\begin{bmatrix} A_b-\lambda E_b & * & * \\ 0 & A_g-\lambda E_g & * \\ 0 & 0 & A_\infty \end{bmatrix}. \tag{10.11}$$

Procedure GSORSF: Specially ordered generalized real Schur form

Inputs : $A-\lambda E$ regular, $\mathbb{C}_g$ and $\mathbb{C}_b$ such that $\mathbb{C}=\mathbb{C}_g\cup\mathbb{C}_b$, $\mathbb{C}_g\cap\mathbb{C}_b=\emptyset$
Outputs: $Q, Z, \widetilde{A}-\lambda\widetilde{E}=Q^T(A-\lambda E)Z$ in (10.9)

1) Compute an orthogonal Z_1 such that $EZ_1=\begin{bmatrix} 0 & E_2\end{bmatrix}$, with E_2 full column rank $r=\operatorname{rank} E$; compute the conformably partitioned $AZ_1=\begin{bmatrix} A_1 & A_2\end{bmatrix}$, with A_1 having full column rank $n-r$.
2) Compute an orthogonal Q_1 such that $Q_1^TA_1=\begin{bmatrix} A_\infty \\ 0\end{bmatrix}$, with A_∞ an $(n-r)\times(n-r)$ invertible upper triangular matrix; compute the conformably partitioned matrices

$$Q_1^TA_2=\begin{bmatrix} A_{12} \\ A_{22}\end{bmatrix},\quad Q_1^TE_2=\begin{bmatrix} E_{12} \\ E_{22}\end{bmatrix}.$$

3) Compute orthogonal Q_2 and Z_2 such that

$$Q_2^T(A_{22}-\lambda E_{22})Z=\begin{bmatrix} A_g-\lambda E_g & A_{gb}-\lambda E_{gb} \\ 0 & A_b-\lambda E_b\end{bmatrix}$$

is in a GRSF, where $\Lambda(A_g-\lambda E_g)\subset\mathbb{C}_g$ and $\Lambda(A_b-\lambda E_b)\subset\mathbb{C}_b$. Compute $A_{12}Q_2=\begin{bmatrix} A_{\infty,g} & A_{\infty,b}\end{bmatrix}$ and $E_{12}Q_2=\begin{bmatrix} E_{\infty,g} & E_{\infty,b}\end{bmatrix}$ conformably partitioned with $Q_2^T(A_{22}-\lambda E_{22})Z$.
4) Set $Q=Q_1\operatorname{diag}(I_{n-r},Q_2)$, $Z=Z_1\operatorname{diag}(I_{n-r},Z_2)$ and define $\widetilde{A}$ and $\widetilde{E}$ from the pencil

$$\widetilde{A}-\lambda\widetilde{E}=\begin{bmatrix} A_\infty & A_{\infty,g}-\lambda E_{\infty,g} & A_{\infty,b}-\lambda E_{\infty,b} \\ 0 & A_g-\lambda E_g & A_{gb}-\lambda E_{gb} \\ 0 & 0 & A_b-\lambda E_b\end{bmatrix}.$$

10.1.5 Controllability and Observability Staircase Forms

Staircase forms represent a large family of block upper triangular condensed forms, which arise from various algorithms which "compress" the numerical data available in single matrices or matrix pairs. All forms already studied, such as the diagonal form (originated from the SVD), upper triangular form (originated from the QR decomposition), the RSF (originated from the Francis QR algorithm) or the GRSF of a matrix pair (originated from the QZ algorithm), can be interpreted as particular staircase forms. For a general rectangular linear pencil, several Kronecker-like staircase forms (see next section) are obtainable using strict pencil similarity transformations using orthogonal transformations. In this section, we discuss two particular staircase forms, the controllability and observability staircase forms, which appear as parts of this form. However, due to their special importance for the computation of irreducible representation of descriptor systems, we dedicate a separate section for the discussion of their properties and also give a numerically stable computational procedure for their determination.

We have the following main result regarding the controllability staircase form.

Theorem 10.6 *Consider the pair $(A - \lambda E, B)$, with $A, E \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$, and assume the pencil $A - \lambda E$ is regular. Then, there exist orthogonal transformation matrices Q and Z such that*

$$\left[\widehat{B}\middle|\widehat{A} - \lambda\widehat{E}\right] := \left[Q^T B\middle|Q^T AZ - \lambda Q^T EZ\right] = \left[\begin{array}{c|cc} B_c & A_c - \lambda E_c & * \\ 0 & 0 & A_{\bar{c}} - \lambda E_{\bar{c}} \end{array}\right], \quad (10.12)$$

is in a generalized controllability staircase form with

$$\left[B_c\middle|A_c\right] = \left[\begin{array}{c|cccccc} A_{1,0} & A_{1,1} & A_{12} & \cdots & A_{1,k-1} & A_{1,k} \\ 0 & A_{2,1} & A_{22} & \cdots & A_{2,k-1} & A_{2,k} \\ 0 & 0 & A_{32} & \cdots & A_{3,k-1} & A_{3,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{k,k-1} & A_{k,k} \end{array}\right], \quad (10.13)$$

where $A_{j,j-1} \in \mathbb{R}^{\nu_j\times\nu_{j-1}}$, with $\nu_0 = m$, are full row rank matrices for $j = 1, \ldots, k$, and the resulting upper triangular matrix E_c has a similar block partitioned form

$$E_c = \begin{bmatrix} E_{1,1} & E_{1,2} & \cdots & E_{1,k-1} & E_{1,k} \\ 0 & E_{2,2} & \cdots & E_{2,k-1} & E_{2,k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & E_{k-1,k-1} & E_{k-1,k} \\ 0 & 0 & \cdots & 0 & E_{k,k} \end{bmatrix}, \quad (10.14)$$

where $E_{j,j} \in \mathbb{R}^{\nu_j \times \nu_j}$. *The resulting block dimensions* ν_j, $j = 0, 1, \ldots, k$, *satisfy*

$$m = \nu_0 \geq \nu_1 \geq \cdots \geq \nu_k > 0\,.$$

The $n_c \times (m + n_c)$ pencil $\left[\,B_c \middle| A_c - \lambda E_c\,\right]$, with $n_c := \sum_{j=1}^{k} \nu_j$, has full row rank for any finite $\lambda \in \mathbb{C}$, and therefore the pair $(A_c - \lambda E_c, B_c)$ is finite controllable. If $n_c < n$, then the $(n - n_c) \times (n - n_c)$ regular pencil $A_{\bar{c}} - \lambda E_{\bar{c}}$ contains the finite uncontrollable eigenvalues of $A - \lambda E$ (and also possibly some infinite ones).

If $m = 1$, then all subdiagonal blocks $A_{j,j-1}$ of A_c are 1×1 and A_c is in a *Hessenberg form*. The pair (A_c, E_c) with A_c in Hessenberg form and E_c upper triangular is in a so-called *generalized Hessenberg form* (GHF). If $m > 1$, then A_c is in a so-called *block Hessenberg form*. If $E = I$, then we can choose $Q = Z$ such that $\widehat{E} = I$.

Remark 10.3 If we partition Q and Z compatibly with the structure of the staircase form (10.12) as $Q = \left[\,Q_1\ Q_2\,\right]$ and, respectively, $Z = \left[\,Z_1\ Z_2\,\right]$, then we can write $AZ_1 = Q_1 A_c$ and $EZ_1 = Q_1 E_c$. It follows that $\dim \mathcal{R}(Q_1) = \dim \mathcal{R}(Z_1)$ and $A\mathcal{R}(Z_1) \subset \mathcal{R}(Q_1)$, $E\mathcal{R}(Z_1) \subset \mathcal{R}(Q_1)$. Thus, $\mathcal{R}(Q_1)$ and $\mathcal{R}(Z_1)$ form a pair of (left and right) deflating subspaces associated to the eigenvalues of $A_c - \lambda E_c$. Additionally we have

$$\mathcal{R}(B) \subset A\mathcal{R}(Z_1) + E\mathcal{R}(Z_1) \tag{10.15}$$

and $\mathcal{C}_f := \mathcal{R}(Z_1)$ is a deflating subspace with least possible dimension satisfying (10.15). We call $\mathcal{C}_f$ the *finite controllability subspace* of the pair $(A - \lambda E, B)$. The pair $(A - \lambda E, B)$ is *finite controllable* if the dimension of $\mathcal{C}_f$ is n. □

We also have the dual result to Theorem 10.6 for the observability staircase form.

Theorem 10.7 *Consider the pair* $(A - \lambda E, C)$, *with* $A, E \in \mathbb{R}^{n \times n}$ *and* $C \in \mathbb{R}^{p \times n}$, *and assume the pencil* $A - \lambda E$ *is regular. Then, there exist orthogonal transformation matrices* Q *and* Z *such that*

$$\left[\begin{array}{c} \widehat{A} - \lambda \widehat{E} \\ \hline \widehat{C} \end{array}\right] := \left[\begin{array}{c} Q^T AZ - \lambda Q^T EZ \\ \hline CZ \end{array}\right] = \left[\begin{array}{cc} A_{\bar{o}} - \lambda E_{\bar{o}} & * \\ 0 & A_o - \lambda E_o \\ \hline 0 & C_o \end{array}\right], \tag{10.16}$$

is in a generalized observability staircase form with

$$\left[\begin{array}{c} A_o \\ \hline C_o \end{array}\right] = \left[\begin{array}{ccccc} A_{\ell,\ell} & A_{\ell,\ell-1} & \cdots & A_{\ell,2} & A_{\ell,1} \\ A_{\ell-1,\ell} & A_{\ell-1,\ell-1} & \cdots & A_{\ell-1,2} & A_{\ell-1,1} \\ 0 & A_{\ell-2,\ell-1} & \cdots & A_{\ell-2,2} & A_{\ell-2,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{1,2} & A_{1,1} \\ \hline 0 & 0 & \cdots & 0 & A_{0,1} \end{array}\right], \tag{10.17}$$

where $A_{j-1,j} \in \mathbb{R}^{\mu_{j-1}\times\mu_j}$, with $\mu_0 = p$, are full column rank matrices for $j = 1, \ldots, \ell$, and the resulting upper triangular matrix E_o has a similar block partitioned form

$$E_o = \begin{bmatrix} E_{\ell,\ell} & E_{\ell,\ell-1} & \cdots & E_{\ell,2} & E_{\ell,1} \\ 0 & E_{\ell-1,\ell-1} & \cdots & E_{\ell-1,2} & E_{\ell-1,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & E_{2,2} & E_{2,1} \\ 0 & 0 & \cdots & 0 & E_{1,1} \end{bmatrix}, \qquad (10.18)$$

with $E_{j,j} \in \mathbb{R}^{\mu_j\times\mu_j}$. The resulting block dimensions μ_j, $j = 0, 1, \ldots, \ell$, satisfy

$$p = \mu_0 \geq \mu_1 \cdots \geq \mu_\ell > 0\,.$$

The $(n_o + p) \times n_o$ pencil $\begin{bmatrix} A_o - \lambda E_o \\ C_o \end{bmatrix}$, with $n_o := \sum_{j=1}^{\ell} \mu_j$, has full column rank for any finite $\lambda \in \mathbb{C}$, and therefore the pair $(A_o - \lambda E_o, C_o)$ is finite observable. If $n_o < n$, then the $(n - n_o) \times (n - n_o)$ regular pencil $A_{\bar{o}} - \lambda E_{\bar{o}}$ contains the finite unobservable eigenvalues of $A - \lambda E$ (and also possibly some infinite ones).

Remark 10.4 If we partition Q and Z compatibly with the structure of the staircase form (10.16) as $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$ and, respectively, $Z = \begin{bmatrix} Z_1 & Z_2 \end{bmatrix}$, then we can write $AZ_1 = Q_1 A_{\bar{o}}$ and $EZ_1 = Q_1 E_{\bar{o}}$. It follows that $\dim \mathcal{R}(Q_1) = \dim \mathcal{R}(Z_1)$ and $A\mathcal{R}(Z_1) \subset \mathcal{R}(Q_1)$, $E\mathcal{R}(Z_1) \subset \mathcal{R}(Q_1)$. Thus, $\mathcal{R}(Q_1)$ and $\mathcal{R}(Z_1)$ form a pair of (left and right) deflating subspaces associated to the eigenvalues of $A_{\bar{o}} - \lambda E_{\bar{o}}$. Additionally, $\overline{\mathcal{O}}_f := \mathcal{R}(Z_1)$ is a deflating subspace with the largest dimension satisfying $\mathcal{R}(Z_1) \subset \mathcal{N}(C)$. We call $\overline{\mathcal{O}}_f$ the *finite unobservable subspace* of the pair $(A - \lambda E, C)$. The pair $(A - \lambda E, C)$ is *finite observable* if the dimension of $\overline{\mathcal{O}}_f$ is zero. □

The following procedure to compute the staircase form (10.12) can be seen as a constructive proof of Theorem 10.6. In view of the main application of this procedure (see Sect. 10.3.1), we included a matrix $C \in \mathbb{R}^{p\times n}$ on which all transformations to the right are also applied.

Procedure GCSF: Generalized controllability staircase form

Inputs : $(A - \lambda E, B, C)$
Outputs: $Q, Z, (A-\lambda E, B, C) := (Q^T AZ - \lambda Q^T EZ, Q^T B, CZ)$; $\nu_j, j = 1, \dots, \ell$

1) Compute an orthogonal matrix Z such that EZ is upper triangular; compute $A \leftarrow AZ$, $E \leftarrow EZ$, $C \leftarrow CZ$.
2) Set $j = 1$, $n_c = 0$, $\nu_0 = m$, $A^{(0)} = A$, $E^{(0)} = E$, $B^{(0)} = B$, $Q = I_n$.
3) Compute orthogonal matrices W and U such that

$$W^T B^{(j-1)} := \begin{array}{c} \left[\begin{array}{c} A_{j,j-1} \\ 0 \end{array}\right] \begin{array}{l} \nu_j \\ \rho \end{array} \\ \nu_{j-1} \end{array},$$

with $A_{j,j-1}$ full row rank and $W^T E^{(j-1)} U$ upper triangular.

4) Compute and partition

$$W^T A^{(j-1)} U := \begin{array}{c} \left[\begin{array}{cc} A_{j,j} & A_{j,j+1} \\ B^{(j)} & A^{(j)} \end{array}\right] \begin{array}{l} \nu_j \\ \rho \end{array} \\ \begin{array}{cc} \nu_j & \rho \end{array} \end{array}, \quad W^T E^{(j-1)} U := \begin{array}{c} \left[\begin{array}{cc} E_{j,j} & E_{j,j+1} \\ O & E^{(j)} \end{array}\right] \begin{array}{l} \nu_j \\ \rho \end{array} \\ \begin{array}{cc} \nu_j & \rho \end{array} \end{array}$$

5) For $i = 1, \dots, j-1$ compute and partition

$$A_{i,j} U := \begin{array}{c} [\, A_{i,j} \;\; A_{i,j+1} \,] \\ \begin{array}{cc} \nu_j & \rho \end{array} \end{array}, \quad E_{i,j} U := \begin{array}{c} [\, E_{i,j} \;\; E_{i,j+1} \,] \\ \begin{array}{cc} \nu_j & \rho \end{array} \end{array}$$

6) $Q \leftarrow Q \operatorname{diag}(I_{n_c}, W)$, $Z \leftarrow Z \operatorname{diag}(I_{n_c}, U)$, $C \leftarrow C \operatorname{diag}(I_{n_c}, U)$.
7) $n_c \leftarrow n_c + \nu_j$; if $\rho = 0$ then $\ell = j$ and **Exit**.
8) If $\nu_j > 0$, then $j \leftarrow j + 1$ and go to Step 3); else, $\ell = j - 1$, and **Exit**.

If the **Procedure GCSF** exits at Step 7), then the original pair $(A - \lambda E, B)$ is finite controllable. However, if the **Procedure GCSF** exits at Step 8), then the original pair $(A - \lambda E, B)$ is not finite controllable. In this case, the trailing $\rho \times \rho$ pencil $A^{(\ell+1)} - \lambda E^{(\ell+1)} =: A_{\bar{c}} - \lambda E_{\bar{c}}$, with $\rho = n - n_c$, contains all uncontrollable finite eigenvalues of $A - \lambda E$.

The **Procedure GCSF** can be implemented such that at Step 1) it exploits any particular shape in the lower triangle of E (e.g., E lower banded). In particular, if E is upper triangular, then the resulting Z is simply $Z = I$ and no further computations are performed at this step. The row compressions at Step 3) are usually performed using rank-revealing QR factorizations with column pivoting (see Sect. 10.1.2). The reductions can be performed using sequences of Givens rotations (instead Householder reflectors), which allow to simultaneously perform the column transformations accumulated in U to maintain the upper triangular form of $E^{(j-1)}$. This reduction technique is described in detail in [125] and is similar to the reduction of a matrix pair to a generalized Hessenberg form. Using this technique, the numerical complexity of **Procedure GCSF** is $\mathcal{O}(n^3)$ (for $m, p \ll n$), provided all transformations are immediately applied without accumulating explicitly W and U. Note that the usage of the more robust rank determinations based on singular values decompositions would increase the overall complexity to $\mathcal{O}(n^4)$ due to the need to accumulate explicitly each W and U. Regarding the numerical properties of **Procedure GCSF**, it is possible to show that the resulting system matrices $\widehat{A}$, $\widehat{E}$, $\widehat{B}$, $\widehat{C}$ are exact for slightly perturbed original data A, E, B, C, while Q and Z are nearly orthogonal matrices. It follows that the **Procedure GCSF** is numerically stable.

To compute the observability staircase form of a pair $(A - \lambda E, C)$, the **Procedure GCSF** can be applied to the dual pair $(A^T - \lambda E^T, C^T)$ to obtain the transformed pair $(\widehat{A}^T - \lambda \widehat{E}^T, \widehat{C}^T)$ in a controllability staircase form. Then, the pair $(P\widehat{A}P - P\widehat{E}P, \widehat{C}P)$, where P is the permutation matrix (10.10), is in an observability staircase form.

10.1.6 Kronecker-Like Forms

Consider the reduction of a general rectangular (or singular) pencil $M - \lambda N$, with $M, N \in \mathbb{R}^{m\times n}$ using strict similarity transformations of the form

$$\widehat{M} - \lambda \widehat{N} = U(M - \lambda N)V,$$

where U and V are invertible matrices. From Lemma 9.9, recall that, using general invertible transformations, we can determine the Kronecker-canonical form (9.44) of the pencil $M - \lambda N$, which basically characterizes the right and left singular structure and the eigenvalue structure of the pencil. The computation of the Kronecker-canonical form may involve the use of ill-conditioned transformations and, therefore, is potentially numerically unstable. Fortunately, alternative staircase forms, called *Kronecker-like forms*, allow to obtain basically the same (or only a part of) structural information on the pencil $M - \lambda N$ by employing exclusively orthogonal transformations (i.e., $U^T U = I$ and $V^T V = I$).

The following result concerns with one of the main Kronecker-like forms.

Theorem 10.8 *Let $M \in \mathbb{R}^{m\times n}$ and $N \in \mathbb{R}^{m\times n}$ be arbitrary real matrices. Then, there exist orthogonal $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$, such that*

$$U(M - \lambda N)V = \begin{bmatrix} M_r - \lambda N_r & * & * \\ 0 & M_{reg} - \lambda N_{reg} & * \\ 0 & 0 & M_l - \lambda N_l \end{bmatrix}, \qquad (10.19)$$

where

(1) The $n_r \times (m_r + n_r)$ pencil $M_r - \lambda N_r$ has full row rank, n_r, for all $\lambda \in \mathbb{C}$ and is in a controllability staircase form

$$M_r - \lambda N_r = \begin{bmatrix} B_r & A_r - \lambda E_r \end{bmatrix}, \qquad (10.20)$$

with $B_r \in \mathbb{R}^{n_r\times m_r}$, $A_r, E_r \in \mathbb{R}^{n_r\times n_r}$, and E_r invertible.

(2) The $n_{reg} \times n_{reg}$ pencil $M_{reg} - \lambda N_{reg}$ is regular and its eigenvalues are the eigenvalues of pencil $M - \lambda N$. The pencil $M_{reg} - \lambda N_{reg}$ may be chosen in a GRSF, with arbitrary-ordered diagonal blocks.

(3) The $(p_l + n_l) \times n_l$ pencil $M_l - \lambda N_l$ has full column rank, n_l, for all $\lambda \in \mathbb{C}$ and is in a observability staircase form

$$M_l - \lambda N_l = \begin{bmatrix} A_l - \lambda E_l \\ C_l \end{bmatrix}, \tag{10.21}$$

with $C_l \in \mathbb{R}^{p_l \times n_l}$, $A_l, E_l \in \mathbb{R}^{n_l \times n_l}$, and E_l invertible.

Let ν_i, $i = 1, \ldots, k$ be the dimensions of the diagonal blocks of $A_r - \lambda E_r$ in the controllability staircase form $\begin{bmatrix} B_r & A_r - \lambda E_r \end{bmatrix}$ and define $\nu_0 = m_r$. These dimensions completely determine the right Kronecker structure of $M - \lambda N$ as follows: there are $\nu_{i-1} - \nu_i$ blocks $L_{i-1}(\lambda)$ of size $(i-1) \times i$, $i = 1, \ldots, k$. Analogously, let μ_i, $i = 1, \ldots, \ell$ be the dimensions of the diagonal blocks of $A_l - \lambda E_l$ in the observability staircase form $\begin{bmatrix} A_l - \lambda E_l \\ C_l \end{bmatrix}$ and define $\mu_0 = p_l$. These dimensions completely determine the left Kronecker structure of $M - \lambda N$ as follows: there are $\mu_{i-1} - \mu_i$ blocks $L^T_{i-1}(\lambda)$ of size $i \times (i-1)$, $i = 1, \ldots, \ell$. We have $n_r = \sum_{i=1}^{k} \nu_i$ and $n_l = \sum_{i=1}^{\ell} \mu_i$, and the normal rank of $M - \lambda N$ is $n_r + n_{reg} + n_l$. The finite Smith zeros of $M - \lambda N$ are the finite eigenvalues of the regular pencil $M_{reg} - \lambda N_{reg}$ and represent the finite values of λ for which $M - \lambda N$ drops its rank below its normal rank.

In Sect. 10.3 several applications of the Kronecker-like forms are presented, such as the computation of minimal nullspace basis, system zeros, inner–outer factorizations and the solution of linear rational equations.

For the computation of the Kronecker-like form (10.19) the standard approach is to achieve successive separations of the structural elements and eigenvalues of the pencil $M - \lambda N$. A typical basic pencil reduction procedure, as **Procedure PREDUCE** presented in this section, uses two orthogonal transformation matrices Q and Z to achieve the following separation:

$$\widetilde{M} - \lambda \widetilde{N} := Q(M - \lambda N)Z = \left[\begin{array}{c|c} M_{r,\infty} - \lambda N_{r,\infty} & * \\ \hline 0 & M_{f,l} - \lambda N_{f,l} \end{array}\right], \tag{10.22}$$

where the $m_{r,\infty} \times n_{r,\infty}$ pencil $M_{r,\infty} - \lambda N_{r,\infty}$ has full row rank for all $\lambda \in \mathbb{C}$ excepting possibly a finite set of infinite values of λ, and the $m_{f,l} \times n_{f,l}$ pencil $M_{f,l} - \lambda N_{f,l}$ has full column rank for all $\lambda \in \mathbb{C}$ excepting possibly a finite set of finite values of λ. Moreover, the pencil $\widetilde{M} - \lambda \widetilde{N}$ is in the following staircase form:

$$\widetilde{M} - \lambda\widetilde{N} = \left[\begin{array}{cccc|c} M_{1,1} & M_{1,2} - \lambda N_{1,2} & \cdots & M_{1,k} - \lambda N_{1,k} & * \\ 0 & M_{2,2} & \cdots & M_{2,k} - \lambda N_{2,k} & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & M_{k,k} & * \\ \hline 0 & 0 & \cdots & 0 & M_{f,l} - \lambda N_{f,l} \end{array}\right] \begin{array}{l} \} \nu_1 \\ \} \nu_2 \\ \\ \} \nu_k \\ \} m_{f,l} \end{array}, \qquad (10.23)$$

$$\underbrace{\quad}_{\mu_1} \quad \underbrace{\qquad}_{\mu_2} \quad \cdots \quad \underbrace{\qquad}_{\mu_k} \quad \underbrace{\qquad}_{n_{f,l}}$$

where $M_{i,i} \in \mathbb{R}^{\nu_i \times \mu_i}$ are full row rank matrices for $i = 1, \ldots, k$, and $N_{i-1,i} \in \mathbb{R}^{\nu_{i-1} \times \mu_i}$ are full column rank matrices for $i = 2, \ldots, k$. The dimensions ν_i and μ_i of the blocks in (10.23) satisfy

$$\mu_1 \geq \nu_1 \geq \mu_2 \geq \nu_2 \geq \cdots \mu_k \geq \nu_k \geq 0$$

and completely determine the right Kronecker structure and the eigenstructure at infinity of the pencil $M - \lambda N$ as follows: there are $\mu_i - \nu_i$ $(i = 1, \ldots, k)$ blocks $L_{i-1}(\lambda)$ of size $(i-1) \times i$ and, with $\mu_{k+1} := 0$, there are $\nu_i - \mu_{i+1}$ $(i = 1, \ldots, k)$ nilpotent Jordan blocks $J_i(0)$ of dimension i (see (9.43)) which correspond to the infinite eigenvalues. The row and column dimensions of pencil $M_{r,\infty} - \lambda N_{r,\infty}$ are given by $m_{r,\infty} = \sum_{i=1}^{k} \nu_i$ and $n_{r,\infty} = \sum_{i=1}^{k} \mu_i$, respectively.

Procedure PREDUCE: Pencil reduction to staircase form

Inputs : $M - \lambda N$ with $M, N \in \mathbb{R}^{m \times n}$
Outputs: $Q, Z, \widetilde{M} - \lambda\widetilde{N} = Q(M - \lambda N)Z$ in the staircase form (10.23).

0) Compute orthogonal matrices Q and Z such that

$$\widetilde{M} = QMZ = \begin{bmatrix} B^{(0)} & A^{(0)} \\ D^{(0)} & C^{(0)} \end{bmatrix}, \qquad \widetilde{N} = QNZ = \begin{bmatrix} 0 & E^{(0)} \\ 0 & 0 \end{bmatrix},$$

where $E^{(0)} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ is upper triangular and invertible with $\tilde{n} = \operatorname{rank} N$, $A^{(0)} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$, $B^{(0)} \in \mathbb{R}^{\tilde{n} \times \widetilde{m}}$ with $\widetilde{m} = n - \tilde{n}$, $C^{(0)} \in \mathbb{R}^{\tilde{p} \times \tilde{n}}$ with $\tilde{p} = m - \tilde{n}$, $D \in \mathbb{R}^{\tilde{p} \times \widetilde{m}}$; set $m_{r,\infty} = 0$, $n_{r,\infty} = 0$, $i = 0$.

Step–i: **while** $\widetilde{m} > 0$, **do**:

1) Compute orthogonal matrices W and Y such that $WD^{(i)}Y = \left[\begin{array}{c:c} D_1^{(i)} & D_2^{(i)} \\ \hdashline 0 & 0 \end{array}\right]$,

where $D_1^{(i)} \in \mathbb{R}^{\tau_i \times \tau_i}$ is invertible and upper triangular. Obtain

$$\left[\begin{array}{c:c|c} B_1^{(i)} & B_2^{(i)} & A^{(i)} - \lambda E^{(i)} \\ \hline D_1^{(i)} & D_2^{(i)} & C_1^{(i)} \\ \hdashline 0 & 0 & C_2^{(i)} \end{array}\right] := \operatorname{diag}(I_{\tilde{n}}, W) \left[\begin{array}{c|c} B^{(i)} & A^{(i)} - \lambda E^{(i)} \\ \hline D^{(i)} & C^{(i)} \end{array}\right] \operatorname{diag}(Y, I_{\tilde{n}}) .$$

Procedure PREDUCE: Pencil reduction to staircase form (continued)

2) Compress the rows of $\left[\begin{array}{c|c} B_1^{(i)} & E^{(i)} \\ D_1^{(i)} & 0 \end{array}\right]$ with orthogonal X such that

$$X\left[\begin{array}{c|c} B_1^{(i)} & E^{(i)} \\ D_1^{(i)} & 0 \end{array}\right] = \left[\begin{array}{c|c} B_{11}^{(i)} & E_1^{(i)} \\ 0 & E_2^{(i)} \end{array}\right],$$

with $B_{11}^{(i)} \in \mathbb{R}^{\tau_i \times \tau_i}$ and $E_2^{(i)} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ invertible and upper triangular. Obtain

$$\left[\begin{array}{c:c|c} B_{11}^{(i)} & B_{12}^{(i)} & A_1^{(i)} - \lambda E_1^{(i)} \\ 0 & B_{22}^{(i)} & A_2^{(i)} - \lambda E_2^{(i)} \\ \hdashline 0 & 0 & C_2^{(i)} \end{array}\right] := \operatorname{diag}(X, I_{\tilde{p}-\tau_i}) \left[\begin{array}{c:c|c} B_1^{(i)} & B_2^{(i)} & A^{(i)} - \lambda E^{(i)} \\ \hline D_1^{(i)} & D_2^{(i)} & C_1^{(i)} \\ \hdashline 0 & 0 & C_2^{(i)} \end{array}\right].$$

3) Compress the rows of $B_{22}^{(i)}$ with orthogonal U such that $UB_{22}^{(i)} = \begin{bmatrix} \widetilde{B}_{22}^{(i)} \\ 0 \end{bmatrix}$, with $\widetilde{B}_{22}^{(i)} \in \mathbb{R}^{\rho_i \times (\widetilde{m}-\tau_i)}$ full row rank, and compute orthogonal V such that $UE_2^{(i)}V$ is upper triangular. Obtain

$$\left[\begin{array}{c:c|c:c} B_{11}^{(i)} & B_{12}^{(i)} & A_{11}^{(i)} - \lambda E_{11}^{(i)} & * \\ \hdashline 0 & \widetilde{B}_{22}^{(i)} & A_{21}^{(i)} - \lambda E_{21}^{(i)} & * \\ \hline 0 & 0 & A_{31}^{(i)} & A_{32}^{(i)} - \lambda E_{32}^{(i)} \\ \hdashline 0 & 0 & C_{21}^{(i)} & C_{22}^{(i)} \end{array}\right] :=$$

$$\operatorname{diag}\left(I_{\tau_i}, U, I_{\tilde{p}-\tau_i}\right) \left[\begin{array}{c:c|c} B_{11}^{(i)} & B_{12}^{(i)} & A_1^{(i)} - \lambda E_1^{(i)} \\ 0 & B_{22}^{(i)} & A_2^{(i)} - \lambda E_2^{(i)} \\ \hdashline 0 & 0 & C_2^{(i)} \end{array}\right] \operatorname{diag}(I_{\widetilde{m}}, V),$$

with $E_{21}^{(i)} \in \mathbb{R}^{\rho_i \times \rho_i}$ and $E_{32}^{(i)} \in \mathbb{R}^{(\tilde{n}-\rho_i)\times(\tilde{n}-\rho_i)}$ invertible and upper triangular.

4) Form $Q^{(i)} = \operatorname{diag}\left(I_{m_{r,\infty}}, \widetilde{Q}\right)$ and $Z^{(i)} = \operatorname{diag}\left(I_{n_{r,\infty}}, \widetilde{Z}\right)$ with

$$\widetilde{Q} = \operatorname{diag}\left(I_{\tau_i}, U, I_{\tilde{p}-\tau_i}\right) \operatorname{diag}\left(X, I_{\tilde{p}-\tau_i}\right) \operatorname{diag}\left(I_{\tilde{n}}, W\right), \qquad \widetilde{Z} = \operatorname{diag}(Y, V),$$

and update $\widetilde{M} \leftarrow Q^{(i)}\widetilde{M}Z^{(i)}$, $\widetilde{N} \leftarrow Q^{(i)}\widetilde{N}Z^{(i)}$, $Q \leftarrow Q^{(i)}Q$, $Z \leftarrow ZZ^{(i)}$
Set $\nu_{i+1} = \rho_i + \tau_i$, $\mu_{i+1} = \widetilde{m}$ and define

$$M_{i+1,i+1} := \left[\begin{array}{c:c} B_{11}^{(i)} & B_{12}^{(i)} \\ \hdashline 0 & \widetilde{B}_{22}^{(i)} \end{array}\right], \quad M_{i+1,i+2} - \lambda N_{i+1,i+2} := \left[\begin{array}{c} A_{11}^{(i)} - \lambda E_{11}^{(i)} \\ \hdashline A_{21}^{(i)} - \lambda E_{21}^{(i)} \end{array}\right],$$

with $M_{i+1,i+1} \in \mathbb{R}^{\nu_{i+1} \times \mu_{i+1}}$ full row rank and $N_{i+1,i+2} \in \mathbb{R}^{\nu_{i+1} \times \rho_i}$ full column rank, and

$$\left[\begin{array}{c:c} B^{(i+1)} & A^{(i+1)} - \lambda E^{(i+1)} \\ \hdashline D^{(i+1)} & C^{(i+1)} \end{array}\right] := \left[\begin{array}{c:c} A_{31}^{(i)} & A_{32}^{(i)} - \lambda E_{32}^{(i)} \\ \hdashline C_{21}^{(i)} & C_{22}^{(i)} \end{array}\right].$$

5) Update $m_{r,\infty} \leftarrow m_{r,\infty} + \nu_{i+1}$, $n_{r,\infty} \leftarrow n_{r,\infty} + \mu_{i+1}$, $\tilde{n} \leftarrow \tilde{n} - \rho_i$, $\widetilde{m} \leftarrow \rho_i$, $\tilde{p} \leftarrow \tilde{p} - \tau_i$.
6) $i \leftarrow i + 1$ and *go to* **Step–i**

At the end of **Procedure PREDUCE** we obtain the $m_{f,l} \times n_{f,l}$ pencil

$$M_{f,l} - \lambda N_{f,l} := \begin{bmatrix} A^{(i)} - \lambda E^{(i)} \\ \hdashline C^{(i)} \end{bmatrix}, \tag{10.24}$$

with $m_{f,l} = \tilde{n} + \tilde{p}$ and $n_{f,l} = \tilde{n}$, and with $E^{(i)}$ upper triangular and invertible. It follows that the pencil $M_{f,l} - \lambda N_{f,l}$ has only finite and left structure. The number of diagonal blocks $M_{j,j}$ of $M_{r,\infty} - \lambda N_{r,\infty}$ in the staircase form (10.23) is $k = i - 1$, where i is the resulting final value of i at the exit of **Procedure PREDUCE**.

The **Procedure PREDUCE** performs exclusively orthogonal transformations on the matrix pair (M, N). It is possible to show that the resulting pair $(\widetilde{M}, \widetilde{N})$ is exact for a slightly perturbed original pair, while Q and Z are nearly orthogonal matrices. It follows that the **Procedure PREDUCE** is numerically stable.

The computational complexity of **Procedure PREDUCE** mainly depends on the details of the computations performed at Step 2) to obtain $E_2^{(i)}$ in an upper triangular form and, at Step 3), to preserve the upper triangular form of $UE_2^{(i)}V$ and to obtain $E_{21}^{(i)}$ and $E_{32}^{(i)}$ invertible and upper triangular. If the transformation matrices U and V are accumulated (e.g., by performing SVD-based row compressions), the worst-case computational complexity of **Procedure PREDUCE** is $\mathcal{O}(n^4)$ (assuming $n \geq m$), which, for large values of n, is unacceptable. However, using the techniques described in [95], these operations can be performed such that a worst-case computational complexity of $\mathcal{O}(n^3)$ can be guaranteed. The main computational ingredients are specially tailored QR decompositions with column pivoting, which provide almost the same reliability as the rank determinations based on the use of SVD. Using specialized QR decompositions, it is possible to implement the row compressions at Steps 2) and 3) such that the preservation of the upper triangular shape of $E^{(i)}$ is simultaneously possible, without the need to explicitly accumulate the intervening transformations. For the rest of necessary row and column compressions at Step 0) and Step 1), the safer SVD-based computations can be still employed, without increasing excessively the computational complexity.

A straightforward application of the **Procedure PREDUCE** is to perform the infinite–finite separation of the eigenvalues of a regular pencil $M - \lambda N$ (i.e., without right and left structures). Since $M - \lambda N$ has no right structure, $M_{r,\infty} - \lambda N_{r,\infty}$ has only infinite eigenvalues. Similarly, since $M - \lambda N$ has no left structure, $M_{f,l} - \lambda N_{f,l}$ contains all finite eigenvalues of the pencil.

A complementary separation of the pencil $M - \lambda N$ can be achieved by applying **Procedure PREDUCE** to the transposed pencil $M^T - \lambda N^T$ and pertranspose the resulted pencil. Recall that the pertranspose M^P of a matrix $M \in \mathbb{R}^{m\times n}$ is defined as $M^P := P_n M^T P_m$, where P_k denotes the $k \times k$ permutation matrix of the form (10.10). The net effect of applying P_n from left is to reverse the order of rows of a matrix, while the application of P_m from right reverses the order of columns of the matrix. If Q and Z are the orthogonal matrices used to reduce $M^T - \lambda N^T$, then overall we obtain

$$P_m Z^T (M - \lambda N) Q^T P_n = \begin{bmatrix} M_{r,f} - \lambda N_{r,f} & * \\ 0 & M_{\infty,l} - \lambda N_{\infty,l} \end{bmatrix},$$

where $M_{r,f} - \lambda N_{r,f}$ contains the right and finite structure and $M_{\infty,l} - \lambda N_{\infty,l}$ contains the infinite and left structure. Moreover, $M_{\infty,l} - \lambda N_{\infty,l}$ is in a dual staircase form, which is obtained by reversing the orders of the blocks in the staircase form (10.23). Sometimes, it is more advantageous to apply **Procedure PREDUCE** to $M^P - \lambda N^P$ instead of $M^T - \lambda N^T$ (e.g., already existing upper block structures are preserved by pertransposition and thus can be further exploited).

For the computation of the complete Kronecker-like form (10.19) of the pencil $M - \lambda N$ we can employ **Procedure PREDUCE** to perform the first separation in (10.22). Then, by applying **Procedure PREDUCE** to the pertransposed pencil $M^P_{r,\infty} - \lambda N^P_{r,\infty}$, we obtain the separation of the right and infinite structures in the form

$$Q_1 (M_{r,\infty} - \lambda N_{r,\infty}) Z_1 = \begin{bmatrix} M_r - \lambda N_r & * \\ 0 & M_\infty - \lambda N_\infty \end{bmatrix}, \qquad (10.25)$$

where Q_1 and Z_1 are orthogonal matrices, the full row rank pencil $M_r - \lambda N_r$ is in the form (10.20) and the regular pencil $M_\infty - \lambda N_\infty$, with M_∞ invertible and N_∞ nilpotent, contains the infinite eigenvalues. Similarly, by applying **Procedure PREDUCE** to the pertransposed pencil $M^P_{f,l} - \lambda N^P_{f,l}$, we obtain the separation of the finite and left structures in the form

$$Q_2 (M_{f,l} - \lambda N_{f,l}) Z_2 = \begin{bmatrix} M_f - \lambda N_f & * \\ 0 & M_l - \lambda N_l \end{bmatrix}, \qquad (10.26)$$

where Q_2 and Z_2 are orthogonal matrices, the regular pencil $M_f - \lambda N_f$ with N_f invertible contains the finite eigenvalues and the full column rank pencil $M_l - \lambda N_l$ is in an observability staircase form (10.21). Overall we achieved

$$\mathrm{diag}(Q_1, Q_2) Q (M - \lambda N) Z \, \mathrm{diag}(Z_1, Z_2) = \begin{bmatrix} M_r - \lambda N_r & * & * & * \\ 0 & M_\infty - \lambda N_\infty & * & * \\ 0 & 0 & M_f - \lambda N_f & * \\ 0 & 0 & 0 & M_l - \lambda N_l \end{bmatrix}$$

from which the regular part $M_{reg} - \lambda N_{reg}$ in (10.19) can be immediately read out. For this separation, it is possible to exploit the structure of the pencil $M_{f,l} - \lambda N_{f,l}$ in (10.24) which results when applying **Procedure PREDUCE**. Since in the pertransposed pencil $M^P_{f,l} - \lambda N^P_{f,l} = \left[\, (C^{(i)})^P \;\; (A^{(i)})^P - \lambda (E^{(i)})^P \, \right]$, the invertible matrix $(E^{(i)})^P$ is already upper triangular, therefore when applying **Procedure PREDUCE** to $M^P_{f,l} - \lambda N^P_{f,l}$ the preliminary reduction at Step 0) is not necessary anymore. Alternatively, the **Procedure GCSF** can be employed to obtain $M_l - \lambda N_l$ in an observability staircase form (10.21). This computation is needed to be additionally performed, to obtain $M_r - \lambda N_r$ in a controllability staircase form (10.20).

The above Kronecker-like form exhibits the main structural elements of an arbitrary pencil $M - \lambda N$. However, in some applications, as the computation of rational left nullspace bases in Sect. 7.4, it is necessary only to know the left Kronecker structure. For this purpose, it is sufficient to apply **Procedure PREDUCE** twice, to obtain the basic separation (10.22) and then the splitting of finite and left structures as in (10.26) to obtain the required form

$$\mathrm{diag}(I, Q_2)Q(M - \lambda N)Z\,\mathrm{diag}(I, Z_2) = \begin{bmatrix} M_{r,\infty} - \lambda N_{r,\infty} & * & * \\ 0 & M_f - \lambda N_f & * \\ 0 & 0 & M_l - \lambda N_l \end{bmatrix}.$$

On the other hand, when all structural details of the Kronecker-like form are necessary, as for example, when solving linear rational equations in Sect. 9.2.9, the separation of right and infinite structure of the pencil $M_{r,\infty} - \lambda N_{r,\infty}$ is necessary. An alternative way to perform this separation is to employ a computational approach proposed in [9] (see Algorithms 3.3.1 and 3.3.2). These algorithms exploit all structural information in the staircase form (10.23) and perform the separation of right and infinite structure by employing exclusively orthogonal transformations, however without making any rank decisions. The resulting subpencils $M_r - \lambda N_r$ and $M_\infty - \lambda N_\infty$ are in staircase forms and the dimensions of the resulting diagonal blocks automatically reveal the right Kronecker indices and infinite eigenvalue structure.

10.2 Solution of Matrix Equations

There are several linear and quadratic matrix equations which play an important role in control theory. In this section, we discuss the computational solutions of some of the main equations and give the conditions for the existence of a solution.

10.2.1 Linear Matrix Equations

We discuss the computational solution of two main classes of liner matrix equations. In the first class, we consider the *generalized Sylvester equation* (GSE) of the form

$$AXG + EXF + Q = 0\,, \tag{10.27}$$

where $A, E \in \mathbb{R}^{n\times n}$, $F, G \in \mathbb{R}^{m\times m}$, $Q \in \mathbb{R}^{n\times m}$, and the desired solution is $X \in \mathbb{R}^{n\times m}$. The Eq. (10.27) has a unique solution if and only if the matrix pencils $A - \lambda E$ and $F - \lambda G$ are regular and $\Lambda(A - \lambda E) \bigcap \Lambda(F + \lambda G) = \emptyset$.

Two special cases of Eq. (10.27) are of particular interest in this book: the *generalized continuous-time Lyapunov equation* (GCLE) of the form

$$AXE^T + EXA^T + Q = 0 \tag{10.28}$$

and the *generalized discrete-time Lyapunov equation* (GDLE) (also called generalized Stein equation)

$$AXA^T - EXE^T + Q = 0\,, \tag{10.29}$$

where Q, and hence also X, are symmetric. The solvability condition of Eq. (10.28) requires that E is invertible and $\lambda_i + \lambda_j \neq 0$, for all $\lambda_i, \lambda_j \in \Lambda(A - \lambda E)$. The solvability condition of Eq. (10.29) requires that $\lambda_i \lambda_j \neq 1$, for all $\lambda_i, \lambda_j \in \Lambda(A - \lambda E)$. In both cases, of special interest are (semi-)positive definite solutions in the case when Q has the form $Q = BB^T \geq 0$ and $\Lambda(A - \lambda E) \in \mathbb{C}_s$. In this case, the solution X can be directly obtained in a Cholesky-factored form $X = SS^T$, with S upper triangular.

For the numerical solution of the above matrix equations the transformation method (developed initially by Bartels and Stewart to solve the Sylvester equation $AX + XB + C = 0$) can be used. Let Q_1 and Z_1 be orthogonal matrices such that the pair $(P, S) := (Q_1^T AZ_1, Q_1^T EZ_1)$ is in a GRSF, and let Q_2 and Z_2 be orthogonal matrices such that the pair $(T, R) := (Q_2^T FZ_2, Q_2^T GZ_2)$ is in a GRSF. The matrices Q_1 and Z_1, and, Q_2 and Z_2, can be obtained by applying the QZ algorithm to the matrix pairs (A, E) and (F, G), respectively. If we define $Y = Z_1^T XQ_2$ and $H = Q_1^T QZ_2$, then the Eq. (10.27) can be rewritten as

$$PYR + SYT + H = 0\,.$$

By exploiting the upper quasi-triangular–upper triangular structures of the pairs (P, S) and (T, R), this equation can be solved by a special (back substitution) technique to obtain the solution Y [47, 54]. Then, the solution of (10.27) is computed as $X = Z_1 YQ_2^T$. The overall computational effort to solve Eq. (10.27) is $\mathcal{O}(n^3 + m^3) + \mathcal{O}(n^2 m + nm^2)$. With obvious simplifications, this approach can be used to solve the GCLE (10.28) and the GDLE (10.29) as well. The overall computational effort to solve these equations is $\mathcal{O}(n^3)$.

The second class of linear equation is the *generalized Sylvester system of equations* (GSSE)

$$\begin{aligned} AX + YF &= C, \\ EX + YG &= D, \end{aligned} \tag{10.30}$$

where $A, E \in \mathbb{R}^{n\times n}$, $F, G \in \mathbb{R}^{m\times m}$, $C, D \in \mathbb{R}^{n\times m}$, and the desired solution is $X, Y \in \mathbb{R}^{n\times m}$. The Eq. (10.30) has a unique solution if and only if the matrix pencils $A - \lambda E$ and $F - \lambda G$ are regular and $\Lambda(A - \lambda E) \bigcap \Lambda(F - \lambda G) = \emptyset$. A transformation method (which is similar to that used for solving (10.27)) can be employed to reduce (10.30) to a simpler form. Let Q_1 and Z_1 be orthogonal matrices such that the pair $(P, S) := (Q_1^T AZ_1, Q_1^T EZ_1)$ is in a GRSF, and let Q_2 and Z_2 be orthogonal matrices such that the pair $(R, T) := (Q_2^T FZ_2, Q_2^T GZ_2)$ is in a GRSF. The matrices Q_1 and Z_1, and Q_2 and Z_2 can be obtained by applying the QZ algorithms to the matrix pairs (A, E) and (F, G), respectively. If we define $X_1 = Z_1^T XZ_2$, $Y_1 = Q_1^T YQ_2$, $C_1 = Q_1^T CZ_2$ and $D_1 = Q_1^T DZ_2$, then the system (10.30) can be rewritten as

$$
\begin{aligned}
PX_1 + Y_1R &= C_1, \\
SX_1 + Y_1T &= D_1.
\end{aligned}
$$

By exploiting the upper quasi-triangular–upper triangular structures of the pairs (P, S) and (R, T), this system of equations can be efficiently solved using methods proposed in [68]. After solving the transformed system for X_1 and Y_1, we obtain the solution of (10.30) as $X = Z_1X_1Z_2^T$ and $Y = Q_1Y_1Q_2^T$. The overall computational effort to solve these equations is $\mathcal{O}(n^3 + m^3) + \mathcal{O}(n^2m + nm^2)$.

10.2.2 Generalized Algebraic Riccati Equations

In this section we address the numerical solution of a class of generalized Riccati equations which appear in various algorithms as the computation of inner–outer factorization (see Sect. 10.3.6) or in spectral factorization problems discussed in Sect. 7.8. We consider a sextuple of matrices (A, E, B, Q, S, R), with the following properties of component matrices: $A \in \mathbb{R}^{n\times n}$, $E \in \mathbb{R}^{n\times n}$ invertible, $B \in \mathbb{R}^{n\times m}$, $Q \in \mathbb{R}^{n\times n}$ symmetric positive semi-definite, $R \in \mathbb{R}^{m\times m}$ symmetric and invertible, and $S \in \mathbb{R}^{n\times m}$. We seek the symmetric positive semi-definite stabilizing solution $X_s \in \mathbb{R}^{n\times n}$ of the *generalized continuous-time algebraic Riccati equation* (GCARE)

$$A^TXE + E^TXA - (E^TXB + S)R^{-1}(B^TXE + S^T) + Q = 0$$

and the corresponding stabilizing state feedback gain $F_s \in \mathbb{R}^{m\times n}$, given by

$$F_s = -R^{-1}(B^TX_sE + S^T) \ ,$$

such that all generalized eigenvalues of the pair $(A + BF_s, E)$ have negative real parts. Similarly, we seek the symmetric positive semi-definite stabilizing solution $X_s \in \mathbb{R}^{n\times n}$ of the *generalized discrete-time algebraic Riccati equation* (GDARE)

$$A^TXA - E^TXE - (A^TXB + S)(R + B^TXB)^{-1}(B^TXA + S^T) + Q = 0$$

and the corresponding stabilizing state feedback gain $F_s \in \mathbb{R}^{m\times n}$, given by

$$F_s = -(R + B^TX_sB)^{-1}(B^TX_sA + S^T) \ ,$$

such that all generalized eigenvalues of the pair $(A + BF_s, E)$ have moduli less than one. Since E is invertible, it is possible to reduce both the GCARE and GDARE to standard Riccati equations, for which there exist standard solution methods. However, to avoid possible accuracy losses due to the need to explicitly invert E, we will indicate methods which directly tackle the above equations, without inverting E.

A unified approach to determine the solutions of the GCARE and GDARE relies on determining an orthogonal basis of the stable deflating subspace of a suitably defined regular matrix pencil $L - \lambda P$. For the solution of the GCARE we have

$$L = \begin{bmatrix} A & 0 & B \\ -Q & -A^T & -S \\ S^T & B^T & R \end{bmatrix}, \qquad P = \begin{bmatrix} E & 0 & 0 \\ 0 & E^T & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{10.31}$$

while for the solution of the GDARE we have

$$L = \begin{bmatrix} A & 0 & B \\ -Q & E^T & -S \\ S^T & 0 & R \end{bmatrix}, \qquad P = \begin{bmatrix} E & 0 & 0 \\ 0 & A^T & 0 \\ 0 & -B^T & 0 \end{bmatrix}. \tag{10.32}$$

Under fairly standard assumptions (e.g., the stabilizability of the pair $(A - \lambda E, B)$ and detectability of the pair $(A - \lambda E, Q - SR^{-1}S^T)$), the existence of the positive semi-definite stabilizing solution X_s is guaranteed. For computational purposes, the main property of the regular pencil $L - \lambda P$ is the existence of an n dimensional (right) deflating subspace corresponding to the stable eigenvalues of $L - \lambda P$. If this subspace is spanned by a $(2n + m) \times n$ matrix Z_1, then we have that $LZ_1 = PZ_1 W$, where W is an $n \times n$ matrix such that $\Lambda(W) \in \mathbb{C}_s$. If we partition Z_1 in accordance with the block column structure of the pencil $L - \lambda P$ as

$$Z_1 = \begin{bmatrix} Z_{11} \\ Z_{21} \\ Z_{31} \end{bmatrix}, \tag{10.33}$$

then the stabilizing positive definite solution X_s of both the GCARE and GDARE, and the corresponding stabilizing feedback F_s can be computed as

$$X_s = Z_{21}(EZ_{11})^{-1}, \qquad F_s = Z_{31}Z_{11}^{-1}.$$

To compute Z_1, we can employ the QZ algorithm to determine an ordered GRSF of the pair (L, P) in the from

$$U^T(L - \lambda P)Z = \begin{bmatrix} L_{11} - \lambda P_{11} & L_{12} - \lambda P_{12} \\ 0 & L_{22} - \lambda P_{22} \end{bmatrix}, \tag{10.34}$$

where U and Z are orthogonal transformation matrices, the $n \times n$ subpencil $L_{11} - \lambda P_{11}$ has only stable eigenvalues, i.e., $\Lambda(L_{11} - \lambda P_{11}) \subset \mathbb{C}_s$, and $\Lambda(L_{22} - \lambda P_{22}) \subset \mathbb{C} \setminus \mathbb{C}_s$. Then, Z_1 is formed from the first n columns of the orthogonal matrix Z. Using this approach, the overall computational effort for solving both the GCARE and GDARE is $\mathcal{O}((n + m)^3)$.

10.3 Algorithms for Descriptor Systems

In this section we present computational procedures for the solution of several basic computational problems for descriptor systems. The theoretical aspects of these problems have been succinctly addressed in Sect. 9.2, where several canonical forms (e.g., Weierstrass, Kronecker) played an important conceptual role in their solutions. However, these canonical forms are not suited to develop reliable numerical algorithms, due to the need of using potentially ill-conditioned transformations for their computation. We present reliable numerical algorithms, which rely on the alternative condensed forms discussed in Sect. 10.1. These forms can be computed using exclusively orthogonal transformations. Therefore, these algorithms are intrinsically numerically reliable and some of them are even numerically stable.

10.3.1 Minimal Realization

Consider a $p \times m$ rational matrix $G(\lambda)$ and let $(A - \lambda E, B, C, D)$ be an n-th order descriptor system realization satisfying

$$G(\lambda) = C(\lambda E - A)^{-1}B + D,$$

with $A - \lambda E$ an $n \times n$ regular pencil. If $Q, Z \in \mathbb{R}^{n \times n}$ are invertible matrices, then it is easy to check that two realizations $(A - \lambda E, B, C, D)$ and $(\widehat{A} - \lambda \widehat{E}, \widehat{B}, \widehat{C}, D)$, whose matrices are related by a similarity transformation of the form

$$\widehat{A} - \lambda \widehat{E} = Q(A - \lambda E)Z, \quad \widehat{B} = QB, \quad \widehat{C} = CZ\,,$$

have the same TFM $G(\lambda)$. Similarity transformations with Q and Z orthogonal matrices can be used to obtain various staircase forms of the system matrices, which allow to extract lower dimensional descriptor realizations of $G(\lambda)$, and finally to arrive to a minimal order realization with the least possible order n.

A minimal realization $(A - \lambda E, B, C, D)$ is characterized by the five conditions (i)–(v) of Theorem 9.2. An irreducible realization fulfils only conditions (i)–(iv) and is thus controllable and observable. In what follows, we describe a two-stage approach which first constructs an irreducible realization of lower order by successively removing the uncontrollable and unobservable eigenvalues of $A - \lambda E$, and in a second stage removes the non-dynamics modes (i.e., the simple infinite eigenvalues of $A - \lambda E$).

The first reduction stage is accomplished in four steps, by employing repeatedly **Procedure GCSF** to successively remove the finite uncontrollable, infinite uncontrollable, finite unobservable and infinite unobservable eigenvalues of $A - \lambda E$. At the first step of this reduction stage, we apply **Procedure GCSF** to the triple $(A - \lambda E, B, C)$ to obtain the orthogonal transformation matrices Q_1 and Z_1, such that the equivalent descriptor realization of $G(\lambda)$ has the form

$$\left[\begin{array}{c|c} Q_1^T(A-\lambda E)Z_1 & Q_1^T B \\ \hline CZ_1 & D \end{array}\right] = \left[\begin{array}{cc|c} A_c^f - \lambda E_c^f & * & B_c^f \\ 0 & A_{\bar{c}}^f - \lambda E_{\bar{c}}^f & 0 \\ \hline C_c^f & C_{\bar{c}}^f & D \end{array}\right]. \qquad (10.35)$$

The finite controllable descriptor system $(A_c^f - \lambda E_c^f, B_c^f, C_c^f, D)$ has the same TFM $G(\lambda)$ and its order $n_c^f \leq n$. By this step we can remove the $n - n_c^f$ uncontrollable eigenvalues of $A_{\bar{c}}^f - \lambda E_{\bar{c}}^f$ from the original descriptor system representation $(A - \lambda E, B, C, D)$. Besides all finite uncontrollable eigenvalues, $\Lambda(A_{\bar{c}}^f - \lambda E_{\bar{c}}^f)$ may also contain some of infinite uncontrollable eigenvalues of $A - \lambda E$.

At the second step of the reduction stage, we apply **Procedure GCSF** to the triple $(E_c^f - \lambda A_c^f, B_c^f, C_c^f)$ (note that A_c^f and E_c^f are interchanged) to obtain the orthogonal transformation matrices Q_2 and Z_2, such that the equivalent descriptor realization of $G(\lambda)$ has the form

$$\left[\begin{array}{c|c} Q_2^T(A_c^f-\lambda E_c^f)Z_2 & Q_2^T B_c^f \\ \hline C_c^f Z_2 & D \end{array}\right] = \left[\begin{array}{cc|c} A_c - \lambda E_c & * & B_c \\ 0 & A_{\bar{c}}^\infty - \lambda E_{\bar{c}}^\infty & 0 \\ \hline C_c & C_{\bar{c}}^\infty & D \end{array}\right]. \qquad (10.36)$$

As before, the controllable descriptor system $(A_c - \lambda E_c, B_c, C_c, D)$ has the same TFM $G(\lambda)$ and its order $n_c \leq n_c^f$. By this step we can remove the $n_c^f - n_c$ uncontrollable infinite eigenvalues of $A_{\bar{c}}^\infty - \lambda E_{\bar{c}}^\infty$ (or equivalently the uncontrollable zero eigenvalues of $E_{\bar{c}}^\infty - \lambda A_{\bar{c}}^\infty$) from the original descriptor system representation $(A - \lambda E, B, C, D)$.

At the third step, we apply **Procedure GCSF** to the dual triple $(E_c^T - \lambda A_c^T, C_c^T, B_c^T)$ to obtain the orthogonal transformation matrices Z_3 and Q_3 (note the changed order), such that the equivalent descriptor realization of $G(\lambda)$ has the form

$$\left[\begin{array}{c|c} P_3 Q_3^T(A_c-\lambda E_c)Z_3 P_3 & P_3 Q_3^T B_c \\ \hline C_c Z_3 P_3 & D \end{array}\right] = \left[\begin{array}{cc|c} A_{c\bar{o}}^f - \lambda E_{c\bar{o}}^f & * & B_{c\bar{o}}^f \\ 0 & A_{co}^f - \lambda E_{co}^f & B_{co}^f \\ \hline 0 & C_{co}^f & D \end{array}\right], \qquad (10.37)$$

where P_3 is the permutation matrix (10.10) of appropriate size. The controllable and finite observable descriptor system $(A_{co}^f - \lambda E_{co}^f, B_{co}^f, C_{co}^f, D)$ has the same TFM $G(\lambda)$ and its order $n_{co}^f \leq n_c$. By this step we can remove the $n_c - n_{co}^f$ unobservable eigenvalues of $A_{c\bar{o}}^f - \lambda E_{c\bar{o}}^f$ from the original descriptor system representation $(A - \lambda E, B, C, D)$.

Finally, at the fourth step, we apply **Procedure GCSF** to the dual triple $((E_{co}^f)^T - \lambda (A_{co}^f)^T, (C_{co}^f)^T, (B_{co}^f)^T)$ (note that A_{co}^f and E_{co}^f are interchanged) to obtain the orthogonal transformation matrices Z_4 and Q_4, such that the equivalent descriptor realization of $G(\lambda)$ has the form

$$\left[\begin{array}{c|c} P_4 Q_4^T (A_{co}^f - \lambda E_{co}^f) Z_4 P_4 & P_4 Q_4^T B_{co}^f \\ \hline C_{co}^f Z_4 P_4 & D \end{array}\right] = \left[\begin{array}{cc|c} A_{c\bar{o}}^\infty - \lambda E_{c\bar{o}}^\infty & * & B_{c\bar{o}}^\infty \\ 0 & A_{co} - \lambda E_{co} & B_{co} \\ \hline 0 & C_{co} & D \end{array}\right], \tag{10.38}$$

where P_4 is a permutation matrix as in (10.10) of appropriate size. The irreducible (i.e., controllable and observable) descriptor system $(A_{co} - \lambda E_{co}, B_{co}, C_{co}, D)$ has the same TFM $G(\lambda)$ and its order $n_{co} \leq n_{co}^f$. By this step we can remove the $n_{co}^f - n_{co}$ unobservable infinite eigenvalues of $A_{c\bar{o}}^\infty - \lambda E_{c\bar{o}}^\infty$ from the original descriptor system representation $(A - \lambda E, B, C, D)$.

With the overall transformation matrices defined as

$$Q := Q_1 \operatorname{diag}(Q_2, I) \operatorname{diag}(Q_3 P_3, I) \operatorname{diag}(I, Q_4 P_4, I),$$

$$Z := Z_1 \operatorname{diag}(Z_2, I) \operatorname{diag}(Z_3 P_3, I) \operatorname{diag}(I, Z_4 P_4, I),$$

we obtained the orthogonally similar system representation

$$(\widetilde{A} - \lambda \widetilde{E}, \widetilde{B}, \widetilde{C}, D) := (Q^T A Z - \lambda Q^T E Z, Q^T B, C Z, D),$$

with

$$\left[\begin{array}{c|c} \widetilde{A} - \lambda \widetilde{E} & \widetilde{B} \\ \hline \widetilde{C} & D \end{array}\right] = \left[\begin{array}{ccccc|c} A_{c\bar{o}}^f - \lambda E_{c\bar{o}}^f & * & * & * & * & B_{c\bar{o}}^f \\ 0 & A_{c\bar{o}}^\infty - \lambda E_{c\bar{o}}^\infty & * & * & * & B_{c\bar{o}}^\infty \\ 0 & 0 & A_{co} - \lambda E_{co} & * & * & B_{co} \\ 0 & 0 & 0 & A_{\bar{c}}^\infty - \lambda E_{\bar{c}}^\infty & * & 0 \\ 0 & 0 & 0 & 0 & A_{\bar{c}}^f - \lambda E_{\bar{c}}^f & 0 \\ \hline 0 & 0 & C_{co} & C_{\bar{c}}^\infty & C_{\bar{c}}^f & D \end{array}\right].$$

This form, obtained using exclusively orthogonal similarity transformations, represents a particular instance of a generalized Kalman decomposition of the descriptor system matrices from which an irreducible realization $(A_{co} - \lambda E_{co}, B_{co}, C_{co}, D)$ can be readily extracted. There are various ways to improve the efficiency of computations. For example, if the original realization corresponds to a proper system, then the second and fourth steps (i.e., removing of uncontrollable or unobservable infinite eigenvalues) can be skipped. Similar simplifications are possible —for example, if the original system description corresponds to a polynomial matrix, or if the original system representation is known to be controllable or observable, or if $A - \lambda E$ has no zero eigenvalues. In the latter case, only the second and fourth steps need to be performed.

The whole computational approach is summarized in the following procedure, which computes for a given triple $(A - \lambda E, B, C)$ an irreducible (i.e., controllable and observable) triple $(A_{co} - \lambda E_{co}, B_{co}, C_{co})$.

Procedure GIR: Generalized irreducible realization algorithm

Input : $(A - \lambda E, B, C)$
Output: Irreducible $(A_{co} - \lambda E_{co}, B_{co}, C_{co})$

1) Perform **Procedure GCSF** on the triple $(A - \lambda E, B, C)$ and extract the finite controllable triple $(A_c^f - \lambda E_c^f, B_c^f, C_c^f)$.
2) Perform **Procedure GCSF** on the triple $(E_c^f - \lambda A_c^f, B_c^f, C_c^f)$ and extract the controllable triple $(A_c - \lambda E_c, B_c, C_c)$.
3) With P an appropriate permutation matrix as in (10.10), perform **Procedure GCSF** on the triple $(PA_c^T P - \lambda PE_c^T P, PC_c^T, B_c^T P)$ and extract the controllable and finite observable triple $(A_{co}^f - \lambda E_{co}^f, B_{co}^f, C_{co}^f)$.
4) With P an appropriate permutation matrix as in (10.10), perform **Procedure GCSF** on the triple $(P(E_{co}^f)^T P - \lambda P(A_{co}^f)^T P, P(C_{co}^f)^T, (B_{co}^f)^T P)$ and build the irreducible triple $(A_{co} - \lambda E_{co}, B_{co}, C_{co})$.

At the end of Step 1), A_c^f is in an upper block Hessenberg form and E_c^f is upper triangular. The upper block Hessenberg shape of A_c^f at Step 2) can be exploited by the **Procedure GCSF**, to reduce the computational burden at the initial reduction of A_c^f to an upper triangular form. The resulting A_c at Step 2) is therefore upper triangular, while E_c is upper block Hessenberg. At Step 3), the use of $PE_c^T P$ instead of E_c^T allows to preserve the upper block Hessenberg form of E_c obtained at the previous step. This is also the case at Step 4), where the upper block Hessenberg structure of A_{co}^f is preserved when using $P(A_{co}^f)^T P$ instead.

The computational effort for **Procedure GIR** is $\mathcal{O}(n^3)$ for $m, p \ll n$. It is possible to show that the computed irreducible descriptor system $(A_{co} - \lambda E_{co}, B_{co}, C_{co}, D)$ is exact for a slightly perturbed original system. Therefore, the **Procedure GIR** can be considered numerically stable.

In the second stage, we have to remove the simple infinite eigenvalues of $A_{co} - \lambda E_{co}$ from the resulting irreducible descriptor representation. For this purpose, we isolate the simple infinite eigenvalues by employing two SVDs. First, we compute the SVD of E_{co} such that

$$U_1^T E_{co} V_1 = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix},$$

with U_1 and V_1 orthogonal matrices and E_{11} a (diagonal) invertible matrix of rank r. Applying the same transformations to A we obtain

$$U_1^T A_{co} V_1 = \begin{bmatrix} A_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{bmatrix}.$$

Now, we compute the SVD of $\widetilde{A}_{22}$ such that

$$U_2^T \widetilde{A}_{22} V_2 = \begin{bmatrix} A_{22} & 0 \\ 0 & 0 \end{bmatrix},$$

with U_2 and V_2 orthogonal matrices and A_{22} a (diagonal) invertible matrix of rank q. With $U = U_1\,\mathrm{diag}(I_r, U_2)$ and $V = V_1\,\mathrm{diag}(I_r, V_2)$ we have the equivalent descriptor realization

$$\left[\begin{array}{c|c} U^T A_{co} V - \lambda U^T E_{co} V & U^T B_{co} \\ C_{co} V & D \end{array}\right] = \left[\begin{array}{ccc|c} A_{11} - \lambda E_{11} & A_{12} & A_{13} & B_1 \\ A_{21} & A_{22} & 0 & B_2 \\ A_{31} & 0 & 0 & B_3 \\ \hline C_1 & C_2 & C_3 & D \end{array}\right].$$

At this step, we have the transformed state vector $\widetilde{x}(t) := V^T x(t)$ partitioned into three components

$$\widetilde{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix},$$

which correspond to the column structure of $U^T A V$. We can eliminate the second component $x_2(t)$ as

$$x_2(t) = -A_{22}^{-1} A_{21} x_1(t) - A_{22}^{-1} B_2 u(t)$$

and obtain a descriptor representation with the reduced state vector $\overline{x}(t) = \begin{bmatrix} x_1(t) \\ x_3(t) \end{bmatrix}$ and the corresponding minimal realization $(\overline{A} - \lambda\overline{E}, \overline{B}, \overline{C}, \overline{D})$ of $G(\lambda)$, with the matrices given by

$$\overline{A} - \lambda\overline{E} = \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} - \lambda E_{11} & A_{13} \\ A_{31} & 0 \end{bmatrix}, \quad \overline{B} = \begin{bmatrix} B_1 - A_{12}A_{22}^{-1}B_2 \\ B_3 \end{bmatrix},$$

$$\overline{C} = \begin{bmatrix} C_1 - C_2 A_{22}^{-1} A_{21} & C_3 \end{bmatrix}, \qquad \overline{D} = D - C_2 A_{22}^{-1} B_2.$$

This final elimination step involves non-orthogonal matrix operations, which can lead to unstable computations if the norm of the intervening matrices is too large or A_{22} is ill conditioned. Fortunately, in most computational algorithms for descriptor systems presented in this book, the elimination of simple infinite eigenvalues is not necessary and we can work with irreducible realizations instead minimal ones. Therefore, we can almost always delay the computation of minimal realizations for the final results of whole computational cycles.

10.3.2 Minimal Proper Rational Nullspace Bases

Let $G(\lambda)$ be a $p \times m$ rational matrix of normal rank r. A proper rational basis of the left nullspace $\mathcal{N}_L(G(\lambda))$ (see Sect. 9.1.3) is any $(p - r) \times p$ proper rational matrix $N_l(\lambda)$ of full row rank such that $N_l(\lambda)G(\lambda) = 0$. Similarly, a proper rational basis of the right nullspace $\mathcal{N}_R(G(\lambda))$ is an $m \times (m - r)$ proper rational matrix $N_r(\lambda)$ of full column rank such that $G(\lambda)N_r(\lambda) = 0$. Of special interest are the *minimal* proper rational bases, which have the least McMillan degree. Assume $G(\lambda)$ has an n-th order descriptor system realization $(A - \lambda E, B, C, D)$, with $A - \lambda E$ regular. In this section we present a numerically reliable computational approach to determine a descriptor system realization of a proper rational left nullspace basis $N_l(\lambda)$ of $G(\lambda)$ and discuss conditions for its minimality. The same approach can be used to determine $N_r(\lambda)$, a proper rational right nullspace basis of $G(\lambda)$, by determining $N_r^T(\lambda)$ as a proper rational left nullspace basis of $G^T(\lambda)$.

The proposed computational approach relies on the fact that $N_l(\lambda)$ is a left nullspace basis of $G(\lambda)$ if and only if, for a suitable $(p - r) \times n$ rational matrix $M_l(\lambda)$,

$$Y_l(\lambda) := [\, M_l(\lambda)\ N_l(\lambda)\,] \tag{10.39}$$

is a left nullspace basis of the system matrix

$$S(\lambda) = \begin{bmatrix} A - \lambda E & B \\ C & D \end{bmatrix}. \tag{10.40}$$

Thus, to compute $N_l(\lambda)$ we can first determine a left nullspace basis $Y_l(\lambda)$ for $S(\lambda)$ and then $N_l(\lambda)$ simply results as

$$N_l(\lambda) = Y_l(\lambda) \begin{bmatrix} 0 \\ I_p \end{bmatrix}. \tag{10.41}$$

As it will be apparent below, the main appeal of this approach is that for the computation of $Y_l(\lambda)$ we can employ powerful pencil manipulation techniques via orthogonal similarity transformations.

Let U and V be orthogonal matrices such that the transformed pencil $\widetilde{S}(\lambda) := US(\lambda)V$ is in the Kronecker-like staircase form (see Sect. 10.1.6)

$$\widetilde{S}(\lambda) = \left[\begin{array}{cc} A_r - \lambda E_r & A_{r,l} - \lambda E_{r,l} \\ \hline 0 & A_l - \lambda E_l \\ \hline 0 & C_l \end{array} \right], \tag{10.42}$$

where the descriptor pair $(A_l - \lambda E_l, C_l)$ is observable, E_l is invertible, and $A_r - \lambda E_r$ has full row rank excepting possibly a finite set of values of λ (i.e., the invariant zeros of $S(\lambda)$). As explained in Sect. 10.1.6, the reduction of $S(\lambda)$ to the form (10.42) can be obtained using (twice) the **Procedure PREDUCE**.

A left nullspace $\widetilde{Y}_l(\lambda)$ of $\widetilde{S}(\lambda)$ in (10.42) can be chosen in the form

$$\widetilde{Y}_l(\lambda) = \left[\ 0 \ \middle|\ C_l(\lambda E_l - A_l)^{-1} \ \middle|\ I\ \right]. \tag{10.43}$$

Then, the left nullspace of $S(\lambda)$ is $Y_l(\lambda) = \widetilde{Y}_l(\lambda)U$ and can be obtained easily after partitioning suitably U as

$$U = \begin{bmatrix} \widehat{B}_{r,l} & B_{r,l} \\ \widehat{B}_l & B_l \\ \widehat{D}_l & D_l \end{bmatrix},$$

where the row partitioning corresponds to the column partitioning of $\widetilde{Y}_l(\lambda)$ in (10.43), while the column partitioning corresponds to the row partitioning of $S(\lambda)$ in (10.40). We obtain

$$Y_l(\lambda) = \left[\begin{array}{c|cc} A_l - \lambda E_l & \widehat{B}_l & B_l \\ \hline C_l & \widehat{D}_l & D_l \end{array}\right] \tag{10.44}$$

and the nullspace of $G(\lambda)$ is

$$N_l(\lambda) = \left[\begin{array}{c|c} A_l - \lambda E_l & B_l \\ \hline C_l & D_l \end{array}\right]. \tag{10.45}$$

To obtain this representation of the nullspace basis, we performed exclusively orthogonal transformations on the system matrices. We can prove that all computed matrices are exact for a slightly perturbed original system matrix (10.40). It follows that this method for the computation of the nullspace basis is numerically backward stable.

When using **Procedure PREDUCE**, as described in Sect. 10.1.6, to determine the Kronecker-like form (10.42), we can assume that the resulting subpencil

$$\begin{bmatrix} A_o - \lambda E_o \\ C_o \end{bmatrix} := \begin{bmatrix} A_l - \lambda E_l \\ C_l \end{bmatrix}, \tag{10.46}$$

which characterizes the left structure of $S(\lambda)$, has the pair $(A_o - \lambda E_o, C_o)$ in an observability staircase form as in (10.17) and (10.18). Let μ_i, $i = 1, \ldots, \ell$ be the dimensions of the diagonal blocks of A_o in (10.17) (and also of E_o in (10.18)), and define $\mu_0 := p_l$ and $\mu_{\ell+1} := 0$ (which corresponds to a fictive full column rank diagonal block $A_{\ell,\ell+1} \in \mathbb{R}^{\mu_\ell \times \mu_{\ell+1}}$ in the leading position of A_o). These dimensions completely determine the left Kronecker structure of $S(\lambda)$ as follows: there are $\mu_{i-1} - \mu_i$ blocks $L^T_{i-1}(\lambda)$ of size $i \times (i-1)$, $i = 1, \ldots, \ell+1$ (see (9.45)). The row dimension of $N_l(\lambda)$ (i.e., the number of linearly independent basis vectors) is given by the total number of $L^T_{\eta_i}(\lambda)$ blocks (see Example 9.1), thus $\sum_{i=1}^{\ell+1}(\mu_{i-1} - \mu_i) = \mu_0$ (i.e., the row dimension of C_l). Applying standard linear algebra results, it follows that $\mu_0 := p - r$.

The following result shows that the resulting staircase form (10.46) provides the complete structural information on any minimal polynomial basis (and also on any simple proper basis constructed from it, see Sect. 9.1.3).

Proposition 10.1 *If the realization $(A - \lambda E, B, C, D)$ of $G(\lambda)$ is controllable and if μ_i, $i = 1, \ldots, \ell$ are the dimensions of the diagonal blocks of A_o in (10.17) (and also of E_o in (10.18)), and $\mu_0 := p_l$ and $\mu_{\ell+1} := 0$, then a minimal polynomial basis of the left nullspace of $G(\lambda)$ has degree $n_l = \sum_{i=1}^{\ell} \mu_i$ and is formed of $\mu_{i-1} - \mu_i$ polynomial vectors of degree $i-1$, for $i = 1, \ldots, \ell+1$.*

Proof The controllability of the descriptor realization ensures that the left Kronecker structure of $G(\lambda)$ and of $S(\lambda)$ are characterized by the same left Kronecker indices. A minimal polynomial basis for the left nullspace of $\widetilde{S}(\lambda)$ can be determined of the form

$$\widehat{Y}_l(\lambda) = \left[\begin{array}{c|c} 0 & \widehat{N}_l(\lambda) \end{array}\right], \tag{10.47}$$

where $\widehat{N}_l(\lambda)$ is a minimal polynomial basis for the left nullspace of $\begin{bmatrix} A_l - \lambda E_l \\ C_l \end{bmatrix}$. To construct $\widehat{N}_l(\lambda)$, the basis vectors can be determined by exploiting the staircase form of this pencil. It was shown in [8, Sect. 4.6.4], in a dual context, that a minimal polynomial basis can be computed by selecting $\mu_{i-1} - \mu_i$ polynomial basis vectors of degree $i-1$, for $i = 1, \ldots, \ell+1$. The degree of this polynomial basis is

$$\begin{aligned} \textstyle\sum_{i=1}^{\ell+1}(\mu_{i-1} - \mu_i)(i-1) &= \textstyle\sum_{i=1}^{\ell+1}\mu_{i-1}(i-1) - \sum_{i=1}^{\ell+1}\mu_i(i-1) \\ &= \textstyle\sum_{i=1}^{\ell}\mu_i i - \sum_{i=1}^{\ell}\mu_i(i-1) \\ &= \textstyle\sum_{i=1}^{\ell}\mu_i\,, \end{aligned}$$

which is equal to n_l, the dimension of the square matrices A_l and E_l. ■

A straightforward consequence of Proposition 10.1 is the following result.

Proposition 10.2 *If the realization $(A - \lambda E, B, C, D)$ of $G(\lambda)$ is controllable, then the rational matrix $N_l(\lambda)$ defined in (10.45) is a minimal proper rational basis of the left nullspace of $G(\lambda)$.*

Proof According to the definition of a minimal proper rational basis (see Sect. 9.1.3), its McMillan degree is given by the degree of a minimal polynomial basis (i.e., the sum of the left minimal indices). By Proposition 10.1, the degree of a minimal polynomial basis is $n_l := \sum_{i=1}^{\ell}\mu_i$, which is thus equal to the dimension of the square matrices A_l and E_l. Therefore, we only need to show that the realization (10.45) is irreducible and $N_l(\lambda)$ defined in (10.45) has no zeros.

The pair $(A_l - \lambda E_l, C_l)$ is observable, by the construction of the Kronecker-like form (10.42). To show that the pair $(A_l - \lambda E_l, B_l)$ is controllable, observe that due to the controllability of the pair $(A - \lambda E, B)$, the subpencil $[\, A - \lambda E \;\; B \,]$ of $S(\lambda)$ in (10.40) has full row rank for all $\lambda \in \mathbb{C}$, and thus the reduced pencil

$$U \begin{bmatrix} A - \lambda E & B & 0 \\ C & D & I_p \end{bmatrix} \left[\begin{array}{c|c} V & 0 \\ 0 & I_p \end{array}\right] = \left[\begin{array}{cc|c} A_r - \lambda E_r & A_{r,l} - \lambda E_{r,l} & B_{r,l} \\ 0 & A_l - \lambda E_l & B_l \\ 0 & C_l & D_l \end{array}\right]$$

has full row rank for all $\lambda \in \mathbb{C}$ as well. It follows that for all $\lambda \in \mathbb{C}$

$$\operatorname{rank} \begin{bmatrix} A_l - \lambda E_l & B_l \end{bmatrix} = n_l$$

and thus the pair $(A_l - \lambda E_l, B_l)$ is controllable.

Since, we also have that

$$\operatorname{rank}\begin{bmatrix} A_l - \lambda E_l & B_l \\ C_l & D_l \end{bmatrix} = n_l + p - r$$

for all $\lambda \in \mathbb{C}$, it follows that $N_l(\lambda)$ has no finite or infinite zeros. Thus, D_l has full row rank $p - r$ and the computed basis is column reduced at $\lambda = \infty$ [122]. ■

In the case, when the realization of $G(\lambda)$ is not controllable, the realization of $N_l(\lambda)$ is not guaranteed to be controllable. The uncontrollable eigenvalues of $A - \lambda E$ may turn partly up as eigenvalues of $A_r - \lambda E_r$ (i.e., invariant zeros) or of $A_l - \lambda E_l$. In the latter case, the resulting proper nullspace basis has not the least possible McMillan degree. Interestingly, a minimal basis cannot be always obtained by simply eliminating the uncontrollable part of the pair $(A_l - \lambda E_l, B_l)$. The reason for this is the lack of the maximal controllability property (see Proposition 10.3).

We can always determine a proper nullspace basis with arbitrarily assigned poles. To show this, consider the transformation matrix

$$\widehat{U} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & K \\ 0 & 0 & I \end{bmatrix} \tag{10.48}$$

and compute $\widehat{S}(\lambda) := \widehat{U}\widetilde{S}(\lambda)$ as

$$\widehat{S}(\lambda) = \left[\begin{array}{cc} A_r - \lambda E_r & A_{r,l} - \lambda E_{r,l} \\ \hline 0 & A_l + KC_l - \lambda E_l \\ \hline 0 & C_l \end{array}\right]. \tag{10.49}$$

We also compute

$$\widehat{U}U\begin{bmatrix} 0 \\ I_p \end{bmatrix} = \begin{bmatrix} B_{r,l} \\ B_l + KD_l \\ D_l\,, \end{bmatrix}$$

and obtain the proper rational left nullspace basis with the alternative realization

$$\widetilde{N}_l(\lambda) = \left[\begin{array}{c|c} A_l + KC_l - \lambda E_l & B_l + KD_l \\ \hline C_l & D_l \end{array}\right]. \tag{10.50}$$

Since the descriptor pair $(A_l - \lambda E_l, C_l)$ is completely observable, there exists an output injection matrix K such that the pair $(A_l + KC_l, E_l)$ has arbitrary assigned generalized eigenvalues. According to Proposition 10.2, the basis (10.50) is minimal provided the realization $(A - \lambda E, B, C, D)$ of $G(\lambda)$ is controllable.

This construction shows that the placement of poles of the left nullspace basis (10.50) can be simply achieved by additionally performing a particular similarity transformation on the reduced pencil $\widetilde{S}(\lambda)$. As a consequence, the output injection may move some of uncontrollable generalized eigenvalues of the pair (A_l, E_l) to other locations and make them controllable. It follows that determining a minimal nullspace basis from a non-minimal one may involve the determination of suitable injection matrix, which makes a maximum number of eigenvalues uncontrollable.

Remark 10.5 The alternative proper left nullspace basis $\widetilde{N}_l(\lambda)$ can be interpreted as the numerator factor of the left coprime factorization

$$N_l(\lambda) = \widetilde{M}_l^{-1}(\lambda)\widetilde{N}_l(\lambda) ,$$

where $\widetilde{M}_l(\lambda)$ has the descriptor system realization

$$\widetilde{M}_l(\lambda) = \left[\begin{array}{c|c} A_l + KC_l - \lambda E_l & K \\ \hline C_l & I_{p-r} \end{array}\right].$$

□

The following result shows that a minimal proper basis of the form (10.45) has the nice property of being *maximally controllable*, that is, the alternative basis (10.50) remains controllable for an arbitrary output injection matrix K, or equivalently, the pair $(A_l + KC_l - \lambda E_l, B_l + KD_l)$ is controllable for all K.

Proposition 10.3 *If the realization $(A - \lambda E, B, C, D)$ of $G(\lambda)$ is controllable, then the realization of $N_l(\lambda)$ defined in (10.45) is maximally controllable.*

Proof We have to show that for an arbitrary output injection matrix K, the pair $(A_l + KC_l - \lambda E_l, B_l + KD_l)$ is controllable. Let K be an arbitrary injection matrix and construct the alternative proper left nullspace basis $\widetilde{N}_l(\lambda)$ with the realization given in (10.50). Since according to Proposition 10.2, $N_l(\lambda)$ is a minimal nullspace basis, the alternative nullspace basis $\widetilde{N}_l(\lambda)$, with the same McMillan degree, is a minimal basis as well. Therefore, the pair $(A_l + KC_l - \lambda E_l, B_l + KD_l)$ is controllable. ■

Even if the resulting rational basis (10.45) has the least possible McMillan degree, and thus is minimal, still, in general, this basis is not a simple basis. The properties of simple proper minimal bases resemble, in many aspects, the properties of minimal polynomial bases. For our purposes, the main use of simple proper nullspace bases is in the nullspace-based synthesis methods of least-order fault detection filters. As it will be shown below, it is possible to obtain a simple basis starting from a non-simple one.

Consider the proper minimal left nullspace basis $N_l(\lambda)$ of $G(\lambda)$, with the descriptor realization given in (10.45), and we denote with $c_{l,i}$ and $d_{l,i}$ the i-th rows of matrices C_l and D_l, respectively. The approach to construct a simple minimal proper rational left nullspace basis is based on the following result.

Proposition 10.4 *For each $i = 1, \ldots, p - r$, let K_i be an output injection matrix such that*

$$v_i(\lambda) := c_{l,i}(\lambda E_l - A_l - K_i C_l)^{-1}(B_l + K_i D_l) + d_{l,i} \qquad (10.51)$$

has the least possible McMillan degree. Then, $\widetilde{N}_l(\lambda)$ formed by stacking the $p-r$ rational row vectors $v_i(\lambda)$ is a simple minimal proper rational left nullspace basis.

Proof According to Proposition 10.3, the realization (10.45) of $N_l(\lambda)$ is maximally controllable, i.e., the pair $(A_l + K_iC_l - \lambda E_l, B_l + K_iD_l)$ is controllable for arbitrary K_i. Therefore, the maximal order reduction of the McMillan degree of $v_i(\lambda)$ can be achieved by making the pair $(A_l + K_iC_l - \lambda E_l, c_{l,i})$ maximally unobservable via an appropriate choice of K_i. For each $i = 1, \ldots, p-r$, the achievable least McMillan degree of $v_i(\lambda)$ is the corresponding minimal index n_i, representing, in a dual setting, the dimension of the least-order controllability subspace of the standard pair $(E_l^{-T}A_l^T, E_l^{-T}C_l^T)$ containing span $(E_l^{-T}c_{l,i}^T)$. This result is the statement of Lemma 6 in [159]. It is easy to check that $v_i(\lambda)G(\lambda) = 0$, thus $\widetilde{N}_l(\lambda)$ is a left annihilator of $G(\lambda)$. Furthermore, the set of vectors $\{ v_1(\lambda), \ldots, v_{p-r}(\lambda) \}$ is linearly independent since the realization of $\widetilde{N}_l(\lambda)$ has the same full row rank matrix D_l as that of $N_l(\lambda)$. It follows that $\widetilde{N}_l(\lambda)$ is a proper left nullspace basis of least dimension $\sum_{i=1}^{p-r} n_i$, with each row $v_i(\lambda)$ of McMillan degree n_i. It follows that $\widetilde{N}_l(\lambda)$ is simple. ■

Let assume that each rational vector $v_i(\lambda)$ has a descriptor realization of the form

$$v_i(\lambda) = \left[\begin{array}{c|c} \widetilde{A}_{l,i} - \lambda\widetilde{E}_{l,i} & \widetilde{B}_{l,i} \\ \hline \widetilde{c}_{l,i} & d_{l,i} \end{array}\right]. \tag{10.52}$$

Then, the simple minimal proper rational basis $\widetilde{N}_l(\lambda)$, constructed by stacking all $v_i(\lambda)$, for $i = 1, \ldots, r$, has the realization

$$\widetilde{N}_l(\lambda) = \left[\begin{array}{c|c} \widetilde{A}_l - \lambda\widetilde{E}_l & \widetilde{B}_l \\ \hline \widetilde{C}_l & D_l \end{array}\right], \tag{10.53}$$

with

$$\widetilde{A}_l - \lambda\widetilde{E}_l = \begin{bmatrix} \widetilde{A}_{l,1} - \lambda\widetilde{E}_{l,1} & & \\ & \ddots & \\ & & \widetilde{A}_{l,p-r} - \lambda\widetilde{E}_{l,p-r} \end{bmatrix}, \quad \widetilde{B}_l = \begin{bmatrix} \widetilde{B}_{l,1} \\ \vdots \\ \widetilde{B}_{l,p-r} \end{bmatrix},$$

$$\widetilde{C}_l = \begin{bmatrix} \widetilde{c}_{l,1} & & \\ & \ddots & \\ & & \widetilde{c}_{l,p-r} \end{bmatrix}.$$

Remark 10.6 The poles of the simple minimal proper rational left nullspace basis $\widetilde{N}_l(\lambda)$ can be arbitrarily placed by performing left coprime rational factorizations using the realizations in (10.52) (see Remark 10.5)

$$v_i(\lambda) = m_i(\lambda)^{-1}\hat{v}_i(\lambda), \tag{10.54}$$

where $m_i(\lambda)$ are polynomials with arbitrary roots in $\mathbb{C}_s$. Therefore, the resulting alternative simple basis $\widehat{N}_l(\lambda) := [\, \hat{v}_1^T(\lambda), \; \ldots, \; \hat{v}_{p-r}^T(\lambda) \,]^T$ can have arbitrarily assigned poles. In particular, a *special simple basis* can be constructed such that each $m_i(\lambda)$ divides $m_j(\lambda)$, if $j < i$. □

Simple rational bases are direct correspondents of polynomial bases and, hence, all operations on polynomial bases have analogous operations on simple rational bases. An important operation (with applications in the synthesis of least-order fault detection filters) is building linear combinations of basis vectors up to a certain McMillan degree. For example, using the special simple basis in Remark 10.6, any linear combination $\sum_{i=1}^{k} h_i \hat{v}_i(\lambda)$ with constant coefficients h_i of the basis vectors of McMillan degree up to a certain value k has McMillan degree at most k.

Consider the proper left nullspace basis $N_l(\lambda)$ constructed in (10.45). From the details of the resulting staircase form (10.46) of the pair $(A_l - \lambda E_l, C_l)$, recall that it is possible to obtain the full column rank matrices $A_{i-1,i} \in \mathbb{R}^{\mu_{i-1} \times \mu_i}$ in the form

$$A_{i-1,i} = \begin{bmatrix} R_{i-1,i} \\ 0 \end{bmatrix},$$

where $R_{i-1,i}$ is an upper triangular invertible matrix of order μ_i. The row dimension $\mu_{i-1} - \mu_i$ of the zero block of $A_{i-1,i}$ gives the number of polynomial vectors of degree $i-1$ in a minimal polynomial basis [8, Sect. 4.6] and thus, also the number of vectors of McMillan degree $i-1$ in a simple basis. It is straightforward to show the following result.

Corollary 10.1 *For a given minimal proper rational left nullspace basis $N_l(\lambda)$ in the form (10.45), let i be a given index such that $1 \leq i < p - r$, and let h be a $(p-r)$-dimensional row vector having only the trailing i components nonzero. Then, a linear combination of the simple proper rational basis vectors, with McMillan degree at most n_i, can be generated as*

$$v(\lambda) := hC_l(\lambda E_l - A_l - KC_l)^{-1}(B_l + KD_l) + hD_l \,, \tag{10.55}$$

where K is an output injection matrix such that $v(\lambda)$ has the least possible McMillan degree.

This result shows that the determination of a linear combination of vectors of a simple proper rational basis up to a given order n_i is possible directly from a proper rational basis determined in the form (10.45). The matrix K together with a minimal realization of $v(\lambda)$ can be computed efficiently using minimal dynamic cover techniques presented in Sect. 10.4.2. The same approach can be applied repeatedly to determine the basis vectors $v_i(\lambda)$, $i = 1, \ldots, p-r$, of a simple basis using the particular choices $h = e_i^T$, where e_i is the i-th column of I_{p-r}.

10.3.3 Poles and Zeros Computation

The computation of poles of a rational matrix $G(\lambda)$, with an irreducible descriptor system realization $(A-\lambda E, B, C, D)$, comes down to compute the eigenvalues of the regular pole pencil $A-\lambda E$. This can be achieved by computing the eigenvalues of $A-\lambda E$ from the GRSF of the pair (A, E). The finite poles are the n_p^f finite eigenvalues of $A-\lambda E$, while there are $n_p^\infty = \operatorname{rank} E - n_p^f$ infinite poles (recall that the multiplicities of infinite eigenvalues are in excess with one with respect to the multiplicities of infinite poles). The McMillan degree of $G(\lambda)$ results as

$$\delta\big(G(\lambda)\big) = n_p^f + n_p^\infty = \operatorname{rank} E\,.$$

A straightforward application of the Kronecker-like form is the computation of the system zeros. Let $G(\lambda)$ be a rational matrix, with an irreducible descriptor system representation $(A-\lambda E, B, C, D)$. The system zeros are those values of λ, where the system pencil

$$S(\lambda) = \begin{bmatrix} A-\lambda E & B \\ C & D \end{bmatrix} := M - \lambda N$$

drops its rank below its normal rank. Thus, the system zeros can be determined from the eigenvalues of the regular pencil $M_{reg} - \lambda N_{reg}$ in the Kronecker-like form (10.19) of the pencil $M-\lambda N$. This can be achieved by computing the eigenvalues of $M_{reg} - \lambda N_{reg}$ from the GRSF of the pair (M_{reg}, N_{reg}). If $M_{reg} - \lambda N_{reg}$ has n_z^f finite eigenvalues, these are the n_z^f finite transmission zeros of the system. Additionally, there are $n_z^\infty = \operatorname{rank} N_{reg} - n_z^f$ infinite zeros (recall that the multiplicities of infinite eigenvalues are in excess with one with respect to the multiplicities of infinite zeros).

10.3.4 Additive Decompositions

Consider a disjunct partition of the complex plane $\mathbb{C}$ as $\mathbb{C} = \mathbb{C}_g \cup \mathbb{C}_b$, where both $\mathbb{C}_g$ and $\mathbb{C}_b$ are symmetrically located with respect to the real axis, $\mathbb{C}_g$ has at least one point on the real axis, and $\mathbb{C}_g \cap \mathbb{C}_b = \emptyset$. Let $G(\lambda)$ be a rational TFM with a descriptor system realization $(A-\lambda E, B, C, D)$. We describe a state-space approach to compute the additive decomposition

$$G(\lambda) = G_g(\lambda) + G_b(\lambda), \tag{10.56}$$

where $G_g(\lambda)$ has only poles in $\mathbb{C}_g$, while $G_b(\lambda)$ has only poles in $\mathbb{C}_b$.

The additive spectral decomposition (10.56) can be computed using a block diagonalization technique of the pole pencil $A-\lambda E$ (9.68). The basic computation is to determine the two invertible matrices U and Z to bring the matrices of the transformed pair (UEZ, UAZ) in suitable block diagonal forms. The following procedure computes the additive decomposition (10.56) using the descriptor realization $(A - \lambda E, B, C, D)$ of $G(\lambda)$, by determining the descriptor realizations $(A_g - \lambda E_g, B_g, C_g, D_g)$ of $G_g(\lambda)$ and $(A_b - \lambda E_b, B_b, C_b, D_b)$ of $G_b(\lambda)$.

Procedure GSDEC: Generalized additive spectral decomposition

Inputs : $G(\lambda) = (A - \lambda E, B, C, D)$, $\mathbb{C}_g$
Outputs: $G_g(\lambda) = (A_g - \lambda E_g, B_g, C_g, D_g)$, $G_b(\lambda) = (A_b - \lambda E_b, B_b, C_b, D_b)$

1) Using the QZ algorithm, compute orthogonal U_1 and V_1, such that the matrix pair $(U_1 A V_1, U_1 E V_1)$ is in an ordered GRSF

$$U_1 A V_1 = \begin{bmatrix} A_g & A_{gb} \\ 0 & A_b \end{bmatrix}, \quad U_1 E V_1 = \begin{bmatrix} E_g & E_{gb} \\ 0 & E_b \end{bmatrix},$$

such that $\Lambda(A_g - \lambda E_g) \subset \mathbb{C}_g$ and $\Lambda(A_b - \lambda E_b) \subset \mathbb{C} \setminus \mathbb{C}_g$.
2) Compute the left and right transformation matrices, U_2 and V_2, respectively, of the form

$$U_2 = \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}, \qquad V_2 = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix},$$

where X and Y satisfy the Sylvester system of equations

$$\begin{aligned} A_g X + Y A_b &= -A_{gb}\,, \\ E_g X + Y E_b &= -E_{gb}\,. \end{aligned}$$

3) Compute

$$\begin{bmatrix} B_g \\ B_b \end{bmatrix} = U_2 U_1 B, \quad \begin{bmatrix} C_g & C_b \end{bmatrix} = C V_1 V_2, \quad D_g = D, \quad D_b = 0,$$

where the row partitioning of $U_2 U_1 B$ and column partitioning of $C V_1 V_2$ are analogous to the row and column partitioning of $U_1 A V_1$.

The resulting pencil $U_2 U_1 (A - \lambda E) V_1 V_2$ at Step 2) is block diagonal. The existence of a unique solution (X, Y) of the Sylvester system to be solved at Step 2) is guaranteed by $\Lambda(A_g - \lambda E_g) \cap \Lambda(A_b - \lambda E_b) = \emptyset$. An efficient solution method, which exploits the GRSFs of the pairs (A_g, E_g) and (A_b, E_b), has been proposed in [68].

10.3.5 Coprime Factorizations

Consider a $p \times m$ rational matrix $G(\lambda)$ having a descriptor system realization $(A - \lambda E, B, C, D)$, for which we will not assume further properties (e.g., minimality or irreducibility). Consider also a disjunct partition of the complex plane as $\mathbb{C} = \mathbb{C}_b \cup \mathbb{C}_g$, $\mathbb{C}_b \cap \mathbb{C}_g = \emptyset$, where $\mathbb{C}_b$ and $\mathbb{C}_g$ denote the "bad" and "good" regions of $\mathbb{C}$, respectively. In this section we present algorithms for the computation of a *right coprime factorization* (RCF) of $G(\lambda)$ in the form $G(\lambda) = N(\lambda)M^{-1}(\lambda)$, where $N(\lambda)$ and $M(\lambda)$ are proper rational matrices with all poles in $\mathbb{C}_g$ and are mutually coprime (see Sect. 9.1.6 for definitions). A special case relevant for many applications is when $\mathbb{C}_g = \mathbb{C}_s$ and $\mathbb{C}_b = \mathbb{C} \setminus \mathbb{C}_g$ and we additionally impose that the denominator factor $M(\lambda)$ is inner. The algorithms to compute RCFs can be equally employed to determine a *left coprime factorization* (LCF) of $G(\lambda)$ in the form $G(\lambda) = M^{-1}(\lambda)N(\lambda)$, where $N(\lambda)$ and $M(\lambda)$ are coprime proper rational matrices with all poles in $\mathbb{C}_g$. We can determine the factors of a LCF factorization from those of a RCF of $G^T(\lambda) = N^T(\lambda)\big(M^T(\lambda)\big)^{-1}$. Therefore, we only discuss algorithms for the computation of RCFs.

The presented algorithms compute RCFs with minimum-degree denominators, by employing a recursive pole dislocation technique (see Sect. 9.1.6), by which all poles of $G(\lambda)$ situated in $\mathbb{C}_b$ are successively moved into $\mathbb{C}_g$, via recursive pole–zero cancellations with elementary denominator factors. To cancel a real pole $\beta \in \mathbb{C}_b$ of $G(\lambda)$, we multiply $G(\lambda)$ from right with an elementary invertible proper factor $\widetilde{M}(\lambda)$ of McMillan degree one, which has β as a zero and $\gamma \in \mathbb{C}_g$ as a pole. For a complex pole, the corresponding $\widetilde{M}(\lambda)$ would contain complex coefficients. Fortunately, we can simultaneously cancel a pair of complex conjugate poles $\beta, \bar{\beta} \in \mathbb{C}_b$ of $G(\lambda)$, by post-multiplying $G(\lambda)$ with an elementary invertible proper factor $\widetilde{M}(\lambda)$ of McMillan degree two, having only real coefficients. This factor has β and $\bar{\beta}$ as zeros and $\gamma_1, \gamma_2 \in \mathbb{C}_g$ as poles (either two real poles or a pair of complex conjugate poles). This pole–zero cancellation technique can be successively employed to dislocate all n_b poles of $G(\lambda)$. The resulting denominator factor can be represented in a product form as

$$M(\lambda) = \widetilde{M}_1(\lambda)\widetilde{M}_2(\lambda)\cdots\widetilde{M}_k(\lambda), \tag{10.57}$$

where each $\widetilde{M}_i(\lambda)$ $(i = 1, \ldots, k)$ is an invertible elementary proper factor with McMillan degree equal to one or two. The computational procedure can be formalized as k successive applications of the updating formula

$$\begin{bmatrix} N_i(\lambda) \\ M_i(\lambda) \end{bmatrix} = \begin{bmatrix} N_{i-1}(\lambda) \\ M_{i-1}(\lambda) \end{bmatrix} \widetilde{M}_i(\lambda), \quad i = 1, \ldots, k\,, \tag{10.58}$$

initialized with $N_0(\lambda) = G(\lambda)$ and $M_0(\lambda) = I_m$. Then, $N(\lambda) = N_k(\lambda)$ and $M(\lambda) = M_k(\lambda)$. By this approach, it is automatically achieved that the resulting $M(\lambda)$ has the least achievable McMillan degree n_b.

We can derive state-space formulas for the efficient implementation of the updating operations in (10.58). Assume $N_{i-1}(\lambda)$ and $M_{i-1}(\lambda)$ have the descriptor realizations

$$\begin{bmatrix} N_{i-1}(\lambda) \\ M_{i-1}(\lambda) \end{bmatrix} = \left[\begin{array}{cc|c} A_{11}-\lambda E_{11} & A_{12}-\lambda E_{12} & B_1 \\ 0 & A_{22}-\lambda E_{22} & B_2 \\ \hline C_{N,1} & C_{N,2} & D_N \\ C_{M,1} & C_{M,2} & D_M \end{array}\right] =: \left[\begin{array}{c|c} \widetilde{A}-\lambda\widetilde{E} & \widetilde{B} \\ \hline \widetilde{C}_N & \widetilde{D}_N \\ \widetilde{C}_M & \widetilde{D}_M \end{array}\right], \quad (10.59)$$

where $\Lambda(A_{22}-\lambda E_{22}) \subset \mathbb{C}_b$. We assume that $A_{22}-\lambda E_{22}$ is a 1×1 pencil in the case when $A_{22}-\lambda E_{22}$ has a real or an infinite eigenvalue, or is a 2×2 pencil, in the case when $A_{22}-\lambda E_{22}$ has a pair of complex conjugate eigenvalues. This form automatically results if the pair $(\widetilde{A}, \widetilde{E})$ is in the specially ordered *generalized real Schur form* (GRSF) determined using **Procedure GSORSF** in Sect. 10.1.4. If $B_2 = 0$, then the eigenvalue(s) of $A_{22}-\lambda E_{22}$ is (are) not controllable, and thus can be removed to obtain realizations of $N_{i-1}(\lambda)$ and $M_{i-1}(\lambda)$ of reduced orders

$$\begin{bmatrix} N_{i-1}(\lambda) \\ M_{i-1}(\lambda) \end{bmatrix} = \left[\begin{array}{c|c} A_{11}-\lambda E_{11} & B_1 \\ \hline C_{N,1} & D_N \\ C_{M,1} & D_M \end{array}\right]. \quad (10.60)$$

After suitable reordering of diagonal blocks of $A_{11}-\lambda E_{11}$ using orthogonal similarity transformations (see Sect. 10.1.4), a new realization $N_{i-1}(\lambda)$ and $M_{i-1}(\lambda)$ can be determined with the matrices again in the form (10.59). If $B_2 \neq 0$, then we have two cases, which are separately discussed in what follows.

If the pencil $A_{22}-\lambda E_{22}$ has finite eigenvalues (i.e., E_{22} is invertible), then the pair $(A_{22}-\lambda E_{22}, B_2)$ is (finite) controllable and there exists F_2 such that the eigenvalues of $A_{22}+B_2F_2-\lambda E_{22}$ can be placed in arbitrary locations in $\mathbb{C}_g$. Assume that such an F_2 has been determined and define the elementary factor $\widetilde{M}_i(\lambda) = (A_{22}+B_2F_2-\lambda E_{22},\ B_2W,\ F_2,\ W)$, where W is chosen to ensure the invertibility of $\widetilde{M}_i(\lambda)$. To compute stable and proper RCFs, the choice $W = I_m$ is always possible. However, alternative choices of W are necessary to ensure, for example, that $\widetilde{M}_i(\lambda)$ is inner. It is easy to check that the updated factors $N_i(\lambda)$ and $M_i(\lambda)$ in (10.58) have the realizations

$$\begin{bmatrix} N_i(\lambda) \\ M_i(\lambda) \end{bmatrix} := \begin{bmatrix} N_{i-1}(\lambda) \\ M_{i-1}(\lambda) \end{bmatrix} \widetilde{M}_i(\lambda) = \left[\begin{array}{cc|c} A_{11}-\lambda E_{11} & A_{12}+B_1F_2-\lambda E_{12} & B_1W \\ 0 & A_{22}+B_2F_2-\lambda E_{22} & B_2W \\ \hline C_{N,1} & C_{N,2}+D_NF_2 & D_NW \\ C_{M,1} & C_{M,2}+D_MF_2 & D_MW \end{array}\right].$$

If we denote $\widetilde{F} = \begin{bmatrix} 0 & F_2 \end{bmatrix}$, then the above relations lead to the following updating formulas:

$$
\begin{aligned}
\widetilde{A} &\leftarrow \widetilde{A} + \widetilde{B}\widetilde{F}, \\
\widetilde{B} &\leftarrow \widetilde{B}W, \\
\widetilde{C}_N &\leftarrow \widetilde{C}_N + \widetilde{D}_N\widetilde{F}, \\
\widetilde{C}_M &\leftarrow \widetilde{C}_M + \widetilde{D}_M\widetilde{F}, \\
\widetilde{D}_N &\leftarrow \widetilde{D}_N W, \\
\widetilde{D}_M &\leftarrow \widetilde{D}_M W.
\end{aligned} \tag{10.61}
$$

If the 1×1 pencil $A_{22} - \lambda E_{22}$ has an infinite eigenvalue (i.e., $E_{22} = 0$), then we choose the elementary factor $\widetilde{M}_i(\lambda) = (\gamma - \lambda,\ B_2,\ F_2,\ W)$, where γ is an arbitrary real eigenvalue in $\mathbb{C}_g$, W is a projection matrix chosen such $B_2 W = 0$ and rank $\begin{bmatrix} B_2 \\ W \end{bmatrix} = m$, and F_2 has been chosen such that $B_2 F_2 = -A_{22}$ and rank$[\, F_2 \ W \,] = m$ (the rank conditions guarantee the invertibility of $\widetilde{M}_i(\lambda)$). Straightforward choices of F_2 and W are, for example, $F_2 = -B_2^T(B_2B_2^T)^{-1}A_{22}$ and $W = I_m - B_2^T(B_2B_2^T)^{-1}B_2$. By this choice of $\widetilde{M}_i(\lambda)$, we made the infinite eigenvalue in the realization of the updated factors $N_i(\lambda)$ and $M_i(\lambda)$ simple, and after its elimination, we obtain the realizations

$$
\begin{bmatrix} N_i(\lambda) \\ M_i(\lambda) \end{bmatrix} := \begin{bmatrix} N_{i-1}(\lambda) \\ M_{i-1}(\lambda) \end{bmatrix} \widetilde{M}_i(\lambda) = \left[\begin{array}{cc|c} A_{11} - \lambda E_{11} & A_{12} + B_1F_2 - \lambda E_{12} & B_1 W \\ 0 & \gamma - \lambda & B_2 \\ \hline C_{N,1} & C_{N,2} + D_N F_2 & D_N W \\ C_{M,1} & C_{M,2} + D_M F_2 & D_M W \end{array}\right].
$$

The above relations lead to the following updating formulas:

$$
\begin{aligned}
\widetilde{A} &\leftarrow \begin{bmatrix} A_{11} & A_{12} + B_1F_2 \\ 0 & \gamma \end{bmatrix}, \\
\widetilde{E} &\leftarrow \begin{bmatrix} E_{11} & E_{12} \\ 0 & 1 \end{bmatrix}, \\
\widetilde{B} &\leftarrow \begin{bmatrix} B_1 W \\ B_2 \end{bmatrix}, \\
\widetilde{C}_N &\leftarrow \begin{bmatrix} C_{N,1} & C_{N,2} + D_N F_2 \end{bmatrix}, \\
\widetilde{C}_M &\leftarrow \begin{bmatrix} C_{M,1} & C_{M,2} + D_M F_2 \end{bmatrix}, \\
\widetilde{D}_N &\leftarrow D_N W, \\
\widetilde{D}_M &\leftarrow D_M W.
\end{aligned} \tag{10.62}
$$

The updating techniques relying on the formulas (10.61) and (10.62) ensure that, if the original pair $(\widetilde{A}, \widetilde{E})$ was in a GRSF, then the updated pair will have a similar form, possibly with $\widetilde{A} - \lambda\widetilde{E}$ having a 2×2 trailing block which corresponds to two real generalized eigenvalues (to recover the GRSF, such a block can be further split into two 1×1 blocks using an orthogonal similarity transformation). By reordering the diagonal blocks in the GRSF of the updated pair $(\widetilde{A}, \widetilde{E})$, we can bring in the trailing position new blocks whose generalized eigenvalues lie in $\mathbb{C}_b$. The described eigenvalue dislocation process is repeated until all eigenvalues are moved into $\mathbb{C}_g$, using suitably chosen elementary denominators.

The following procedure computes a proper and stable RCF of an arbitrary rational TFM $G(\lambda)$ with respect to a given partition $\mathbb{C} = \mathbb{C}_b \cup \mathbb{C}_g$ as $G(\lambda) = N(\lambda)M^{-1}(\lambda)$, where the resulting factors $N(\lambda)$ and $M(\lambda)$ have the realizations $N(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}_N, \widetilde{D}_N)$ and $M(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}_M, \widetilde{D}_M)$.

Procedure GRCF: Generalized stable right coprime factorization

Inputs : $G(\lambda) = (A - \lambda E, B, C, D)$ with $A, E \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$; $\mathbb{C}_g$ and $\mathbb{C}_b$, such that $\mathbb{C} = \mathbb{C}_b \cup \mathbb{C}_g$, $\mathbb{C}_b \cap \mathbb{C}_g = \emptyset$

Outputs: $N(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}_N, \widetilde{D}_N)$ and $M(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}_M, \widetilde{D}_M)$ such that $G(\lambda) = N(\lambda)M^{-1}(\lambda)$, all finite eigenvalues of $\widetilde{A} - \lambda\widetilde{E}$ are in $\mathbb{C}_g$ and all infinite eigenvalues of $\widetilde{A} - \lambda\widetilde{E}$ are simple

1) Compute, using **Procedure GSORSF**, the orthogonal matrices Q and Z to reduce the pair (A, E) to the special ordered GRSF

$$\widetilde{A} := Q^T AZ = \begin{bmatrix} A_\infty & * & * \\ 0 & A_g & * \\ 0 & 0 & A_b \end{bmatrix}, \qquad \widetilde{E} := Q^T EZ = \begin{bmatrix} 0 & * & * \\ 0 & E_g & * \\ 0 & 0 & E_b \end{bmatrix},$$

where $A_\infty \in \mathbb{R}^{(n-r)\times(n-r)}$ is invertible and upper triangular, with $r = \operatorname{rank} E$, $\Lambda(A_g - \lambda E_g) \subset \mathbb{C}_g$ with $A_g, E_g \in \mathbb{R}^{n_g\times n_g}$ and $\Lambda(A_b - \lambda E_b) \subset \mathbb{C}_b$ with $A_b, E_b \in \mathbb{R}^{n_b\times n_b}$. Compute $\widetilde{B} := Q^T B$, $\widetilde{C}_N := CZ$, $\widetilde{C}_M = 0$, $\widetilde{D}_N = D$, $\widetilde{D}_M = I_m$. Set $q = n - n_b$.
2) If $q = n$, **Exit**.
3) Let (A_{22}, E_{22}) be the last $k \times k$ diagonal blocks of the GRSF of $(\widetilde{A}, \widetilde{E})$ (with $k=1$ or $k=2$) and let B_2 be the $k \times m$ matrix formed from the last k rows of $\widetilde{B}$. If $\|B_2\| \leq \varepsilon$ (a given tolerance), then remove the parts corresponding to the uncontrollable eigenvalues $\Lambda(A_{22} - \lambda E_{22})$:
$\widetilde{A} \leftarrow \widetilde{A}(1:n-k, 1:n-k)$, $\widetilde{E} \leftarrow \widetilde{E}(1:n-k, 1:n-k)$,
$\widetilde{B} \leftarrow \widetilde{B}(1:n-k, 1:m)$, $\widetilde{C}_N \leftarrow \widetilde{C}_N(1:p, 1:n-k)$,
$\widetilde{C}_M \leftarrow \widetilde{C}_M(1:p, 1:n-k)$; update $n \leftarrow n-k$, $q \leftarrow q-k$ and go to Step 2).
4) If $E_{22} \neq 0$, determine F_2 such that $\Lambda(A_{22} + B_2F_2 - \lambda E_{22}) \subset \mathbb{C}_g$. Set $\widetilde{F} = [\,0 \;\; F_2\,]$ and compute $\widetilde{A} \leftarrow \widetilde{A} + \widetilde{B}\widetilde{F}$, $\widetilde{C}_N \leftarrow \widetilde{C}_N + \widetilde{D}_N\widetilde{F}$, $\widetilde{C}_M \leftarrow \widetilde{C}_M + \widetilde{D}_M\widetilde{F}$.
5) If $E_{22} = 0$, compute $F_2 = -B_2^T(B_2B_2^T)^{-1}A_{22}$ and $W = I_m - B_2^T(B_2B_2^T)^{-1}B_2$. Choose $\gamma \in \mathbb{C}_g$ and update $\widetilde{A}$, $\widetilde{E}$, $\widetilde{B}$, $\widetilde{C}_N$, $\widetilde{D}_N$, $\widetilde{C}_M$ and $\widetilde{D}_M$ using (10.62).
6) Compute the orthogonal matrices $\widetilde{Q}$ and $\widetilde{Z}$ to move the last blocks of $(\widetilde{A}, \widetilde{E})$ to positions $(q+1, q+1)$ by interchanging the diagonal blocks of the GRSF. Compute $\widetilde{A} \leftarrow \widetilde{Q}^T\widetilde{A}\widetilde{Z}$, $\widetilde{E} \leftarrow \widetilde{Q}^T\widetilde{E}\widetilde{Z}$, $\widetilde{B} \leftarrow \widetilde{Q}^T\widetilde{B}$, $\widetilde{C}_N \leftarrow \widetilde{C}_N\widetilde{Z}$, $\widetilde{C}_M \leftarrow \widetilde{C}_M\widetilde{Z}$. Put $q \leftarrow q + k$ and go to Step 2).

This algorithm is completely general, being applicable regardless the original descriptor realization is $\mathbb{C}_b$-stabilizable or not, is infinite controllable or not. The resulting pair $(\widetilde{A}, \widetilde{E})$ is in a special GRSF with $n - r$ simple infinite eigenvalues in the leading $n - r$ positions (no such block exists if E is invertible). A minimal realization of the least McMillan degree denominator $M(\lambda)$ can be easily determined. The resulting $\widetilde{C}_M$ has always the form

$$\widetilde{C}_M = [\,0 \;\; \widetilde{C}_{M,2}\,], \tag{10.63}$$

where the number of columns of $\widetilde{C}_{M,2}$ is equal to the number of controllable generalized eigenvalues of the pair (A, E) lying in $\mathbb{C}_b$. By partitioning accordingly the resulting $\widetilde{E}$, $\widetilde{A}$ and $\widetilde{B}$

$$\widetilde{A} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad \widetilde{E} = \begin{bmatrix} E_{11} & E_{12} \\ 0 & E_{22} \end{bmatrix}, \quad \widetilde{B} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \tag{10.64}$$

then $(A_{22} - \lambda E_{22}, B_2, \widetilde{C}_{M,2}, \widetilde{D}_M)$ is a minimal descriptor system realization of $M(\lambda)$. Notice however that the order of the minimal realization of $M(\lambda)$ can be higher than the least possible McMillan degree if some eigenvalues of $A - \lambda E$ in $\mathbb{C}_b$ are controllable but not observable.

The **Procedure GRCF** can be interpreted as an extension of the generalized pole assignment algorithm of [127], which generalizes the pole assignment algorithm of [123] for standard systems. The roundoff error analysis of this latter algorithm [124] revealed that if each gain matrix F_2 computed at Step 4) or Step 5) satisfies $\|F_2\| \leq \kappa \|A\|/\|B\|$, with κ having moderate values (say $\kappa < 100$), then the standard pole assignment algorithm is numerically backward stable. This condition is also applicable in our case, because it is independent of the presence of E. We note however that, unfortunately, this condition cannot be always fulfilled if large gains are necessary to stabilize the system. This can arise either if the unstable poles are too "far" from the stable region or if these poles are weekly controllable. Nevertheless, the **Procedure GRCF** can be considered a reliable algorithm, since the above condition can be checked at each computational step and therefore the potential loss of numerical stability can be easily detected.

A similar recursive procedure can be developed to compute RCFs with inner denominators. In this case, we use the partition of the complex plane with $\mathbb{C}_g = \mathbb{C}_s$ and $\mathbb{C}_b = \mathbb{C} \setminus \mathbb{C}_s$. A necessary and sufficient condition for the existence of such a factorization is that $G(\lambda)$ has no poles in $\partial\mathbb{C}_s$ (the boundary of $\mathbb{C}_s$). In the continuous-time case, this means that the pencil $A - sE$ has no finite eigenvalues on the imaginary axis and all infinite eigenvalues of $A - sE$ are simple. In the discrete-time case, $A - zE$ has no eigenvalues on the unit circle centred in the origin. However, for the sake of generality, $G(z)$ can be improper, thus $A - zE$ may have multiple infinite eigenvalues.

For the computation of the RCF with inner denominators we use a similar recursive pole dislocation technique as in the case of a general RCF, using elementary inner factors. The denominator factor results in the factored form (10.57), where each $\widetilde{M}_i(\lambda)$ $(i = 1, \ldots, k)$ is an elementary inner factor with McMillan degree equal to one or two. These factors are used to reflect the unstable poles of $G(\lambda)$ to stable symmetric positions with respect to the imaginary axis, in the case of a continuous-time system, or with respect to the unit circle in the origin, in the case of a discrete-time system.

In what follows, we give the formulas to determine the elementary inner factors to be used in (10.57) and derive appropriate updating formulas of the factors. We assume $N_{i-1}(\lambda)$ and $M_{i-1}(\lambda)$ have the descriptor realizations in (10.59) and $B_2 \neq 0$ (otherwise the uncontrollable part $A_{22} - \lambda E_{22}$ can be removed from the realization, see (10.60)). In the case when $A_{22} - \lambda E_{22}$ has finite eigenvalues (i.e., E_{22} is invertible) we choose the elementary inner factor as $\widetilde{M}_i(\lambda) = (A_{22} + B_2F_2 - \lambda E_{22},\ B_2W,\ F_2,\ W)$. The updating formulas for this case are the same as those employed in **Procedure GRCF** given in (10.61). For the computation of F_2 and W we have the following results.

Lemma 10.1 *Let* $(A_{22} - sE_{22}, B_2)$ *be a controllable continuous-time descriptor pair with* E_{22} *invertible and* $\Lambda(A_{22} - sE_{22}) \subset \mathbb{C}_u$. *Then the elementary denominator factor* $\widetilde{M}_i(s) = (A_{22} + B_2F_2 - sE_{22},\ B_2W,\ F_2,\ W)$ *is inner by choosing* F_2 *and* W *as*

$$\begin{array}{l} A_{22}YE_{22}^T + E_{22}YA_{22}^T - B_2B_2^T = 0, \\ F_2 = -B_2^T(YE_{22}^T)^{-1}, \quad W = I_m. \end{array}$$

Lemma 10.2 *Let* $(A_{22} - zE_{22}, B_2)$ *be a controllable discrete-time descriptor pair with* E_{22} *invertible and* $\Lambda(A_{22} - zE_{22}) \subset \mathbb{C}_u$. *Then the elementary denominator factor* $\widetilde{M}_i(z) = (A_{22} + B_2F_2 - zE_{22},\ B_2W,\ F_2,\ W)$ *is inner by choosing* F_2 *and* W *as*

$$\begin{array}{l} A_{22}YA_{22}^T - B_2B_2^T = E_{22}YE_{22}^T, \\ F_2 = -B_2^T(YA_{22}^T)^{-1}, \\ W^T(I + B_2^T(E_{22}YE_{22}^T)^{-1}B_2)W = I. \end{array}$$

If the 1×1 pencil $A_{22} - zE_{22}$ has an infinite eigenvalue (i.e., $E_{22} = 0$), then we have the following result for the choice of the elementary inner factor.

Lemma 10.3 *Let* $(A_{22} - zE_{22}, B_2)$ *be an infinite controllable discrete-time descriptor pair with* $E_{22} = 0$, *and* A_{22} *nonzero. Then the elementary denominator factor* $\widetilde{M}_i(z) = (0 + zA_{22},\ B_2,\ F_2,\ W)$ *is inner by choosing* F_2 *and* W *as*

$$\begin{array}{l} F_2 = -B_2^T(B_2B_2^T)^{-1}A_{22}, \\ W = I - B_2^T(B_2B_2^T)^{-1}B_2. \end{array}$$

By this choice of $\widetilde{M}_i(z)$, we made the infinite eigenvalue in the realization of the updated factors $N_i(z)$ and $M_i(z)$ simple, and after its elimination, we obtain the realizations

$$\begin{bmatrix} N_i(z) \\ M_i(z) \end{bmatrix} := \begin{bmatrix} N_{i-1}(z) \\ M_{i-1}(z) \end{bmatrix} \widetilde{M}_i(z) = \left[\begin{array}{cc|c} A_{11} - zE_{11} & A_{12} + B_1F_2 - zE_{12} & B_1W \\ 0 & zA_{22} & B_2 \\ \hline C_{N,1} & C_{N,2} + D_NF_2 & D_NW \\ C_{M,1} & C_{M,2} + D_MF_2 & D_MW \end{array}\right],$$

The above relations lead to the following updating formulas:

$$
\begin{aligned}
\widetilde{A} &\leftarrow \begin{bmatrix} A_{11} & A_{12}+B_1F_2 \\ 0 & 0 \end{bmatrix}, \quad \widetilde{E} \leftarrow \begin{bmatrix} E_{11} & E_{12} \\ 0 & -A_{22} \end{bmatrix}, \quad \widetilde{B} \leftarrow \begin{bmatrix} B_1 W \\ B_2 \end{bmatrix}, \\
\widetilde{C}_N &\leftarrow \big[\, C_{N,1} \;\; C_{N,2}+D_N F_2 \,\big], \quad \widetilde{C}_M \leftarrow \big[\, C_{M,1} \;\; C_{M,2}+D_M F_2 \,\big], \\
\widetilde{D}_N &\leftarrow D_N W, \quad \widetilde{D}_M \leftarrow D_M W.
\end{aligned}
\tag{10.65}
$$

The following procedure computes a stable RCF with inner denominator of a rational TFM $G(\lambda)$, without poles in $\partial\mathbb{C}_s$, as $G(\lambda) = N(\lambda)M^{-1}(\lambda)$, where the resulting factors $N(\lambda)$ and $M(\lambda)$ have the realizations $N(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}_N, \widetilde{D}_N)$ and $M(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}_M, \widetilde{D}_M)$.

Procedure GRCFID: Generalized RCF with inner denominator

Inputs : $G(\lambda) = (A - \lambda E, B, C, D)$ with $A, E \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$

Outputs: $N(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}_N, \widetilde{D}_N)$ and $M(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}_M, \widetilde{D}_M)$ such that $G(\lambda) = N(\lambda)M^{-1}(\lambda)$, $M(\lambda)$ is inner, all finite eigenvalues of $\widetilde{A} - \lambda\widetilde{E}$ are in $\mathbb{C}_s$ and all infinite eigenvalues of $\widetilde{A} - \lambda\widetilde{E}$ are simple

1) Compute using **Procedure GSORSF**, the orthogonal matrices Q and Z to reduce the pair (A, E) to the special ordered GRSF

$$
\widetilde{A} := Q^T A Z = \begin{bmatrix} A_\infty & * & * \\ 0 & A_s & * \\ 0 & 0 & A_u \end{bmatrix}, \qquad \widetilde{E} := Q^T E Z = \begin{bmatrix} 0 & * & * \\ 0 & E_s & * \\ 0 & 0 & E_u \end{bmatrix},
$$

where $A_\infty \in \mathbb{R}^{(n-r)\times(n-r)}$ is invertible and upper triangular, with $r = \operatorname{rank} E$, $\Lambda(A_s - \lambda E_s) \subset \mathbb{C}_s$ with $A_s, E_s \in \mathbb{R}^{n_s\times n_s}$ and $\Lambda(A_u - \lambda E_u) \subset \mathbb{C}_u$ with $A_u, E_u \in \mathbb{R}^{n_u\times n_u}$. Compute $\widetilde{B} := Q^T B$, $\widetilde{C}_N := CZ$, $\widetilde{C}_M = 0$, $\widetilde{D}_N = D$, $\widetilde{D}_M = I_m$. Set $q = n - n_u$.
2) If $q = n$, **Exit**.
3) Let (A_{22}, E_{22}) be the last $k \times k$ diagonal blocks of the GRSF of $(\widetilde{A}, \widetilde{E})$ (with $k = 1$ or $k = 2$) and let B_2 be the matrix formed from the last k rows of $\widetilde{B}$. If $\|B_2\| \le \varepsilon$ (a given tolerance), then remove the parts corresponding to the uncontrollable eigenvalues $\Lambda(A_{22} - \lambda E_{22})$: $\widetilde{A} \leftarrow \widetilde{A}(1:n-k, 1:n-k)$, $\widetilde{E} \leftarrow \widetilde{E}(1:n-k, 1:n-k)$, $\widetilde{B} \leftarrow \widetilde{B}(1:n-k, 1:m)$, $\widetilde{C}_N \leftarrow \widetilde{C}_N(1:p, 1:n-k)$, $\widetilde{C}_M \leftarrow \widetilde{C}_M(1:p, 1:n-k)$; update $n \leftarrow n-k$, $q \leftarrow q-k$ and go to Step 2).
4) If $E_{22} \neq 0$, compute F_2 and W according to Lemma 10.1 in the continuous-time case or according to Lemma 10.2 in the discrete-time case. Set $\widetilde{F} = [\,0 \;\; F_2\,]$ and update $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}_N$, $\widetilde{D}_N$, $\widetilde{C}_M$ and $\widetilde{D}_M$ using (10.61).
5) If $E_{22} = 0$, compute $F_2 = -B_2^T(B_2B_2^T)^{-1}A_{22}$ and $W = I_m - B_2^T(B_2B_2^T)^{-1}B_2$, and update $\widetilde{A}$, $\widetilde{E}$, $\widetilde{B}$, $\widetilde{C}_N$, $\widetilde{D}_N$, $\widetilde{C}_M$ and $\widetilde{D}_M$ using (10.65).
6) Compute the orthogonal matrices $\widetilde{Q}$ and $\widetilde{Z}$ to move the last blocks of $(\widetilde{A}, \widetilde{E})$ to positions $(q+1, q+1)$ by interchanging the diagonal blocks of the GRSF. Compute $\widetilde{A} \leftarrow \widetilde{Q}^T\widetilde{A}\widetilde{Z}$, $\widetilde{E} \leftarrow \widetilde{Q}^T\widetilde{E}\widetilde{Z}$, $\widetilde{B} \leftarrow \widetilde{Q}^T\widetilde{B}$, $\widetilde{C}_N \leftarrow \widetilde{C}_N\widetilde{Z}$, $\widetilde{C}_M \leftarrow \widetilde{C}_M\widetilde{Z}$. Put $q \leftarrow q+k$ and go to Step 2).

The resulting inner factor $M(\lambda)$ has least McMillan degree, only if all unstable generalized eigenvalues of the pair (E, A) are observable. A minimal realization of $M(\lambda)$ can be explicitly determined as $(A_{22} - \lambda E_{22}, B_2, \widetilde{C}_{M,2}, \widetilde{D}_M)$, where the matrices of the realization are defined in (10.64) and (10.63).

The numerical properties of **Procedure GRCFID** are similar to those of **Procedure GRCF**, as long as the matrix gains $\|F_2\|$ at Steps 4) and 5) are reasonably small. However, this condition for numerical reliability may not always be fulfilled due to the lack of freedom in assigning the poles. Recall that the unstable poles are reflected in symmetrical position with respect to $\partial\mathbb{C}_s$, and this may occasionally require large gains.

10.3.6 Inner–Outer Factorization

In the light of the needs of the synthesis algorithms presented in Chap. 5, we discuss the computation of the inner–outer factorization of a particular $p \times m$ rational matrix $G(\lambda)$, namely which is proper and has full column rank. Assume that $G(\lambda)$ has an n-th order descriptor system realization

$$G(\lambda) = \left[\begin{array}{c|c} A - \lambda E & B \\ \hline C & D \end{array}\right], \tag{10.66}$$

with E an invertible $n \times n$ matrix. Consider the disjunct partition of the complex plane as $\mathbb{C} = \mathbb{C}_u \cup \overline{\mathbb{C}}_s$. We discuss the computation of the inner–outer factorization of $G(\lambda)$ either in the compact form

$$G(\lambda) = G_{i,1}(\lambda)G_o(\lambda)\,, \tag{10.67}$$

or in the extended form

$$G(\lambda) = \left[\, G_{i,1}(\lambda) \;\; G_{i,2}(\lambda) \,\right] \begin{bmatrix} G_o(\lambda) \\ 0 \end{bmatrix} = G_i(\lambda) \begin{bmatrix} G_o(\lambda) \\ 0 \end{bmatrix}, \tag{10.68}$$

where $G_i(\lambda) := \left[\, G_{i,1}(\lambda) \;\; G_{i,2}(\lambda) \,\right]$ is a square inner TFM (i.e., with $G_{i,1}(\lambda)$ inner too), and $G_o(\lambda)$ is an invertible *quasi*-outer TFM, having all zeros in $\overline{\mathbb{C}}_s$. The stability of $G_o(\lambda)$ is ensured, provided $G(\lambda)$ is stable. The component $G_{i,2}(\lambda)$ is a complementary inner factor (also called an "orthogonal" complement of $G_{i,1}(\lambda)$) (see Sect. 9.1.8).

For the computation of inner–outer factorization of $G(\lambda)$, a special reduced form of the system matrix will be instrumental.

Proposition 10.5 *Let $G(\lambda)$ be a $p \times m$ proper rational matrix of full column rank with a stabilizable realization given in (10.66). Then, there exist orthogonal matrices U and Z such that*

$$\begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} \left[\begin{array}{cc} A - \lambda E & B \\ \hline C & D \end{array}\right] Z = \left[\begin{array}{ccc} A_s - \lambda E_s & * & * \\ 0 & A_{u\ell} - \lambda E_{u\ell} & B_{u\ell} \\ \hline 0 & C_{u\ell} & D_{u\ell} \end{array}\right], \tag{10.69}$$

where

(a) *The regular pencil* $A_s - \lambda E_s$ *contains the zeros of* $G(\lambda)$ *in* $\overline{\mathbb{C}}_s$*;*
(b) *The descriptor system defined by*

$$G_{u\ell}(\lambda) = \left[\begin{array}{c|c} A_{u\ell} - \lambda E_{u\ell} & B_{u\ell} \\ \hline C_{u\ell} & D_{u\ell} \end{array}\right] \tag{10.70}$$

is proper, with the $n_\ell \times n_\ell$ *matrix* $E_{u\ell}$ *invertible, is stabilizable, has full column rank and has only zeros in* $\mathbb{C}_u$*.*

Proof This proposition is a simplified version of a slight variation of Theorem 3.1 of [97] combined with Theorem 2.2 in [94], where constructive proofs are also given to determine the orthogonal matrices U and Z, as well as the condensed form (10.69), using numerically stable computational algorithms. For convenience, we describe the main computational steps of this reduction for the considered particular case. Let us denote the initial system pencil as

$$S_0(\lambda) := \left[\begin{array}{cc} A - \lambda E & B \\ \hline C & D \end{array}\right]$$

and observe that $S_0(\lambda)$ has full column rank $n+m$ and furthermore $[\, A - \lambda E \;\; B\,]$ has full row rank for all $\lambda \in \mathbb{C}_u$. The reduction algorithm has three computational steps, which are presented in what follows.

(1) Compute orthogonal Z_1 such that

$$[\; C \;\; D \;] Z_1 = [\; 0 \;\; C_2^{(1)} \;] \,,$$

with $C_2^{(1)}$ of full column rank and define

$$S_1(\lambda) := S_0(\lambda) Z_1 = \left[\begin{array}{cc} A_{11}^{(1)} - \lambda E_{11}^{(1)} & A_{12}^{(1)} - \lambda E_{12}^{(1)} \\ \hline 0 & C_2^{(1)} \end{array}\right].$$

Since E is invertible, it follows that $[\, E_{11}^{(1)} \;\; E_{12}^{(1)}\,]$ has full row rank n.

(2) Compute orthogonal U and Z_2 to reduce the pencil $A_{11}^{(1)} - \lambda E_{11}^{(1)}$ to a Kronecker-like form (see Sect. 10.1.6)

$$U\big(A_{11}^{(1)} - \lambda E_{11}^{(1)}\big) Z_2 = \begin{bmatrix} A_s - \lambda E_s & * & * \\ 0 & A_u - \lambda E_u & * \\ 0 & 0 & \overline{A}_\ell - \lambda \overline{E}_\ell \end{bmatrix} := \begin{bmatrix} A_s - \lambda E_s & * \\ 0 & \overline{A}_{u\ell} - \lambda \overline{E}_{u\ell} \end{bmatrix},$$

where $\Lambda(A_s - \lambda E_s) \subset \overline{\mathbb{C}}_s$, $\Lambda(A_u - \lambda E_u) \subset \mathbb{C}_u$, and $\overline{A}_\ell - \lambda \overline{E}_\ell$ has full column rank for all $\lambda \in \mathbb{C}$. Define

$$S_2(\lambda) := \operatorname{diag}(U, I) S_1(\lambda) \operatorname{diag}(Z_2, I) = \left[\begin{array}{ccc} A_s - \lambda E_s & * & * \\ 0 & \overline{A}_{u\ell} - \lambda \overline{E}_{u\ell} & \overline{B}_{u\ell} - \lambda \overline{F}_{u\ell} \\ \hline 0 & 0 & C_2^{(1)} \end{array}\right].$$

It easy to show that $[\,\overline{A}_{u\ell}-\lambda\overline{E}_{u\ell}\;\; \overline{B}_{u\ell}-\lambda\overline{F}_{u\ell}\,]$ has full row rank for all $\lambda\in\mathbb{C}_u$ and also $[\,\overline{E}_{u\ell}\;\;\overline{F}_{u\ell}\,]$ has full row rank.

(3) Compute orthogonal Z_3 such that

$$\left[\,\overline{E}_{u\ell}\;\overline{F}_{u\ell}\,\right] Z_3 = \begin{bmatrix} E_{u\ell} & 0 \\ C_{u\ell} & D_{u\ell} \end{bmatrix},$$

with $E_{u\ell}$ invertible. Define

$$S_3(\lambda) = S_2(\lambda)\,\mathrm{diag}(I, Z_3) = \left[\begin{array}{ccc} A_s-\lambda E_s & * & * \\ 0 & A_{u\ell}-\lambda E_{u\ell} & B_{u\ell} \\ \hline 0 & C_{u\ell} & D_{u\ell} \end{array}\right].$$

The properties (a) and (b) follow immediately from the above properties of the blocks of the reduced final form. The overall transformation matrix Z is defined as

$$Z = Z_1\,\mathrm{diag}(Z_2, I)\,\mathrm{diag}(I, Z_3)\,.$$

■

Remark 10.7 This proposition extracts from the original system (10.66) a proper system (10.70) which has a standard inner–outer factorization. It can be shown that there exists an invertible $G_r(\lambda)$ with zeros only in $\overline{\mathbb{C}}_s$ such that

$$G_{u\ell}(\lambda)G_r(\lambda) = G(\lambda)\,.$$

It follows that $G_{u\ell}(\lambda)$ and $G(\lambda)$ have the same inner factor. Assume $G_i(\lambda)$ is a square inner TFM such that

$$G_{u\ell}(\lambda) = G_i(\lambda)\begin{bmatrix} G_{o,1}(\lambda) \\ 0 \end{bmatrix}$$

is an extended standard inner–outer factorization, where $G_{o,1}(\lambda)$ has only zeros in $\mathbb{C}_s$. Then with $G_o(\lambda) := G_{o,1}(\lambda)G_r(\lambda)$ we immediately obtain an inner–quasi-outer factorization of $G(\lambda)$ in the form (10.68). □

We discuss now the computation of the inner–outer factorization separately for the continuous-time and discrete-times cases.

In the continuous-time case, we can further refine the reduced form (10.69) by observing that $D_{u\ell}$ is full column rank (otherwise $G_{u\ell}(s)$ would have infinite zeros). Therefore, we can compress $D_{u\ell}$ to a full row rank matrix using an orthogonal transformation matrix V, such that

$$V^T D_{u\ell} = \begin{bmatrix} D_\ell \\ 0 \end{bmatrix}, \qquad C_\ell := V^T C_{u\ell} = \begin{bmatrix} C_{\ell,1} \\ C_{\ell,2} \end{bmatrix}, \tag{10.71}$$

where D_ℓ is invertible. With this, we have the following result from [97].

Proposition 10.6 *Let $G(s)$ be a $p \times m$ proper full column rank rational matrix with a stabilizable realization (10.66), let U and Z be orthogonal transformation matrices such that (10.69) holds and let V be an orthogonal transformation matrix which compresses $D_{u\ell}$ as in (10.71). Let X_s be the positive definite stabilizing solution of the generalized continuous-time Riccati equation (GCARE)*

$$
\begin{aligned}
&A_{u\ell}^T X E_{u\ell} + E_{u\ell}^T X A_{u\ell} - (E_{u\ell}^T X B_{u\ell} + C_{u\ell}^T D_{u\ell}) \\
&\times (D_{u\ell}^T D_{u\ell})^{-1} (B_{u\ell}^T X E_{u\ell} + D_{u\ell}^T C_{u\ell}) + C_{u\ell}^T C_{u\ell} = 0
\end{aligned} \tag{10.72}
$$

and let F_s be the corresponding stabilizing feedback

$$
F_s = -R^{-1}(B_{u\ell}^T X_s E_{u\ell} + D_{u\ell}^T C_{u\ell}),
$$

with $R := D_{u\ell}^T D_{u\ell} > 0$. Then, the factors of the inner–quasi-outer factorization (10.68) are given by

$$
G_i(s) = \begin{bmatrix} G_{i,1}(s) & G_{i,2}(s) \end{bmatrix} = V \left[\begin{array}{c|cc} A_{u\ell} + B_{u\ell}F_s - sE_{u\ell} & B_{u\ell}D_\ell^{-1} & -X_s^{-1}E_{u\ell}^{-T}C_{\ell,2}^T \\ \hline C_{\ell,1} + D_\ell F_s & I & 0 \\ C_{\ell,2} & 0 & I \end{array} \right]
$$

and

$$
G_o(s) = \left[\begin{array}{c|c} A - sE & B \\ \hline \widetilde{C} & \widetilde{D} \end{array} \right],
$$

where $\begin{bmatrix} \widetilde{C} & \widetilde{D} \end{bmatrix} := R^{1/2} \begin{bmatrix} 0 & F_s & I \end{bmatrix} Z^T$.

In the discrete-time case, we have the following result from [94].

Proposition 10.7 *Let $G(z)$ be a $p \times m$ proper full column rank rational matrix with a stabilizable realization (10.66), and let U and Z be orthogonal transformation matrices such that (10.69) holds. Let X_s be the stabilizing solution of the generalized discrete-time Riccati equation (GDARE)*

$$
\begin{aligned}
&A_{u\ell}^T X A_{u\ell} - E_{u\ell}^T X E_{u\ell} - (A_{u\ell}^T X B_{u\ell} + C_{u\ell}^T D_{u\ell}) \\
&\times (D_{u\ell}^T D_{u\ell} + B_{u\ell}^T X B_{u\ell})^{-1} (B_{u\ell}^T X A_{u\ell} + D_{u\ell}^T C_{u\ell}) + C_{u\ell}^T C_{u\ell} = 0
\end{aligned} \tag{10.73}
$$

and let F_s be the corresponding stabilizing feedback

$$
F_s = -R^{-1}(B_{u\ell}^T X_s A_{u\ell} + D_{u\ell}^T C_{u\ell}),
$$

with $R := D_{u\ell}^T D_{u\ell} + B_{u\ell}^T X_s B_{u\ell} > 0$. Then, the factors of the inner–quasi-outer factorization (10.67) are given by

$$G_{i,1}(z) = \left[\begin{array}{c|c} A_{u\ell} + B_{u\ell}F_s - zE_{u\ell} & B_{u\ell}R^{-\frac{1}{2}} \\ \hline C_{u\ell} + D_{u\ell}F_s & D_{u\ell}R^{-\frac{1}{2}} \end{array}\right]$$

and

$$G_o(z) = \left[\begin{array}{c|c} A - zE & B \\ \hline \widetilde{C} & \widetilde{D} \end{array}\right],$$

where $\left[\, \widetilde{C}\ \widetilde{D} \,\right] := R^{1/2} \left[\, 0\ F_s\ I \,\right] Z^T$.

Remark 10.8 The complementary inner factor $G_{i,2}(z)$ can be computed in the form [164]

$$G_{i,2}(z) = \left[\begin{array}{c|c} A_{u\ell} + B_{u\ell}F_s - zE_{u\ell} & Y \\ \hline C_{u\ell} + D_{u\ell}F_s & W \end{array}\right],$$

where Y and W satisfy

$$\begin{aligned} A_{u\ell}^T X_s Y + C_{u\ell}^T W &= 0\,, \\ B_{u\ell}^T X_s Y + D_{u\ell}^T W &= 0\,, \\ W^T W + Y^T X_s Y &= I\,. \end{aligned}$$

To compute Y and W we can determine first an orthogonal nullspace basis $\begin{bmatrix} \widetilde{Y} \\ \widetilde{W} \end{bmatrix}$ satisfying

$$\begin{bmatrix} A_{u\ell}^T X_s & C_{u\ell}^T \\ B_{u\ell}^T X_s & D_{u\ell}^T \end{bmatrix} \begin{bmatrix} \widetilde{Y} \\ \widetilde{W} \end{bmatrix} = 0$$

and then compute $Y = \widetilde{Y}L^{-1}$ and $W = \widetilde{W}L^{-1}$, where L is a Cholesky factor satisfying

$$\widetilde{W}^T\widetilde{W} + \widetilde{Y}^T X_s \widetilde{Y} = L^T L\,.$$

A numerically reliable way to compute the orthogonal nullspace is via the singular value decomposition

$$\begin{bmatrix} A_{u\ell}^T X_s & C_{u\ell}^T \\ B_{u\ell}^T X_s & D_{u\ell}^T \end{bmatrix} = \left[\, U_1\ U_2 \,\right] \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \left[\, V_1\ V_2 \,\right]^T,$$

where Σ is an invertible $k \times k$ diagonal matrix and $\left[\, U_1\ U_2 \,\right]$ and $\left[\, V_1\ V_2 \,\right]$ are compatibly partitioned orthogonal matrices. Then we can set

$$\begin{bmatrix} \widetilde{Y} \\ \widetilde{W} \end{bmatrix} = V_2\,,$$

where V_2 is a matrix whose orthonormal columns span the right nullspace basis. □

10.3.7 Linear Rational Matrix Equations

Several synthesis algorithms presented in Chap. 5 (see Sect. 7.9) involve the solution of linear rational equations of the form

$$G(\lambda)X(\lambda) = F(\lambda), \tag{10.74}$$

where $G(\lambda)$ and $F(\lambda)$ are given $p \times m$ and $p \times q$ rational matrices, respectively, and $X(\lambda)$ is the $m \times q$ rational matrix sought, which must have the least possible McMillan degree. It is a well-known fact that the system (10.74) has a solution provided the rank condition

$$\operatorname{rank} G(\lambda) = \operatorname{rank}[\, G(\lambda)\ F(\lambda)\,] \tag{10.75}$$

is fulfilled. We assume in what follows that this condition holds.

The general solution of (10.74) can be expressed as

$$X(\lambda) = X_0(\lambda) + X_N(\lambda)Y(\lambda), \tag{10.76}$$

where $X_0(\lambda)$ is any particular solution of (10.74), $X_N(\lambda)$ is a rational matrix whose columns form a basis for the right nullspace of $G(\lambda)$, and $Y(\lambda)$ is an arbitrary rational matrix with compatible dimensions. In the case when both $X_0(\lambda)$ and $X_N(\lambda)$ are proper, a possible approach to compute a solution $X(\lambda)$ of least McMillan degree is to determine a suitable proper $Y(\lambda)$ to achieve this goal. A geometric control theoretic method for this purpose has been developed in [88], based on computing minimum dynamic covers. This method has been turned into an efficient and numerically reliable state-space computational approach in [133], which can be used to determine a least McMillan degree solution of (10.74) for this particular case.

Since $X_N(\lambda)$ can always be chosen proper (see Sect. 7.4), the main difficulty using the above approach is the computation of an appropriate $Y(\lambda)$ in the case when there is no proper solution of (10.74), and thus $X_0(\lambda)$ cannot be chosen proper. To overcome this difficulty we can determine $X_0(\lambda)$ so that its polynomial part corresponds to a minimal number of infinite poles. These infinite poles originate from the intrinsic improper nature of any solution of (10.74) and are related to the common infinite zeros of $G(\lambda)$ and $F(\lambda)$. In what follows, we show how to determine a special particular solution $X_0(\lambda)$ with minimum number of infinite poles. Then, we determine a rational basis $X_N(\lambda)$ for the right nullspace of $G(\lambda)$ which will serve to determine a solution $X(\lambda)$ of least McMillan degree. This goal is achieved by employing an approach similar to that of [88] to determine a proper $Y(\lambda)$ to reduce the McMillan degree of the proper part of $X_0(\lambda)$. This approach relies on the generalized minimum cover algorithm of [136].

Computation of $X_0(\lambda)$

Let assume that the rational matrices $G(\lambda)$ and $F(\lambda)$ have descriptor realizations of order n of the forms

$$G(\lambda) := \left[\begin{array}{c|c} A-\lambda E & B_G \\ \hline C & D_G \end{array}\right], \quad F(\lambda) := \left[\begin{array}{c|c} A-\lambda E & B_F \\ \hline C & D_F \end{array}\right], \tag{10.77}$$

where we only assume that the pencil $A - \lambda E$ is regular. Such realizations, which share the pair $(A - \lambda E, C)$, automatically result from a minimal realization of the compound TFM $\left[\, G(\lambda)\; F(\lambda)\,\right]$.

Let $S_G(\lambda)$ and $S_F(\lambda)$ be the system matrix pencils associated to the realizations of $G(\lambda)$ and $F(\lambda)$

$$S_G(\lambda) = \begin{bmatrix} A-\lambda E & B_G \\ C & D_G \end{bmatrix}, \quad S_F(\lambda) = \begin{bmatrix} A-\lambda E & B_F \\ C & D_F \end{bmatrix}.$$

Using the straightforward relations

$$\begin{bmatrix} A-\lambda E & B_G \\ 0 & G(\lambda) \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -C(A-\lambda E)^{-1} & I_p \end{bmatrix} S_G(\lambda),$$

$$\begin{bmatrix} A-\lambda E & B_F \\ 0 & F(\lambda) \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -C(A-\lambda E)^{-1} & I_p \end{bmatrix} S_F(\lambda),$$

it is easy to see that $X(\lambda)$ is a solution of $G(\lambda)X(\lambda) = F(\lambda)$ if and only if

$$Y(\lambda) = \begin{bmatrix} Y_{11}(\lambda) & Y_{12}(\lambda) \\ Y_{21}(\lambda) & X(\lambda) \end{bmatrix}$$

is a solution of

$$S_G(\lambda)Y(\lambda) = S_F(\lambda). \tag{10.78}$$

The existence of the solution of (10.78) is guaranteed by (10.75), which is equivalent to

$$\operatorname{rank} S_G(\lambda) = \operatorname{rank}[\, S_G(\lambda)\; S_F(\lambda)\,]. \tag{10.79}$$

It follows that, instead of solving the rational equation $G(\lambda)X(\lambda) = F(\lambda)$, we can solve the polynomial equation (10.78) and take

$$X(\lambda) = \left[\, 0\; I_m \,\right] Y(\lambda) \begin{bmatrix} 0 \\ I_q \end{bmatrix}.$$

In fact, since we are only interested in the second block column $Y_2(\lambda)$ of $Y(\lambda)$, we need only to solve

$$\begin{bmatrix} A-\lambda E & B_G \\ C & D_G \end{bmatrix} Y_2(\lambda) = \begin{bmatrix} B_F \\ D_F \end{bmatrix} \tag{10.80}$$

and compute $X(\lambda)$ as

$$X(\lambda) = \left[\, 0\; I_m \,\right] Y_2(\lambda).$$

The condition (10.79) for the existence of a solution becomes

$$\operatorname{rank}\begin{bmatrix} A-\lambda E & B_G \\ C & D_G \end{bmatrix} = \operatorname{rank}\begin{bmatrix} A-\lambda E & B_G & B_F \\ C & D_G & D_F \end{bmatrix}. \tag{10.81}$$

To solve (10.80), we isolate a full rank part of $S_G(\lambda)$ by reducing it to a particular Kronecker-like form. Let Q and Z be orthogonal matrices to reduce $S_G(\lambda)$ to the Kronecker-like form

$$\overline{S}_G(\lambda) := QS_G(\lambda)Z = \begin{bmatrix} B_r & A_r-\lambda E_r & A_{r,reg}-\lambda E_{r,reg} & * \\ 0 & 0 & A_{reg}-\lambda E_{reg} & * \\ 0 & 0 & 0 & A_l-\lambda E_l \end{bmatrix}, \tag{10.82}$$

where $A_{reg}-\lambda E_{reg}$ is a regular subpencil, the pair $(A_r-\lambda E_r, B_r)$ is controllable with E_r invertible and the subpencil $A_l-\lambda E_l$ has full column rank for all $\lambda \in \mathbb{C}$. The above reduction can be computed by employing numerically stable algorithms, as those described in Sect. 10.1.6.

If $\overline{Y}_2(\lambda)$ is a solution of the reduced equation

$$\overline{S}_G(\lambda)\overline{Y}_2(\lambda) = Q\begin{bmatrix} B_F \\ D_F \end{bmatrix}, \tag{10.83}$$

then $Y_2(\lambda) = Z\overline{Y}_2(\lambda)$, and thus

$$X(\lambda) = \begin{bmatrix} 0 & I_m \end{bmatrix} Z\overline{Y}_2(\lambda)$$

is a solution of the equation $G(\lambda)X(\lambda) = F(\lambda)$. Partition

$$Q\begin{bmatrix} -B_F \\ -D_F \end{bmatrix} = \begin{bmatrix} \overline{B}_1 \\ \overline{B}_2 \\ \overline{B}_3 \end{bmatrix}$$

in accordance with the row structure of $\overline{S}_G(\lambda)$. Since $A_l-\lambda E_l$ has full column rank, it follows from (10.81) that $\overline{B}_3 = 0$ (otherwise no solution exists). Thus, $\overline{Y}_2(\lambda)$ has the form

$$\overline{Y}_2(\lambda) = \begin{bmatrix} \overline{Y}_{12}(\lambda) \\ \overline{Y}_{22}(\lambda) \\ \overline{Y}_{32}(\lambda) \\ 0 \end{bmatrix},$$

where the partitioning of $\overline{Y}_2(\lambda)$ corresponds to the column partitioning of $\overline{S}_G(\lambda)$. To determine a particular solution $X_0(\lambda)$, we can freely choose $\overline{Y}_{12}(\lambda) = 0$ and determine $\overline{Y}_{22}(\lambda)$ and $\overline{Y}_{32}(\lambda)$ by solving

$$\begin{bmatrix} \overline{Y}_{22}(\lambda) \\ \overline{Y}_{32}(\lambda) \end{bmatrix} = \begin{bmatrix} \lambda E_r - A_r & \lambda E_{r,reg} - A_{r,reg} \\ 0 & \lambda E_{reg} - A_{reg} \end{bmatrix}^{-1} \begin{bmatrix} \overline{B}_1 \\ \overline{B}_2 \end{bmatrix}.$$

Let partition $[\,0\;\; I_m\,]Z$ in accordance with the column structure of $S_G(\lambda)$ as

$$[\,0\;\; I_m\,]Z = [\,D_r\;\; C_r\;\; C_{reg}\;\; C_l\,] \tag{10.84}$$

and denote

$$\overline{A} - \lambda \overline{E} = \begin{bmatrix} A_r - \lambda E_r & A_{r,reg} - \lambda E_{r,reg} \\ 0 & A_{reg} - \lambda E_{reg} \end{bmatrix}, \quad \overline{B} = \begin{bmatrix} \overline{B}_1 \\ \overline{B}_2 \end{bmatrix}, \quad \overline{C} = [\,C_r\;\; C_{reg}\,]\,. \tag{10.85}$$

Then, a particular solution $X_0(\lambda)$ of the equation $G(\lambda)X(\lambda) = F(\lambda)$ can be determined with the descriptor system realization

$$X_0(\lambda) := \left[\begin{array}{c|c} \overline{A} - \lambda \overline{E} & \overline{B} \\ \hline \overline{C} & 0 \end{array}\right]. \tag{10.86}$$

Some properties of $X_0(\lambda)$ can be easily deduced from the computed Kronecker-like form. The pair $(\overline{C}, \overline{A} - \lambda \overline{E})$ is always observable, but, in general, the pair $(\overline{A} - \lambda \overline{E}, \overline{B})$ may be uncontrollable. The poles of $X_0(\lambda)$ are among the generalized eigenvalues of the pair $(\overline{A}, \overline{E})$ and are partly freely assignable and partly fixed. The generalized eigenvalues of the pair (A_r, E_r) are called the "spurious" poles, and they originate from the column singularity of $G(\lambda)$. These poles are freely assignable by appropriate choice of a (non-orthogonal) right transformation matrix [131]. The fixed poles are the controllable eigenvalues of the pair $(A_{reg} - \lambda E_{reg}, \overline{B}_2)$. If $G(\lambda)$ and $F(\lambda)$ have no common poles and zeros then the pair $(A_{reg} - \lambda E_{reg}, \overline{B}_2)$ is controllable. In this case $X_0(\lambda)$ has the minimum possible poles at infinity.

According to the dual of Lemma 9.5, there exists a solution $X_0(\lambda)$ without a pole in γ (finite or infinite) if the pole and zero structures of $G(\lambda)$ and $[G(\lambda)\; F(\lambda)]$ at γ coincide. For practical computations, this implies that some or all of common poles and zeros of $G(\lambda)$ and $[G(\lambda)\; F(\lambda)]$ will cancel. This cancellation can be done explicitly by removing the uncontrollable eigenvalues (finite and infinite) of the pair $(A_{reg} - \lambda E_{reg}, \overline{B}_2)$.

Removing the uncontrollable eigenvalues of the pair $(A_{reg} - \lambda E_{reg}, \overline{B}_2)$ can be done using the generalized controllability staircase form algorithm of **Procedure GCSF** described in Sect. 10.1.5 (see also Sect. 10.3.1). By applying this algorithm, two orthogonal matrices Q_{reg} and Z_{reg} are determined such that all uncontrollable finite eigenvalues are separated in the trailing part of the transformed regular pencil $Q_{reg}(A_{reg} - \lambda E_{reg})Z_{reg}$, while the corresponding rows of $Q_{reg}\overline{B}_2$ are zero. The uncontrollable part of the triple $(\overline{A} - \lambda \overline{E}, \overline{B}, \overline{C})$ can be thus eliminated by removing the appropriate trailing rows and columns from the matrices of the transformed triple $(\overline{Q}(\overline{A} - \lambda \overline{E})\overline{Z}, \overline{QB}, \overline{CZ})$, where $\overline{Q} = \mathrm{diag}(I, Q_{reg})$ and $\overline{Z} = \mathrm{diag}(I, Z_{reg})$. The same technique can be used to remove the uncontrollable infinite eigenvalues by simply interchanging the roles of matrices $\overline{A}$ and $\overline{E}$, thus working on the triple $(\overline{E} - \lambda \overline{A}, \overline{B}, \overline{C})$. For the sake of simplicity we reuse the same notation (with bar) by assuming that the pair $(\overline{A} - \lambda \overline{E}, \overline{B})$ is already controllable, thus the resulting $X_0(\lambda)$ fulfils the requirement for a minimal number of poles at infinity.

To compute the particular solution $X_0(\lambda)$ we employed exclusively orthogonal similarity transformations to determine the matrices of a descriptor realization in (10.86). Therefore, this computation is numerically stable, because we can easily show that the computed system matrices in the presence of roundoff errors are exact for an original problem with slightly perturbed data.

In view of the order reduction step described later, we need to enforce a block diagonal descriptor matrix $\overline{E}$ in (10.85) (i.e., with $E_{r,reg} = 0$). This can be easily achieved by performing an additional non-orthogonal column transformation using the transformation matrix

$$V = \begin{bmatrix} I & -E_r^{-1}E_{r,reg} \\ 0 & I \end{bmatrix}.$$

The transformed system $(\overline{A}V - \lambda\overline{E}V, \overline{B}, \overline{C}V, 0)$, representing also $X_0(\lambda)$, has thus a block diagonal descriptor matrix $\overline{E}V$. To simplify the presentation we will reuse the notation with bar and assume in what follows that $E_{r,reg} = 0$ in (10.85).

Computation of $X_N(\lambda)$

Using the same reduction of $S_G(\lambda)$ to $\overline{S}_G(\lambda)$ as in (10.82), a right nullspace basis $X_N(\lambda)$ of $G(\lambda)$ can be computed from a right nullspace basis $\overline{Y}_N(\lambda)$ of $\overline{S}_G(\lambda)$ as

$$X_N(\lambda) = [\, 0 \; I_m \,] Z \overline{Y}_N(\lambda)\, .$$

We can determine $\overline{Y}_N(\lambda)$ in the form

$$\overline{Y}_N(\lambda) = \begin{bmatrix} I \\ (\lambda E_r - A_r)^{-1}B_r \\ 0 \\ 0 \end{bmatrix}.$$

With C_r and D_r defined in (10.84), we obtain a descriptor realization of $X_N(\lambda)$ as

$$X_N(\lambda) := \left[\begin{array}{c|c} A_r - \lambda E_r & B_r \\ \hline C_r & D_r \end{array}\right].$$

Obviously $X_N(\lambda)$ is proper and controllable. Furthermore, according to Proposition 10.2 applied to the dual realization of $X_N^T(\lambda)$, the realization of $X_N(\lambda)$ is observable, provided the realization of $G(\lambda)$ in (10.77) is observable. Moreover, the poles of $X_N(\lambda)$ are freely assignable by appropriately choosing the transformation matrices Q and Z to reduce the system pencil $S_G(\lambda)$. Note that, to obtain this nullspace basis, we performed exclusively orthogonal transformations on the system matrices. We can prove that all computed matrices are exact for a slightly perturbed original system. It follows that the algorithm to compute the nullspace basis is numerically stable.

Computation of a Least-Order Solution $X(\lambda)$

We can represent $X_N(\lambda)$ to have the same state, descriptor and output matrices as $X_0(\lambda)$. Let these realizations of $X_0(\lambda)$ and $X_N(\lambda)$ be

$$\begin{bmatrix} X_0(\lambda) \ X_N(\lambda) \end{bmatrix} := \left[\begin{array}{c|cc} \overline{A}-\lambda\overline{E} & \overline{B} & \overline{B}_r \\ \hline \overline{C} & \overline{D} & \overline{D}_r \end{array}\right] := \left[\begin{array}{cc|cc} A_r-\lambda E_r & A_{r,reg} & \overline{B}_1 & B_r \\ 0 & A_{reg}-\lambda E_{reg} & \overline{B}_2 & 0 \\ \hline C_r & C_{reg} & 0 & D_r \end{array}\right], \tag{10.87}$$

where E_r is invertible.

We consider first the case when $X_0(\lambda)$ is proper, that is, all eigenvalues of the pencil $A_{reg} - \lambda E_{reg}$ are finite and thus $\overline{E}$ is invertible. In this case, it was shown in [88] that a solution with least McMillan degree can be determined as $X(\lambda) = X_0(\lambda) + X_N(\lambda)Y(\lambda)$ by choosing an appropriate proper $Y(\lambda)$. This can be done by determining a suitable feedback matrix $\overline{F}_r$ and a feedforward matrix $\overline{L}_r$ to cancel the maximum number of unobservable and uncontrollable poles of

$$X(\lambda) := \left[\begin{array}{c|c} \overline{A}+\overline{B}_r\overline{F}_r-\lambda\overline{E} & \overline{B}+\overline{B}_r\overline{L}_r \\ \hline \overline{C}+\overline{D}_r\overline{F}_r & \overline{D}+\overline{D}_r\overline{L}_r \end{array}\right]. \tag{10.88}$$

It can be shown that if we start with a minimal realization of $[\, G(\lambda) \ F(\lambda)\,]$, then we cannot produce any unobservable poles in $X(\lambda)$ via state feedback. Therefore, we only need to determine the matrices $\overline{F}_r$ and $\overline{L}_r$ to cancel the maximum number of uncontrollable poles.

This problem has been solved in [88] by reformulating it as a minimal order dynamic cover design problem. We denote $\widetilde{A} := \overline{E}^{\,-1}\overline{A}$, $\widetilde{B} := \overline{E}^{\,-1}\overline{B}$, and $\widetilde{B}_r := \overline{E}^{\,-1}\overline{B}_r$, and also $\widetilde{\mathcal{B}} = \operatorname{span} \widetilde{B}$ and $\widetilde{\mathcal{B}}_r = \operatorname{span} \widetilde{B}_r$. Consider the set

$$\mathcal{J} = \{\mathcal{V} : \widetilde{\mathcal{B}} + \widetilde{A}\mathcal{V} \subset \widetilde{\mathcal{B}}_r + \mathcal{V}\},$$

and let $\mathcal{J}^*$ denote the set of subspaces in $\mathcal{J}$ of least dimension. If $\mathcal{V} \in \mathcal{J}^*$, then a pair $(\overline{F}_r, \overline{L}_r)$ can be determined such that

$$(\widetilde{A} + \widetilde{B}_r\overline{F}_r)\mathcal{V} + \operatorname{span}(\widetilde{B} + \widetilde{B}_r\overline{L}_r) \subset \mathcal{V}.$$

Thus, determining a minimal dimension $\mathcal{V}$ is equivalent to a minimal order cover design problem, and a conceptual geometric approach to solve it has been indicated in [88]. The outcome of his method is, besides $\mathcal{V}$, the pair $(\overline{F}_r, \overline{L}_r)$ which achieves a maximal order reduction by forcing pole–zero cancellations. This approach, in the case of standard systems (i.e., $\overline{E} = I$), has been turned into a numerically reliable procedure in [133] and extended to the descriptor case with invertible $\overline{E}$ in [136]. In this latter procedure, $\overline{F}_r$ and $\overline{L}_r$ are determined from a special controllability staircase form of the pair $(\overline{A} - \lambda\overline{E}, [\,\overline{B}_r \ \overline{B}\,])$ obtained using a numerically reliable method relying on both orthogonal and non-orthogonal similarity transformations. Details of this algorithm are given in Sect. 10.4.3.

It is possible to refine this approach by exploiting the structure of matrices in (10.87). Assuming $\overline{F}_r = [\, F_r \; F_{reg} \,]$ is partitioned according to the structure of $\overline{A}$, we get from (10.88)

$$X(\lambda) := \left[\begin{array}{cc|c} A_r + B_r F_r - \lambda E_r & A_{r,reg} + B_r F_{reg} & \overline{B}_1 + B_r \overline{L}_r \\ 0 & A_{reg} - \lambda E_{reg} & \overline{B}_2 \\ \hline C_r + D_r F_r & C_{reg} + D_r F_{reg} & \overline{D} + D_r \overline{L}_r \end{array}\right].$$

Since the eigenvalues of $A_{reg} - \lambda E_{reg}$ are not controllable via $\overline{B}_r$, the state feedback $\overline{F}_r$ affects only the blocks $A_r - \lambda E_r$ and $A_{r,reg}$. To make a maximum number of eigenvalues of $A_r + B_r F_r - \lambda E_r$ uncontrollable we can alternatively solve a minimum dynamic cover problem of lower dimension for the system

$$\left[\, X_{0,r}(\lambda) \; X_N(\lambda) \,\right] := \left[\begin{array}{c|cc} A_r - \lambda E_r & \left[\, A_{r,reg} \; \overline{B}_1 \,\right] & B_r \\ \hline C_r & \left[\, C_{r,reg} \;\; \overline{D} \,\right] & D_r \end{array}\right],$$

by determining an appropriate state feedback matrix F_r and a feedforward matrix $[\, F_{reg} \; \overline{L}_r \,]$. Besides lower size of the computational problem, the main advantage of this approach is that it is applicable regardless $A_{reg} - \lambda E_{reg}$ has infinite eigenvalues or not.

10.4 Special Algorithms

In this section we describe several algorithms, which are instrumental in addressing least-order synthesis problems of fault detection and isolation filters and the solution of the Nehari problem, which is encountered in solving least distance problems.

10.4.1 Special Controllability Staircase Form Algorithm

The computational methods of minimum dynamic covers, presented in Sects. 10.4.2 and 10.4.3, rely on a special controllability staircase form (see Sect. 10.1.5) involving a controllable descriptor pair $(A - \lambda E, [\, B_1 \; B_2 \,])$, where $A, E \in \mathbb{R}^{n\times n}$ with E invertible, $B_1 \in \mathbb{R}^{n\times m_1}$, $B_2 \in \mathbb{R}^{n\times m_2}$. The main difference to the reduction performed in **Procedure GCSF** is in exploiting, at the j-th reduction step, the partitioned form of the matrix $B^{(j-1)} := [\, B_1^{(j-1)} \; B_2^{(j-1)} \,]$, by compressing its rows in two steps. In the first step, the rows of $B_1^{(j-1)}$ are compressed, while in the second step, those columns of the updated $B_2^{(j-1)}$ are compressed, which are linearly independent of the columns of $B_1^{(j-1)}$. All row compressions can be performed using orthogonal similarity transformations.

The following procedure determines for a descriptor triple $(A - \lambda E, [\, B_1 \; B_2 \,], C)$, two orthogonal transformation matrices Q and Z such that for the resulting triple $(Q^T AZ - \lambda Q^T EZ, [\, Q^T B_1 \; Q^T B_2 \,], CZ)$, the pencil $[\, Q^T B_1 \; Q^T B_2 \; Q^T AZ - \lambda Q^T EZ \,]$

is in a special controllability staircase form with $Q^T EZ$ upper triangular.

Procedure GSCSF: Generalized special controllability staircase form

Input : $(A - \lambda E, [\, B_1 \; B_2 \,], C)$
Outputs: $Q, Z, (A-\lambda E, [\, B_1 \; B_2 \,], C) := (Q^T AZ - \lambda Q^T EZ, [\, Q^T B_1 \; Q^T B_2 \,], CZ),$
$(\nu_{1,j}, \nu_{2,j}),\ j = 1, \ldots, \ell$

1) Compute an orthogonal matrix Q such that $Q^T E$ is upper triangular; compute $A \leftarrow Q^T A$, $E \leftarrow Q^T E$, $B_1 \leftarrow Q^T B_1$, $B_2 \leftarrow Q^T B_2$. Set $Z = I_n$.
2) Set $j=1$, $r=0$, $\nu_{1,0} = m_1$, $\nu_{2,0} = m_2$, $A^{(0)} = A$, $E^{(0)} = E$, $B_1^{(0)} = B_1$, $B_2^{(0)} = B_2$.
3) Compute orthogonal matrices W and U such that

$$W^T \left[B_1^{(j-1)} \middle| B_2^{(j-1)} \right] := \begin{array}{c} \left[\begin{array}{c|c} A_{2j-1,2j-3} & A_{2j-1,2j-2} \\ 0 & A_{2j,2j-2} \\ 0 & 0 \end{array} \right] \begin{array}{l} \nu_{1,j} \\ \nu_{2,j} \\ \rho \end{array} \\ \begin{array}{cc} \nu_{1,j-1} & \quad \nu_{2,j-1} \end{array} \end{array}$$

with $A_{2j-1,2j-3}$ and $A_{2j,2j-2}$ full row rank matrices and $W^T E^{(j-1)} U$ is upper triangular.
4) Compute and partition

$$W^T A^{(j-1)} U := \begin{array}{c} \left[\begin{array}{ccc} A_{2j-1,2j-1} & A_{2j-1,2j} & A_{2j-1,2j+1} \\ A_{2j,2j-1} & A_{2j,2j} & A_{2j,2j+1} \\ B_1^{(j)} & B_2^{(j)} & A^{(j)} \end{array} \right] \begin{array}{l} \nu_{1,j} \\ \nu_{2,j} \\ \rho \end{array} \\ \begin{array}{ccc} \nu_{1,j} & \quad \nu_{2,j} & \quad \rho \end{array} \end{array}$$

$$W^T E^{(j-1)} U := \begin{array}{c} \left[\begin{array}{ccc} E_{2j-1,2j-1} & E_{2j-1,2j} & E_{2j-1,2j+1} \\ 0 & E_{2j,2j} & E_{2j,2j+1} \\ 0 & 0 & E^{(j)} \end{array} \right] \begin{array}{l} \nu_{1,j} \\ \nu_{2,j} \\ \rho \end{array} \\ \begin{array}{ccc} \nu_{1,j} & \quad \nu_{2,j} & \quad \rho \end{array} \end{array}$$

5) For $i = 1, \ldots, 2j-2$ compute and partition

$$A_{i,2j-1} U := \begin{array}{c} [\, A_{i,2j-1} \;\; A_{i,2j} \;\; A_{i,2j+1} \,] \\ \begin{array}{ccc} \nu_{1,j} & \nu_{2,j} & \rho \end{array} \end{array}$$

$$E_{i,2j-1} U := \begin{array}{c} [\, E_{i,2j-1} \;\; E_{i,2j} \;\; E_{i,2j+1} \,] \\ \begin{array}{ccc} \nu_{1,j} & \nu_{2,j} & \rho \end{array} \end{array}$$

6) $Q \leftarrow Q \operatorname{diag}(I_r, W)$, $Z \leftarrow Z \operatorname{diag}(I_r, U)$, $C \leftarrow C \operatorname{diag}(I_r, U)$.
7) $r \leftarrow r + \nu_{1,j} + \nu_{2,j}$; if $\rho = 0$, then $\ell = j$ and **Exit**;
else, $j \leftarrow j + 1$ and go to Step 3).

At the end of this algorithm we have $\widehat{A} - \lambda\widehat{E} := Q^T(A - \lambda E)Z$, $\widehat{B} := [\, Q^T B_1 \;\; Q^T B_2 \,]$, $\widehat{C} := CZ$, $\widehat{E}$ is upper triangular, and the pair $(\widehat{A}, \widehat{B})$ is in the *special staircase form*

$$
\big[\,\widehat{B}\,|\,\widehat{A}\,\big] = \left[\begin{array}{cc|cccccccc}
A_{1,-1} & A_{1,0} & A_{11} & A_{12} & \cdots & A_{1,2\ell-3} & A_{1,2\ell-2} & A_{1,2\ell-1} & A_{1,2\ell} \\
0 & A_{2,0} & A_{21} & A_{22} & \cdots & A_{2,2\ell-3} & A_{2,2\ell-2} & A_{2,2\ell-1} & A_{2,2\ell} \\
0 & 0 & A_{31} & A_{32} & \cdots & A_{3,2\ell-3} & A_{3,2\ell-2} & A_{3,2\ell-1} & A_{3,2\ell} \\
0 & 0 & 0 & A_{42} & \cdots & A_{4,2\ell-3} & A_{4,2\ell-2} & A_{4,2\ell-1} & A_{4,2\ell} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & A_{2\ell-1,2\ell-3} & A_{2\ell-1,2\ell-2} & A_{2\ell-1,2\ell-1} & A_{2\ell-1,2\ell} \\
0 & 0 & 0 & 0 & \cdots & 0 & A_{2\ell,2\ell-2} & A_{2\ell,2\ell-1} & A_{2\ell,2\ell}
\end{array}\right], \qquad (10.89)
$$

where $A_{2j-1,2j-3} \in \mathbb{R}^{\nu_{1,j}\times\nu_{1,j}}$ and $A_{2j,2j-2} \in \mathbb{R}^{\nu_{2,j}\times\nu_{2,j}}$ are full row rank matrices for $j = 1, \ldots, \ell$. The resulting upper triangular matrix $\widehat{E}$ has a similar block partitioned form

$$
\widehat{E} = \begin{bmatrix}
E_{11} & E_{12} & \cdots & E_{1,2\ell-1} & E_{1,2\ell} \\
0 & E_{22} & \cdots & E_{2,2\ell-1} & E_{2,2\ell} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & E_{2\ell-1,2\ell-1} & E_{2\ell-1,2\ell} \\
0 & 0 & \cdots & 0 & E_{2\ell,2\ell}
\end{bmatrix}. \qquad (10.90)
$$

The resulting block dimensions $(\nu_{1,j}, \nu_{2,j})$, $j = 1, \ldots, \ell$, satisfy

$$
m_1 = \nu_{1,0} \ge \nu_{1,1} \ge \cdots \ge \nu_{1,\ell} \ge 0
$$

and

$$
m_2 = \nu_{2,0} \ge \nu_{2,1} \ge \cdots \ge \nu_{2,\ell} \ge 0
$$

and represents the dimensions $n_1 := \sum_{i=1}^{\ell} \nu_{1,j}$ and $n_2 := \sum_{j=1}^{\ell} \nu_{2,j}$ of two subspaces, which underlie the computation of appropriate minimal dynamic covers in the next sections.

When implementing **Procedure GSCSF**, the row compressions at Step 3) are usually performed using rank-revealing QR factorizations with column pivoting. This computation can be done in two steps, first by compressing the r rows of $B_1^{(j-1)}$ to a full row rank matrix $A_{2j-1,2j-3}$ using an orthogonal matrix W_1 (i.e., as $W_1^T B_1^{(j-1)}$), and then by compressing the trailing $r - \nu_{1,j}$ rows of $W_1^T B_2^{(j-1)}$ to a full row rank matrix $A_{2j,2j-2}$ using a second orthogonal matrix W_2. The overall transformation W at Step 3) results as $W = W_1 \,\mathrm{diag}(I_{\nu_{1,j}}, W_2)$. Both reductions can be performed using sequences of Givens rotations, which allow to simultaneously perform the column transformations accumulated in U to maintain the upper triangular form of $E^{(j-1)}$. This reduction technique is described in detail in [125]. Using this technique, the numerical complexity of **Procedure GSCSF** is $\mathcal{O}(n^3)$, provided all transformations are immediately applied without accumulating explicitly W and U. The usage of the more robust rank determinations based on singular values decompositions would increase the overall complexity to $\mathcal{O}(n^4)$ due to the need to accumulate explicitly W and U. Regarding the numerical properties of **Procedure GSCSF**, it is possible to show that the resulting system matrices $\widehat{A}$, $\widehat{E}$, $\widehat{B}$, $\widehat{C}$ are exact for slightly perturbed original data A, E, B, C, while Q and Z are nearly orthogonal matrices. It follows that the **Procedure GSCSF** is numerically stable. In the standard case we have $E = I$, and therefore $Q = Z$ and $\widehat{E} = I$.

Example 10.1 For $\ell = 3$, $[\,\widehat{B}\ \ \widehat{A}\,]$ and $\widehat{E}$ have similarly block partitioned forms

$$[\,\widehat{B}\,|\,\widehat{A}\,] = \left[\begin{array}{cc|cccccc} A_{1,-1} & A_{1,0} & A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} \\ 0 & A_{2,0} & A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} \\ 0 & 0 & A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} \\ 0 & 0 & 0 & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} \\ 0 & 0 & 0 & 0 & A_{53} & A_{54} & A_{55} & A_{56} \\ 0 & 0 & 0 & 0 & 0 & A_{64} & A_{65} & A_{66} \end{array}\right], \quad \widehat{E} = \left[\begin{array}{cccc} E_{11} & E_{12} & \cdots & E_{16} \\ 0 & E_{22} & \cdots & E_{26} \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & E_{66} \end{array}\right].$$

◊

10.4.2 Order Reduction Using Minimum Dynamic Covers of Type I

The computational problem which we address in this section is the following: given a descriptor pair $(A - \lambda E, B)$ with $A, E \in \mathbb{R}^{n\times n}$ and E invertible, $B \in \mathbb{R}^{n\times m}$, and B partitioned as $B = [\,B_1\ B_2\,]$ with $B_1 \in \mathbb{R}^{n\times m_1}$, $B_2 \in \mathbb{R}^{n\times m_2}$, determine the matrix $F \in \mathbb{R}^{m_2\times n}$ such that the pair $(A + B_2F - \lambda E, B_1)$ is *maximally uncontrollable* (i.e., $A + B_2F - \lambda E$ has maximal number of uncontrollable eigenvalues).

This computation is useful to determine least-order solutions of linear rational equations using state feedback techniques. Consider the compatible linear rational system of equations $G(\lambda)X(\lambda) = F(\lambda)$, where $G(\lambda)$ and $F(\lambda)$ are given and $X(\lambda)$ is sought. Assume $X_1(\lambda)$ and $X_2(\lambda)$ are two proper TFMs, which generate all solutions of the rational system of equation $G(\lambda)X(\lambda) = F(\lambda)$ in the form

$$X(\lambda) = X_1(\lambda) + X_2(\lambda)Y(\lambda), \tag{10.91}$$

where $X_1(\lambda)$ is any particular solution satisfying $G(\lambda)X_1(\lambda) = F(\lambda)$, $X_2(\lambda)$ is a proper rational basis of the right nullspace of $G(\lambda)$ (i.e., $G(\lambda)X_2(\lambda) = 0$), and $Y(\lambda)$ is arbitrary, having appropriate dimensions. Assume $X_1(\lambda)$ and $X_2(\lambda)$ have the descriptor system realizations

$$\big[\,X_1(\lambda)\ X_2(\lambda)\,\big] = \left[\begin{array}{c|cc} A - \lambda E & B_1 & B_2 \\ \hline C & D_1 & D_2 \end{array}\right], \tag{10.92}$$

with the descriptor pair $(A - \lambda E, [\,B_1\ B_2\,])$ controllable and E invertible. Let F be a state feedback gain and define the TFMs

$$\left[\,\widetilde{X}_1(\lambda)\ \widetilde{X}_2(\lambda)\,\right] := \left[\begin{array}{c|cc} A+B_2F-\lambda E & B_1 & B_2 \\ \hline C+D_2F & D_1 & D_2 \end{array}\right]. \tag{10.93}$$

It is straightforward to check that

$$\widetilde{X}_1(\lambda) = X_1(\lambda) + X_2(\lambda)Y(\lambda), \quad \widetilde{X}_2(\lambda) = X_2(\lambda)\widetilde{Y}(\lambda), \tag{10.94}$$

where $Y(\lambda)$ and $\widetilde{Y}(\lambda)$ have the descriptor system realizations

$$\left[\,Y(\lambda)\ \widetilde{Y}(\lambda)\,\right] = \left[\begin{array}{c|cc} A+B_2F-\lambda E & B_1 & B_2 \\ \hline F & 0 & I \end{array}\right]. \tag{10.95}$$

Therefore, $\widetilde{X}_1(\lambda)$ and $\widetilde{X}_2(\lambda)$ also generate all solutions, because $\widetilde{X}_1(\lambda)$ is another particular solution, while $\widetilde{X}_2(\lambda)$ is another right nullspace basis, because $\widetilde{Y}(\lambda)$ is invertible. If F is determined such that the pair $(A+B_2F-\lambda E, B_1)$ is maximally uncontrollable, then the resulting realization of $\widetilde{X}_1(\lambda)$ contains a maximum number of uncontrollable eigenvalues which can be eliminated using minimal realization techniques. Thus, $\widetilde{X}_1(\lambda)$ represents another particular solution with a reduced McMillan degree.

Remark 10.9 The above approach achieves the maximum order reduction for $\widetilde{X}_1(\lambda)$ provided the descriptor system realization $(A-\lambda E, B_2, C, D_2)$ is *maximally observable*, i.e., the pair $(A+B_2F-\lambda E, C+D_2F)$ is observable for any F [88]. If this condition is not fulfilled, then the least -order can be achieved after a preliminary order reduction, where a maximum number of unobservable eigenvalues are eliminated using a suitable choice of F. If $E=I$ and $D_2=0$, a numerically stable algorithm proposed in [116] to compute the maximal (A, B_2)-invariant subspace contained in the kernel of C can be employed for this purpose. If E is a general invertible matrix, then the same algorithm can be applied to the triple $(E^{-1}A, E^{-1}B_2, C)$, provided E is not too ill conditioned. The case $D_2 \neq 0$ can be addressed using the extended system technique suggested in [6, p. 240]. □

An important application of the above order reduction technique is to determine least-order combinations of a left nullspace basis vectors, which satisfy additional fault detectability conditions (see Sect. 7.5). In this case, we deal with a homogeneous equation $Q(\lambda)G(\lambda)=0$ and find a suitable fault detection filter $Q(\lambda)$ in the form

$$Q(\lambda) = HN_l(\lambda) + Y(\lambda)N_l(\lambda), \tag{10.96}$$

where $N_l(\lambda)$ is a proper rational left nullspace basis of $G(\lambda)$ and H is a constant matrix (to be appropriately selected to fulfil the fault detectability condition). Assuming $N_l(\lambda)$ has the observable descriptor realization

$$N_l(\lambda) = \left[\begin{array}{c|c} A_l-\lambda E_l & B_l \\ \hline C_l & D_l \end{array}\right],$$

this leads to a *dual* problem to be solved in Sect. 7.5, which involves an observable pair $(A_l - \lambda E_l, \widetilde{C}_l)$ with invertible E_l and with a $\widetilde{C}_l$ matrix partitioned as

$$\widetilde{C}_l = \begin{bmatrix} HC_l \\ C_l \end{bmatrix}.$$

In this case, a matrix K is sought such that the pair $(A_l + KC_l - \lambda E_l, HC_l)$ is *maximally unobservable*. For this purpose, the algorithm described in this section can be applied to the controllable pair $(A_l^T - \lambda E_l^T, [\, HC_l^T \;\; C_l^T \,])$ to determine a suitable "state feedback" K^T, which cancels the maximum number of uncontrollable eigenvalues.

We denote $\overline{A} = E^{-1}A$, $\overline{B}_1 = E^{-1}B_1$, $\overline{B}_2 = E^{-1}B_2$, and also and $\overline{\mathcal{B}}_1 = \operatorname{span} \overline{B}_1$ and $\overline{\mathcal{B}}_2 = \operatorname{span} \overline{B}_2$. The problem to determine F which makes the pair $(A + B_2F - \lambda E, B_1)$ maximally uncontrollable is equivalent [162] to compute a subspace $\mathcal{V}$ of least possible dimension satisfying

$$(\overline{A} + \overline{B}_2F)\mathcal{V} \subset \mathcal{V}, \quad \overline{\mathcal{B}}_1 \subset \mathcal{V}. \tag{10.97}$$

This subspace is the least-order $(\overline{A}, \overline{B}_2)$-invariant subspace which contains $\overline{\mathcal{B}}_1$ [162]. The above condition can be equivalently rewritten as a condition defining $\mathcal{V}$ as a *Type I* minimum dynamic cover [40, 71]

$$\overline{A}\mathcal{V} \subset \mathcal{V} + \overline{\mathcal{B}}_2, \quad \overline{\mathcal{B}}_1 \subset \mathcal{V}. \tag{10.98}$$

In this section we describe a computational method for determining minimal dynamic covers, which relies on the reduction of the descriptor system pair $(A - \lambda E, [B_1, B_2])$ to a particular condensed form, for which the solution of the problem (i.e., the choice of appropriate F) is simple. This reduction is performed in two stages. The first stage is the orthogonal reduction performed with the **Procedure GSCSF** presented in Sect. 10.4.1. In the second stage, additional zero blocks are generated in the reduced matrices using non-orthogonal transformations. With additional blocks zeroed via a specially chosen state feedback F, the least-order $(\overline{A}, \overline{B}_2)$-invariant subspace containing $\overline{\mathcal{B}}_1$ can be identified as the linear span of the leading columns of the resulting right transformation matrix. In what follows we present in detail the second reduction stage as well as the determination of F.

We assume that after performing the **Procedure GSCSF**, we obtained the orthogonal transformation matrices Q and Z, such that the transformed system triple

$$(\widehat{A} - \lambda\widehat{E}, [\, \widehat{B}_1 \; \widehat{B}_2 \,], \widehat{C}) := (Q^TAZ - \lambda Q^TEZ, [\, Q^TB_1 \; Q^TB_2 \,], CZ) \tag{10.99}$$

has the pair $(\widehat{A}, \widehat{B})$, with $\widehat{B} = [\, \widehat{B}_1 \; \widehat{B}_2, \,]$, in the staircase form (10.89) and the matrix $\widehat{E}$ in the block structured form (10.90). The dimensions of the first 2ℓ diagonal blocks of $\widehat{A}$ and $\widehat{E}$ are determined by the two sets of dimensions $\nu_{1,j}$ and $\nu_{2,j}$ for $j = 1, \ldots, \ell$, and define the dimensions $n_1 := \sum_{j=1}^{\ell} \nu_{1,j}$ and $n_2 := \sum_{j=1}^{\ell} \nu_{2,j}$. Additionally, partition the columns of the resulting $\widehat{C}$ in accordance with the column structure of $\widehat{A}$ in (10.89)

$$\widehat{C} = \left[\, C_1 \; C_2 \; \cdots \; C_{2\ell-1} \; C_{2\ell} \,\right]. \tag{10.100}$$

In the second reduction stage we use non-orthogonal upper triangular left and right transformation matrices W and U, respectively, to annihilate the minimum number of blocks in $\widehat{A}$ and $\widehat{E}$ which allows to solve the minimum cover problem. Assume W and U have block structures identical to $\widehat{E}$. By exploiting the full rank of submatrices $A_{2k,2k-2}$ we can introduce zero blocks in the block row $2k$ of $\widehat{A}$ by annihilating the blocks $A_{2k,2j-1}$, for $j = k, k+1, \ldots, \ell$. Similarly, by exploiting the invertibility of $E_{2j-1,2j-1}$, we can introduce zero blocks in the block row $2k-2$ of E by annihilating the blocks $E_{2k-2,2j-1}$, for $j = k, k+1, \ldots, \ell$ of $\widehat{E}$. This computation is performed for $k = \ell, \ell-1, \ldots, 2$. Let $\widetilde{A} := W\widehat{A}U$, $\widetilde{E} := W\widehat{E}U$, $[\,\widetilde{B}_1\ \widetilde{B}_2\,] := W[\,\widehat{B}_1\ \widehat{B}_2\,] = [\,\widehat{B}_1\ \widehat{B}_2\,]$, and $\widetilde{C} = \widehat{C}U$ be the system matrices resulted after this (non-orthogonal) reduction. Define also the feedback matrix $\widetilde{F} \in \mathbb{R}^{m_2 \times n}$ partitioned column-wise compatibly with $\widehat{A}$

$$\widetilde{F} = [\,F_1\ 0\ F_3\ \cdots\ 0\ F_{2\ell-1}\ 0\,],$$

where $F_{2j-1} \in \mathbb{R}^{m_2 \times \nu_{1,j}}$ are such that $A_{2,0}F_{2j-1} + A_{2,2j-1} = 0$ for $j = 1, \ldots, \ell$. With this feedback we introduced ℓ zero blocks in the second block row of $\widetilde{A} + \widetilde{B}_2\widetilde{F}$. Finally, consider the permutation matrix defined by

$$P = \left[\begin{array}{cc|cc|c|cc} I_{\nu_{1,1}} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & I_{\nu_{1,2}} & 0 & \cdots & 0 & 0 \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & 0 & 0 & \cdots & I_{\nu_{1,\ell}} & 0 \\ \hline 0 & I_{\nu_{2,1}} & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & 0 & I_{\nu_{2,2}} & \cdots & 0 & 0 \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & 0 & 0 & \cdots & 0 & I_{\nu_{2,\ell}} \end{array}\right]. \qquad (10.101)$$

If we define $L = PWQ^T$, $V = ZUP^T$ and $F = \widetilde{F}V^{-1}$, then overall we obtained the reduced system $(\breve{A} - \lambda\breve{E}, [\,\breve{B}_1\ \breve{B}_2\,], \breve{C}, [\,D_1\ D_2\,])$ defined with

$$\begin{aligned} \breve{A} - \lambda\breve{E} &:= L(A + B_2F - \lambda E)V = \left[\begin{array}{c|c} \breve{A}_{11} - \lambda\breve{E}_{11} & \breve{A}_{12} - \lambda\breve{E}_{12} \\ \hline 0 & \breve{A}_{22} - \lambda\breve{E}_{22} \end{array}\right], \\ \left[\,\breve{B}_1 \,\middle|\, \breve{B}_2\,\right] &:= L\left[\,B_1 \,\middle|\, B_2\,\right] = \left[\begin{array}{c|c} \breve{B}_{11} & \breve{B}_{12} \\ \hline 0 & \breve{B}_{22} \end{array}\right], \\ \breve{C} &:= (C + D_2F)V = \left[\,\breve{C}_1 \,\middle|\, \breve{C}_2\,\right], \end{aligned} \qquad (10.102)$$

where, by construction, the pairs $(\breve{A}_{11} - \lambda\breve{E}_{11}, \breve{B}_{11})$ and $(\breve{A}_{22} - \lambda\breve{E}_{22}, \breve{B}_{22})$ are in controllable staircase form. Thus, by the above choice of F, we made n_2 of the n eigenvalues of the pencil $A + B_2F - \lambda E$ uncontrollable via B_1. It is straightforward to show that the matrix V_1 formed from the the first n_1 columns of V satisfies

$$\overline{A}V_1 = V_1\breve{E}_{11}^{-1}\breve{A}_{11} - \overline{B}_2FV_1, \quad \overline{B}_1 = V_1\breve{E}_{11}^{-1}\breve{B}_{11}\,.$$

Thus, according to (10.98), $\mathcal{V} := \operatorname{span} V_1$ is a dynamic cover of *Type I* of dimension n_1. It can be shown using the results of [71] that $\mathcal{V}$ has minimum dimension.

To illustrate the computational procedure, we consider the reduced system in Example 10.1. First, the following zero blocks are introduced: A_{65}, E_{45}, A_{43}, A_{45}, E_{23}, E_{25} (in this order). The resulting $\widetilde{A}$ and $\widetilde{E}$ are

$$\widetilde{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} \\ 0 & A_{42} & 0 & A_{44} & 0 & A_{46} \\ 0 & 0 & A_{53} & A_{54} & A_{55} & A_{56} \\ 0 & 0 & 0 & A_{64} & 0 & A_{66} \end{bmatrix}, \quad \widetilde{E} = \begin{bmatrix} E_{11} & E_{12} & E_{13} & E_{14} & E_{15} & E_{16} \\ 0 & E_{22} & 0 & E_{24} & 0 & E_{26} \\ 0 & 0 & E_{33} & E_{34} & E_{35} & E_{36} \\ 0 & 0 & 0 & E_{44} & 0 & E_{46} \\ 0 & 0 & 0 & 0 & E_{55} & E_{56} \\ 0 & 0 & 0 & 0 & 0 & E_{66} \end{bmatrix}.$$

Additional blocks are zeroed using the feedback $\widetilde{F}$ to obtain

$$\widetilde{A} + \widetilde{B}_2\widetilde{F} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} \\ 0 & A_{22} & 0 & A_{24} & 0 & A_{26} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} \\ 0 & A_{42} & 0 & A_{44} & 0 & A_{46} \\ 0 & 0 & A_{53} & A_{54} & A_{55} & A_{56} \\ 0 & 0 & 0 & A_{64} & 0 & A_{66} \end{bmatrix}.$$

Finally, after block permutations, we obtained the controllable staircase forms

$$\left[\,\breve{B}_{11}\,\middle|\,\breve{A}_{11} - \lambda\breve{E}_{11}\,\right] = \left[\begin{array}{c|ccc} A_{1,-1} & A_{11} - \lambda E_{11} & A_{13} - \lambda E_{13} & A_{15} - \lambda E_{15} \\ 0 & A_{31} & A_{33} - \lambda E_{33} & A_{35} - \lambda E_{35} \\ 0 & 0 & A_{53} & A_{55} - \lambda E_{55} \end{array}\right],$$

$$\left[\,\breve{B}_{22}\,\middle|\,\breve{A}_{22} - \lambda\breve{E}_{22}\,\right] = \left[\begin{array}{c|ccc} A_{2,0} & A_{22} - \lambda E_{22} & A_{24} - \lambda E_{24} & A_{26} - \lambda E_{26} \\ 0 & A_{42} & A_{44} - \lambda E_{44} & A_{46} - \lambda E_{46} \\ 0 & 0 & A_{64} & A_{66} - \lambda E_{66} \end{array}\right].$$

The above approach to compute a minimum dynamic cover of Type I is the basis of **Procedure GRMCOVER1**, presented in what follows. This procedure determines, for a pair of generators $(X_1(\lambda), X_2(\lambda))$ with the descriptor realizations given in (10.92), explicit minimal realizations for $\widetilde{X}_1(\lambda)$ and $Y(\lambda)$ (see (10.93) and (10.95)) in the form

$$\widetilde{X}_1(\lambda) = \left[\begin{array}{c|c} \lambda\breve{E}_{11} - \breve{A}_{11} & \breve{B}_{11} \\ \hline \breve{C}_1 & D_1 \end{array}\right], \qquad Y(\lambda) = \left[\begin{array}{c|c} \lambda\breve{E}_{11} - \breve{A}_{11} & \breve{B}_{11} \\ \hline \breve{F}_1 & 0 \end{array}\right],$$

where $\widetilde{F}P^T =: \begin{bmatrix} \breve{F}_1 & \breve{F}_2 \end{bmatrix}$, with $\breve{F}_1$ having n_1 columns.

Procedure GRMCOVER1: Order reduction using dynamic covers of Type I

Inputs : $X_1(\lambda) = (A - \lambda E, B_1, C, D_1)$ and $X_2(\lambda) = (A - \lambda E, B_2, C, D_2)$
Outputs: $\widetilde{X}_1(\lambda) = (\breve{A}_{11} - \lambda \breve{E}_{11}, \breve{B}_{11}, \breve{C}_1, D_1)$ and
$Y(\lambda) = (\breve{A}_{11} - \lambda \breve{E}_{11}, \breve{B}_{11}, \breve{F}_1, 0)$ such that
$\widetilde{X}_1(\lambda) = X_1(\lambda) + X_2(\lambda)Y(\lambda)$ has least McMillan degree.

1) Apply **Procedure GSCSF** to the system triple $(A-\lambda E, [\, B_1 \; B_2 \,], C)$ to determine the orthogonally similar system triple $(\widehat{A} - \lambda \widehat{E}, [\, \widehat{B}_1 \; \widehat{B}_2 \,], \widehat{C})$ defined in (10.99) and (10.100), and the dimensions $\nu_{1,j}$ and $\nu_{2,j}$ for $j = 1, \ldots, \ell$; set $n_1 := \sum_{j=1}^{\ell} \nu_{1,j}$.
2) With $\widehat{A}$ partitioned as in (10.89) and $\widehat{E}$ partitioned as in (10.90), perform the second stage of the special reduction for Type I covers:

   Set $W = I$, $U = I$, and partition W and U in blocks analogous to $\widehat{E}$ in (10.90).

   for $k = \ell, \ell - 1, \ldots, 2$
   Comment. Annihilate blocks $A_{2k,2j-1}$, for $j = k, k+1, \ldots, \ell$.
   for $j = k, k+1, \ldots, \ell$
   Compute $U_{2k-2,2j-1}$ such that $A_{2k,2k-2}U_{2k-2,2j-1} + A_{2k,2j-1} = 0$.
   $A_{i,2j-1} \leftarrow A_{i,2j-1} + A_{i,2k-2}U_{2k-2,2j-1}$, $i = 1, 2, \ldots, 2k$.
   $E_{i,2j-1} \leftarrow E_{i,2j-1} + E_{i,2k-2}U_{2k-2,2j-1}$, $i = 1, 2, \ldots, 2k - 2$.
   $C_{2j-1} \leftarrow C_{2j-1} + C_{2k-2}U_{2k-2,2j-1}$.
   $U_{i,2j-1} \leftarrow U_{i,2j-1} + U_{i,2k-2}U_{2k-2,2j-1}$, $i = 1, 2, \ldots, 2\ell$.
   end
   Comment. Annihilate blocks $E_{2k-2,2j-1}$, for $j = k, k+1, \ldots, \ell$.
   for $j = k, k+1, \ldots, \ell$
   Compute $W_{2k-2,2j-1}$ such that $W_{2k-2,2j-1}E_{2j-1,2j-1} + E_{2k-2,2j-1} = 0$.
   $A_{2k-2,i} \leftarrow A_{2k-2,i} + W_{2k-2,2j-1}A_{2j-1,i}$, $i = 2j-2, 2j-1, \ldots, 2\ell$.
   $E_{2k-2,i} \leftarrow E_{2k-2,i} + W_{2k-2,2j-1}E_{2j-1,i}$, $i = 2j, 2j+1, \ldots, 2\ell$.
   $W_{2k-2,i} \leftarrow W_{2k-2,i} + W_{2k-2,2j-1}W_{2j-1,i}$, $i = 1, 2, \ldots, 2\ell$.
   end
   end

   Denote $\widetilde{A} - \lambda \widetilde{E} := W\widehat{A}U - \lambda W\widehat{E}U$, $[\, \widetilde{B}_1 \; \widetilde{B}_2 \,] := W[\, \widehat{B}_1 \; \widehat{B}_2 \,]$, $\widetilde{C} := \widehat{C}U$.
3) Compute $\widetilde{F} = [\, F_1 \; 0 \; F_3 \; \cdots \; 0 \; F_{2\ell-1} \; 0 \,]$, where $F_{2j-1} \in \mathbb{R}^{m_2 \times \nu_1^{(j)}}$ are such that $A_{2,0}F_{2j-1} + A_{2,2j-1} = 0$ for $j = 1, \ldots, \ell$.
4) With P defined in (10.101), compute $\breve{A} - \lambda \breve{E} = P(\widetilde{A} + \widetilde{B}_2\widetilde{F} - \lambda \widetilde{E})P^T$, $\breve{B}_1 = P\widetilde{B}_1$, $\breve{C} = (\widetilde{C} + D_2\widetilde{F})P^T$ and $\breve{F} = \widetilde{F}P^T$.
5) Set $\widetilde{X}_1(\lambda) = \big(\breve{A}(1\!:\!n_1, 1\!:\!n_1) - \lambda \breve{E}(1\!:\!n_1, 1\!:\!n_1), \breve{B}_1(1\!:\!n_1, :), \breve{C}(:, 1\!:\!n_1), D_1\big)$ and $Y(\lambda) = \big(\breve{A}(1\!:\!n_1, 1\!:\!n_1) - \lambda \breve{E}(1\!:\!n_1, 1\!:\!n_1), \breve{B}_1(1\!:\!n_1, :), \breve{F}(:, 1\!:\!n_1), 0\big)$.

As stated in Sect. 10.4.1, the reduction of system matrices to the special controllability form at Step 1) can be performed using exclusively orthogonal similarity transformations. It can be shown that the computed condensed matrices $\widehat{A}$, $\widehat{E}$, $\widehat{B}_1$, $\widehat{B}_2$ and $\widehat{C}$ are exact for matrices which are nearby to the original matrices A, E, B_1, B_2 and C respectively. Thus this part of the reduction is *numerically backward stable*.

The computations performed at Step 2), representing the second stage of the special reduction and the computation of the feedback matrix $\widetilde{F}$ at Step 3) involve the solution of many, generally overdetermined, linear equations. Therefore, these steps are generally not numerically stable. In spite of this, the numerical reliability of the overall computations can be guaranteed, as long as W and U, the block upper triangular transformation matrices employed at Step 2), have no excessively large condition numbers. The condition numbers can be approximated as $\kappa(L) \approx \|W\|_F^2$ and $\kappa(V) \approx \|U\|_F^2$. It follows that if these norms are relatively small (e.g., $\leq$10,000) then practically there is no danger for a significant loss of accuracy due to performing non-orthogonal reductions. On contrary, large values of these norms provide a clear hint of potential accuracy losses. In practice, it suffices only to look at the largest magnitudes of the generated elements of W and U at Step 2) to obtain equivalent information. For the computation of $\widetilde{F}$, condition numbers for solving the underlying equations can be also easily estimated. However, a large norm of $\widetilde{F}$ is an indication of possible accuracy losses. For Step 2) of the reduction, a simple operation count is possible by assuming all blocks are 1×1, and this indicates a computational complexity of $\mathcal{O}(n^3)$. Thus, the overall computational complexity of **Procedure GRMCOVER1** is also $\mathcal{O}(n^3)$.

10.4.3 Order Reduction Using Minimum Dynamic Covers of Type II

The computational problem which we address in this section is the following: given the descriptor system pair $(A - \lambda E, B)$ with $A, E \in \mathbb{R}^{n\times n}$ and E invertible, $B \in \mathbb{R}^{n\times m}$, and B partitioned as $B = [\, B_1 \; B_2 \,]$ with $B_1 \in \mathbb{R}^{n\times m_1}$, $B_2 \in \mathbb{R}^{n\times m_2}$, determine the matrices F and G such that the pair $(A + B_2F - \lambda E, B_1 + B_2G)$ has maximal number of uncontrollable eigenvalues.

This computation is useful to determine least-order solutions of linear rational equations using state feedback and feedforward techniques. For the compatible linear rational system of equations $G(\lambda)X(\lambda) = F(\lambda)$, considered also in Sect. 10.4.2, assume there exists a particular solution $X_1(\lambda)$ which is proper. Then, the general solution can be expressed as in (10.91), where $X_2(\lambda)$ is a proper rational basis of the right nullspace of $G(\lambda)$. The proper TFMs $X_1(\lambda)$ and $X_2(\lambda)$ thus generate all solutions of $G(\lambda)X(\lambda) = F(\lambda)$. Assume $X_1(\lambda)$ and $X_2(\lambda)$ have the controllable descriptor realizations in (10.92) with invertible E. Let F be a state feedback gain and let G be a feedforward gain. Then, the TFMs defined as

$$\left[\, \widetilde{X}_1(\lambda) \; \widetilde{X}_2(\lambda) \,\right] := \left[\begin{array}{c|cc} A + B_2F - \lambda E & B_1 + B_2G & B_2 \\ \hline C + D_2F & D_1 + D_2G & D_2 \end{array}\right] \tag{10.103}$$

generate also all solutions. It is straightforward to check that

$$\widetilde{X}_1(\lambda) = X_1(\lambda) + X_2(\lambda)Y(\lambda), \quad \widetilde{X}_2(\lambda) = X_2(\lambda)\widetilde{Y}(\lambda)\,, \tag{10.104}$$

where $Y(\lambda)$ and $\widetilde{Y}(\lambda)$ have the descriptor system realizations

$$\left[\, Y(\lambda)\ \widetilde{Y}(\lambda)\,\right] = \left[\begin{array}{c|cc} A + B_2F - \lambda E & B_1 + B_2G & B_2 \\ \hline F & G & I \end{array}\right]. \tag{10.105}$$

It follows that $\widetilde{X}_1(\lambda)$ is another particular solution, while $\widetilde{X}_2(\lambda)$ is another right nullspace basis, because $\widetilde{Y}(\lambda)$ is invertible. If the gains F and G are determined such that the pair $(A + B_2F - \lambda E, B_1 + B_2G)$ is maximally uncontrollable, then the resulting realizations of $\widetilde{X}_1(\lambda)$ and $Y(\lambda)$ contain a maximum number of uncontrollable eigenvalues which can be eliminated using minimal realization techniques. Thus, $\widetilde{X}_1(\lambda)$ represents another particular solution with a reduced McMillan degree. An important application of the above order reduction technique addressed in Sect. 7.9 is to determine a least-order solution of the EMMP by solving a dual linear rational equation $G(\lambda) = X(\lambda)H(\lambda)$ using the techniques presented in Sect. 10.3.7.

The problem to determine the matrices F and G, which make the descriptor system pair $(A + B_2F - \lambda E, B_1 + B_2G)$ maximally uncontrollable, is essentially equivalent [88] to compute a subspace $\mathcal{V}$ having least possible dimension and satisfying

$$(\overline{A} + \overline{B}_2F)\mathcal{V} \subset \mathcal{V}, \quad \text{span}\,(\overline{B}_1 + \overline{B}_2G) \subset \mathcal{V}\,, \tag{10.106}$$

where $\overline{A} = E^{-1}A$, $\overline{B}_1 = E^{-1}B_1$, and $\overline{B}_2 = E^{-1}B_2$. If we denote $\overline{\mathcal{B}}_1 = \text{span}\ \overline{B}_1$ and $\overline{\mathcal{B}}_2 = \text{span}\ \overline{B}_2$, then the above condition can be equivalently rewritten also as a condition defining a *Type II* minimum dynamic cover [40, 71] of the form

$$\overline{A}\mathcal{V} \subset \mathcal{V} + \overline{\mathcal{B}}_2, \qquad \overline{\mathcal{B}}_1 \subset \mathcal{V} + \overline{\mathcal{B}}_2. \tag{10.107}$$

The computation of the minimal dynamic covers of Type II can be done in two stages using a similar technique as for the Type I covers presented in Sect. 10.4.2. The first stage is identical to the reduction performed for covers of Type I and is performed using **Procedure GSCSF**. Two orthogonal transformation matrices Q and Z are determined, such that the transformed system triple

$$(\widehat{A} - \lambda\widehat{E}, [\,\widehat{B}_2\ \widehat{B}_1\,], \widehat{C}) := (Q^TAZ - \lambda Q^TEZ, [\,Q^TB_2\ Q^TB_1\,], CZ) \tag{10.108}$$

has the pair $(\widehat{A}, \widehat{B})$, with $\widehat{B} = [\,\widehat{B}_2\ \widehat{B}_1, \,]$, in the staircase form (10.89) and the matrix $\widehat{E}$ in the block structured form (10.90). The dimensions of the first 2ℓ diagonal blocks of $\widehat{A}$ and $\widehat{E}$ are determined by the two sets of dimensions $\nu_{1,j}$ and $\nu_{2,j}$ for $j = 1, \ldots, \ell$, and define the dimensions $n_1 := \sum_{j=1}^{\ell} \nu_{1,j}$ and $n_2 := \sum_{j=1}^{\ell} \nu_{2,j}$. Additionally, partition the columns of the resulting $\widehat{C}$ in accordance with the column structure of $\widehat{A}$ in (10.89)

$$\widehat{C} = \left[\, C_1\ C_2\ \cdots\ C_{2\ell-1}\ C_{2\ell}\,\right]. \tag{10.109}$$

In the second reduction stage we use non-orthogonal upper triangular left and right transformation matrices W and U, respectively, to annihilate the minimum number of blocks in $\widehat{A}$ and $\widehat{E}$ which allows to solve the minimum cover problem. Assume W and U have block structures identical to $\widehat{E}$. By exploiting the invertibility of the diagonal blocks $E_{2j,2j}$, we can introduce zero blocks in the block row $2k-1$ of E by annihilating the blocks $E_{2k-1,2j}$, for $j = k, k+1, \ldots, \ell$ of $\widehat{E}$. Similarly, by exploiting the full rank of submatrices $A_{2k-1,2k-3}$, we can introduce zero blocks in the block row $2k-1$ of $\widehat{A}$ by annihilating the blocks $A_{2k-1,2j}$, for $j = k-1, k, \ldots, \ell$. Let $\widetilde{A} := W\widehat{A}U$, $\widetilde{E} := W\widehat{E}U$, $[\,\widetilde{B}_2\ \widetilde{B}_1\,] := W[\,\widehat{B}_2\ \widehat{B}_1\,] = [\,\widehat{B}_2\ \widehat{B}_1\,]$ and $\widetilde{C} = \widehat{C}U$ be the system matrices resulted after this (non-orthogonal) reduction.

Choose the feedforward matrix $G \in \mathbb{R}^{m_2\times m_1}$ such that $A_{1,-1}G + A_{1,0} = 0$ and the feedback matrix $\widetilde{F} \in \mathbb{R}^{m_2\times n}$ partitioned column-wise compatibly with $\widehat{E}$ as

$$\widetilde{F} = [\,0\ F_2\ \cdots\ F_{2\ell-2}\ 0\ F_{2\ell}\ 0\,],$$

where F_{2j} are such that $A_{1,-1}F_{2j} + A_{1,2j} = 0$ for $j = 1, \ldots, \ell$. With the permutation matrix

$$P = \left[\begin{array}{cc|cc|c|cc} 0 & I_{\nu_{2,1}} & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & 0 & I_{\nu_{2,2}} & \cdots & 0 & 0 \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & 0 & 0 & \cdots & 0 & I_{\nu_{2,\ell}} \\ \hline I_{\nu_{1,1}} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & I_{\nu_{1,2}} & 0 & \cdots & 0 & 0 \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & 0 & 0 & \cdots & I_{\nu_{1,\ell}} & 0 \end{array}\right], \tag{10.110}$$

we define $L = PWQ^T$, $V = ZUP^T$ and $F = \widetilde{F}V^{-1}$. Overall we obtain the reduced system $(\breve{A} - \lambda\breve{E}, [\,\breve{B}_2\ \breve{B}_1\,], \breve{C}, [\,\breve{D}_2\ \breve{D}_1\,])$ defined with

$$\begin{aligned} \breve{A} - \lambda\breve{E} &:= L(A + B_2F - \lambda E)V = \left[\begin{array}{c|c} \breve{A}_{11} - \lambda\breve{E}_{11} & \breve{A}_{12} - \lambda\breve{E}_{12} \\ \hline 0 & \breve{A}_{22} - \lambda\breve{E}_{22} \end{array}\right], \\ \big[\,\breve{B}_2\big|\breve{B}_1\,\big] &:= L\big[\,B_2\big|B_1 + B_2G\,\big] = \left[\begin{array}{c|c} 0 & \breve{B}_{12} \\ \hline \breve{B}_{21} & 0 \end{array}\right], \\ \breve{C} &:= (C + D_2F)V = \big[\,\breve{C}_1\big|\breve{C}_2\,\big], \\ \big[\,\breve{D}_2\big|\breve{D}_1\,\big] &:= \big[\,D_2\big|D_1 + D_2G\,\big], \end{aligned} \tag{10.111}$$

where, by construction, the pairs $(\breve{A}_{11} - \lambda\breve{E}_{11}, \breve{B}_{12})$ and $(\breve{A}_{22} - \lambda\breve{E}_{22}, \breve{B}_{21})$ are in controllable staircase form. Thus, by the above choice of F and G, we made n_1 of eigenvalues of the pair $(A + B_2F - \lambda E, B_1 + B_2G)$ uncontrollable. The first n_2 columns V_1 of V, satisfy

$$\overline{A}V_1 = V_1\breve{E}_{11}^{-1}\breve{A}_{11} - \overline{B}_2FV_1, \quad \overline{B}_2G = V_1\breve{E}_{11}^{-1}\breve{B}_{12} - \overline{B}_1$$

and thus, according to (10.107), span a *Type II* dynamic cover of dimension n_2 for the pair $(\overline{A}, [\,\overline{B}_1\ \overline{B}_2\,])$. It can be shown using the results of [71] that the resulting *Type II* dynamic cover $\mathcal{V}$ has minimum dimension.

To illustrate the computational procedure, we consider the reduced system in Example 10.1. First, the following zero blocks are introduced: E_{56}, A_{54}, A_{56}, E_{34}, E_{36}, $A_{3,2}$, A_{34}, A_{36}, E_{12}, E_{14} and E_{16} (in this order). We obtain

$$\left[\widetilde{B}_2\ \widetilde{B}_1 \,\middle|\, \widetilde{A}\right] = \left[\begin{array}{cc|cccccc} A_{1,-1} & A_{1,0} & A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} \\ 0 & A_{2,0} & A_{2,1} & A_{22} & A_{2,3} & A_{24} & A_{2,5} & A_{26} \\ 0 & 0 & A_{31} & 0 & A_{33} & 0 & A_{35} & 0 \\ 0 & 0 & 0 & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} \\ 0 & 0 & 0 & 0 & A_{53} & 0 & A_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & A_{64} & A_{65} & A_{66} \end{array}\right],$$

$$\widetilde{E} = \left[\begin{array}{cccccc} E_{11} & 0 & E_{13} & 0 & E_{15} & 0 \\ 0 & E_{22} & E_{23} & E_{24} & E_{25} & E_{26} \\ 0 & 0 & E_{33} & 0 & E_{35} & 0 \\ 0 & 0 & 0 & E_{44} & E_{45} & E_{46} \\ 0 & 0 & 0 & 0 & E_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & E_{66} \end{array}\right].$$

Additional blocks are zeroed using the feedback $\widetilde{F}$ and feedforward gain G to obtain

$$\left[\widetilde{B}_2\ \widetilde{B}_1 + \widetilde{B}_2 G \,\middle|\, \widetilde{A} + \widetilde{B}_2\widetilde{F}\right] = \left[\begin{array}{cc|cccccc} A_{1,-1} & 0 & A_{11} & 0 & A_{13} & 0 & A_{15} & 0 \\ 0 & A_{2,0} & A_{2,1} & A_{22} & A_{2,3} & A_{24} & A_{2,5} & A_{26} \\ 0 & 0 & A_{31} & 0 & A_{33} & 0 & A_{35} & 0 \\ 0 & 0 & 0 & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} \\ 0 & 0 & 0 & 0 & A_{53} & 0 & A_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & A_{64} & A_{65} & A_{66} \end{array}\right].$$

Finally, after block permutations, we obtained the controllable staircase forms

$$\left[\breve{B}_1 \,\middle|\, \breve{A}_1 - \lambda \breve{E}_1\right] = \left[\begin{array}{c|ccc} A_{2,0} & A_{2,2} - \lambda E_{2,2} & A_{2,4} - \lambda E_{2,4} & A_{2,6} - \lambda E_{2,6} \\ 0 & A_{4,2} & A_{4,4} - \lambda E_{4,4} & A_{4,6} - \lambda E_{4,6} \\ 0 & 0 & A_{6,4} & A_{6,6} - \lambda E_{6,6} \end{array}\right],$$

$$\left[\breve{B}_2 \,\middle|\, \breve{A}_2 - \lambda \breve{E}_2\right] = \left[\begin{array}{c|ccc} A_{1,-1} & A_{1,1} - \lambda E_{1,1} & A_{1,3} - \lambda E_{1,3} & A_{1,5} - \lambda E_{1,5} \\ 0 & A_{3,1} & A_{3,3} - \lambda E_{3,3} & A_{3,5} - \lambda E_{3,5} \\ 0 & 0 & A_{5,3} & A_{5,5} - \lambda E_{5,5} \end{array}\right].$$

The above approach to compute a minimum dynamic cover of Type II is the basis of **Procedure GRMCOVER2**, presented in what follows. This procedure determines, for a pair of generators $(X_1(\lambda), X_2(\lambda))$ with the descriptor realizations given in (10.92), explicit minimal realizations for $\widetilde{X}_1(\lambda)$ and $Y(\lambda)$ (see (10.103) and (10.105)) in the form $\widetilde{X}_1(\lambda) = (\breve{A}_{11} - \lambda \breve{E}_{11}, \breve{B}_{12}, \breve{C}_1, \breve{D}_1)$ and $Y(\lambda) = (\breve{A}_{11} - \lambda \breve{E}_{11}, \breve{B}_{12}, \breve{F}_1, G)$, where $\widetilde{F} P^T =: \left[\,\breve{F}_1\ \breve{F}_2\,\right]$, with $\breve{F}_1$ having n_2 columns.

Procedure GRMCOVER2: Order reduction using dynamic covers of Type II

Inputs : $X_1(\lambda) = (A - \lambda E, B_1, C, D_1)$ and $X_2(\lambda) = (A - \lambda E, B_2, C, D_2)$
Outputs: $\widetilde{X}_1(\lambda) = (\breve{A}_{11} - \lambda \breve{E}_{11}, \breve{B}_{12}, \breve{C}_1, \breve{D}_1)$ and $Y(\lambda) = (\breve{A}_{11} - \lambda \breve{E}_{11}, \breve{B}_{12}, \breve{F}_1, G)$ such that $\widetilde{X}_1(\lambda) = X_1(\lambda) + X_2(\lambda)Y(\lambda)$ has least McMillan degree.

1) Apply **Procedure GSCSF** to the system triple $(A - \lambda E, [\, B_2 \; B_1 \,], C)$ to determine the orthogonally similar system triple $(\widehat{A} - \lambda \widehat{E}, [\, \widehat{B}_2 \; \widehat{B}_1 \,], \widehat{C})$ defined in (10.108) and (10.109), and the dimensions $\nu_{1,j}$ and $\nu_{2,j}$ for $j = 1, \ldots, \ell$; set $n_2 := \sum_{j=1}^{\ell} \nu_{2,j}$.
2) With $\widehat{A}$ partitioned as in (10.89) and $\widehat{E}$ partitioned as in (10.90), perform the second stage of the special reduction for Type II covers:

    Set $W = I_n$, $U = I_n$ and partition W and U in blocks analogous to $\widehat{E}$ in (10.90).

    for $k = \ell, \ell - 1, \ldots, 1$
      Comment. Annihilate blocks $E_{2k-1,2j}$, for $j = k, k+1, \ldots, \ell$.
      for $j = k, k+1, \ldots, \ell$
        Compute $W_{2k-1,2j}$ such that $W_{2k-1,2j}E_{2j,2j} + E_{2k-1,2j} = 0$.
        $A_{2k-1,i} \leftarrow A_{2k-1,i} + W_{2k-1,2j}A_{2j,i}$, $i = 2j-2, 2j-1, \ldots, 2\ell$.
        $E_{2k-1,i} \leftarrow E_{2k-1,i} + W_{2k-1,2j}E_{2j,i}$, $i = 2j, 2j+1, \ldots, 2\ell$.
        $W_{2k-1,i} \leftarrow W_{2k-1,i} + W_{2k-1,2j}W_{2j,i}$, $i = 1, 2, \ldots, 2\ell$.
      end
      if $k > 1$ **then**
        Comment. Annihilate blocks $A_{2k-1,2j}$, for $j = k-1, k, \ldots, \ell$.
        for $j = k-1, k, \ldots, \ell$
          Compute $U_{2k-3,2j}$ such that $A_{2k-1,2k-3}U_{2k-3,2j} + A_{2k-1,2j} = 0$.
          $A_{i,2j} \leftarrow A_{i,2j} + A_{i,2k-3}U_{2k-3,2j}$, $i = 1, 2, \ldots, 2k-1$.
          $E_{i,2j} \leftarrow E_{i,2j} + E_{i,2k-3}U_{2k-3,2j}$, $i = 1, 2, \ldots, 2k-3$.
          $C_{2j} \leftarrow C_{2j} + C_{2k-3}U_{2k-3,2j}$.
          $U_{i,2j} \leftarrow U_{i,2j} + U_{i,2k-3}U_{2k-3,2j}$, $i = 1, 2, \ldots, 2\ell$.
        end
      end if
    end

    Denote $\widetilde{A} - \lambda \widetilde{E} = W\widehat{A}U - \lambda W\widehat{E}U$, $[\, \widetilde{B}_2 \; \widetilde{B}_1 \,] = W[\, \widehat{B}_2 \; \widehat{B}_1 \,]$, $\widetilde{C} = \widehat{C}U$.
3) Compute $\widetilde{F} = [\, 0 \; F_2 \; \cdots \; F_{2\ell-2} \; 0 \; F_{2\ell} \; 0 \,]$, where F_{2j} are such that $A_{1,-1}F_{2j} + A_{1,2j} = 0$ for $j = 1, \ldots, \ell$; compute G such that $A_{1,-1}G + A_{1,0} = 0$.
4) With P in (10.110), compute $\breve{A} - \lambda \breve{E} = P(\widetilde{A} + \widetilde{B}_2\widetilde{F} - \lambda \widetilde{E})P^T$, $\breve{B}_1 = P(\widetilde{B}_1 + \widetilde{B}_2 G)$, $\breve{C} = (\widetilde{C} + D_2\widetilde{F})P^T$, $\breve{D}_1 = D_1 + D_2 G$ and $\breve{F} = \widetilde{F}P^T$.
5) Set $\widetilde{X}_1(\lambda) = \big(\breve{A}(1\!:\!n_2, 1\!:\!n_2) - \lambda \breve{E}(1\!:\!n_2, 1\!:\!n_2), \breve{B}_1(1\!:\!n_2, :), \breve{C}(:, 1\!:\!n_2), \breve{D}_1\big)$ and $Y(\lambda) = \big(\breve{A}(1\!:\!n_2, 1\!:\!n_2) - \lambda \breve{E}(1\!:\!n_2, 1\!:\!n_2), \breve{B}_1(1\!:\!n_2, :), \breve{F}(:, 1\!:\!n_2), G\big)$.

The numerical properties of **Procedure GRMCOVER2** are the same as those of **Procedure GRMCOVER1**, which are discussed in Sect. 10.4.2.

10.4.4 Minimal Realization Using Balancing Techniques

The aim of the algorithm presented in this section is to determine minimal order realizations of stable systems in a descriptor state-space form, by exploiting the concept of balanced realization. For a balanced realization, the controllability and observability properties are perfectly equilibrated. This is expressed by the fact that the controllability and observability gramians are equal and diagonal. The eigenvalues of the gramian of a balanced system are called the *Hankel singular values*. The largest singular value represents the *Hankel norm* of the corresponding TFM of the system, while the smallest one can be interpreted as a measure of the nearness of the system to a non-minimal one. Important applications of balanced realizations are to ensure minimum sensitivity to roundoff errors of real-time filter models or to perform model order reduction, by reducing large order models to lower order approximations. The order reduction can be performed by simply truncating the system state to a part corresponding to the "large" singular values, which significantly exceed the rest of "small" singular values. In what follows we present a procedure to compute minimal balanced realizations of stable descriptor systems. This procedure is instrumental in solving the Nehari approximation problem (see **Procedure GNEHARI** in Sect. 10.4.5).

For a stable state-space system $(A - \lambda E, B, C, D)$ with E invertible, the controllability gramian P and observability gramian Q satisfy appropriate generalized Lyapunov equations. In the continuous-time case P and Q satisfy

$$\begin{aligned} APE^T + EPA^T + BB^T &= 0\,, \\ A^TQE + E^TQA + C^TC &= 0\,, \end{aligned} \tag{10.112}$$

while in the discrete-time case

$$\begin{aligned} APA^T - EPE^T + BB^T &= 0\,, \\ A^TQA - E^TQE + C^TC &= 0\,. \end{aligned} \tag{10.113}$$

Since for a stable system both gramians P and Q are positive semi-definite matrices, in many applications it is advantageous to determine these matrices directly in (Cholesky) factored forms as $P = SS^T$ and $Q = R^TR$, where both S and R can be chosen upper triangular matrices. Algorithms to compute directly these factors have been proposed in [59] for standard systems (i.e., with $E = I$) and extended to descriptor systems in [102]. The following minimal realization procedure proposed in [113] extends to descriptor systems the algorithms proposed in [114] for standard systems. This procedure determines for a stable system $(A - \lambda E, B, C, D)$ the minimal balanced realization $(\widetilde{A} - \lambda I, \widetilde{B}, \widetilde{C}, D)$ and the corresponding balanced diagonal gramian matrix $\widetilde{\Sigma}$. The nonzero Hankel singular values are the decreasingly ordered diagonal elements of $\widetilde{\Sigma}$ and the largest Hankel singular value is $\|G(\lambda)\|_H$, the Hankel norm of the corresponding TFM $G(\lambda) = C(\lambda E - A)^{-1}B + D$.

Procedure GBALMR: Balanced minimal realization of stable systems

Input : $(A - \lambda E, B, C, D)$ such that $\Lambda(A, E) \subset \mathbb{C}_s$
Outputs: Minimal realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, D)$, $\widetilde{\Sigma}$

1) Compute the upper triangular factors S and R such that $P = SS^T$ and $Q = R^T R$ satisfy the appropriate Lyapunov equations (10.112) or (10.113), in accordance with the system type, continuous- or discrete-time.
2) Compute the singular value decomposition

$$RES = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \widetilde{\Sigma} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix},$$

where $\widetilde{\Sigma} > 0$.
3) With the projection matrices $T_l = \widetilde{\Sigma}^{-1/2} U_1^T R$ and $T_r = SV_1 \widetilde{\Sigma}^{-1/2}$, compute the matrices of the minimal realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, D)$ with

$$\widetilde{A} = T_l A T_r, \quad \widetilde{B} = T_l B, \quad \widetilde{C} = C T_r\,.$$

Remark 10.10 The projection matrices satisfy $T_l E T_r = I$ and for a minimal standard system (A, B, C, D) we have $T_l = T_r^{-1}$. The reduction of a linear state-space model to a balanced minimal realization may involve the usage of ill-conditioned transformations (or projections) for systems which are nearly non-minimal or nearly unstable. This is why, for the computation of minimal realizations, the so-called *balancing-free* approaches, as proposed in [126] for standard systems and in [113] for descriptor systems, are generally more accurate. In this case, we can avoid any inversion using at Step 3) the projection matrices $T_l = U_1^T R$ and $T_r = SV_1$ to obtain the descriptor minimal realization $(\widetilde{A} - \lambda \widetilde{E}, \widetilde{B}, \widetilde{C}, D)$ with the invertible $\widetilde{E} = T_l E T_r$. □

10.4.5 Solution of Nehari Problems

In this section we consider the solution of the following optimal Nehari problem: Given $R(\lambda)$ such that $R^{\sim}(\lambda) \in \mathcal{H}_\infty$, find a $Y(\lambda) \in \mathcal{H}_\infty$ which is the closest to $R(\lambda)$ and satisfies

$$\|R(\lambda) - Y(\lambda)\|_\infty = \|R^{\sim}(\lambda)\|_H\,. \tag{10.114}$$

This computation is encountered in the solution of the AMMP formulated in Sect. 9.1.10. As shown in [51], to solve the Nehari problem (10.114), we can solve instead for $Y^{\sim}(\lambda)$ the optimal zeroth-order Hankel-norm approximation problem

$$\|R^{\sim}(\lambda) - Y^{\sim}(\lambda)\|_\infty = \|R^{\sim}(\lambda)\|_H\,. \tag{10.115}$$

In what follows, we only give a solution procedure for the solution of (10.114) in the continuous-time setting. The corresponding procedure for discrete-time systems is much more involved (see [58]) and therefore we prefer the approach based on employing a bilinear transformation as suggested in [51]. To solve the continuous-time Nehari problem (10.114), we solve the optimal zeroth-order Hankel-norm approximation problem to determine $Y(-s)$ such that

$$\|R(-s) - Y(-s)\|_\infty = \|R^\sim(s) - Y^\sim(s)\|_\infty = \|R^\sim(s)\|_H\,. \tag{10.116}$$

The following procedure is a straightforward adaptation of the general Hankel-norm approximation procedure proposed in [51] and [108] for *square* $R(\lambda)$ with poles only in $\mathbb{C}_u$. Assuming $(A - \lambda E, B, C, D)$ is a state-space realization of $R(\lambda)$ (not necessarily minimal), this procedure computes the optimal stable Nehari approximation $Y(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}, \widetilde{D})$.

Procedure GNEHARI: Generalized optimal Nehari approximation

Input : $R(\lambda) = (A - \lambda E, B, C, D)$
Output: $Y(\lambda) = (\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ such that $\|R(\lambda) - Y(\lambda)\|_\infty = \|R^\sim(\lambda)\|_H$.

1) For a discrete-time system employ the bilinear transformation $z = \frac{1+s}{1-s}$:

$$(E, A, B, C, D) \leftarrow (E + A, A - E, \sqrt{2}B, \sqrt{2}C(E + A)^{-1}E, D - C(E + A)^{-1}B)\,.$$

2) Compute using the **Procedure GBALMR** the balanced minimal realization $(\widehat{A}, \widehat{B}, \widehat{C}, D)$ of the system $(-A - sE, -B, C, D)$ and the corresponding diagonal Gramian $\widehat{\Sigma}$ of the balanced system satisfying $\widehat{A}\widehat{\Sigma} + \widehat{\Sigma}\widehat{A}^T + \widehat{B}\widehat{B}^T = 0$ and $\widehat{A}^T\widehat{\Sigma} + \widehat{\Sigma}\widehat{A} + \widehat{C}^T\widehat{C} = 0$.
3) Partition $\widehat{\Sigma}$ in the form $\widehat{\Sigma} = \text{diag}(\sigma_1 I, \widehat{\Sigma}_2)$, such that $\widehat{\Sigma}_2 - \sigma_1 I < 0$ and partition $\widehat{A}$, $\widehat{B}$ and $\widehat{C}$ conformably with $\widehat{\Sigma}$, as

$$\widehat{A} = \begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} \\ \widehat{A}_{21} & \widehat{A}_{22} \end{bmatrix}, \quad \widehat{B} = \begin{bmatrix} \widehat{B}_1 \\ \widehat{B}_2 \end{bmatrix}, \quad \widehat{C} = \begin{bmatrix} \widehat{C}_1 & \widehat{C}_2 \end{bmatrix};$$

compute an orthogonal U such that $U\widehat{B}_1^T = -\widehat{C}_1$.
4) Compute the descriptor system realization $(\widetilde{A} - \lambda\widetilde{E}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ of $Y(s)$ as

$$\begin{aligned}
\widetilde{E} &= \widehat{\Sigma}_2^2 - \sigma_1^2 I\,,\\
\widetilde{A} &= -(\sigma_1^2\widehat{A}_{22} + \widehat{\Sigma}_2\widehat{A}_{22}\widehat{\Sigma}_2 - \sigma_1\widehat{C}_2^T U\widehat{B}_2^T)\,,\\
\widetilde{B} &= -(\Sigma_2\widehat{B}_2 + \sigma_1\widehat{C}_2^T U)\,,\\
\widetilde{C} &= \widehat{C}_2\Sigma_2 + \sigma_1 U\widehat{B}_2^T\,,\\
\widetilde{D} &= D - \sigma_1 U\,.
\end{aligned}$$

5) For a discrete-time system employ the bilinear transformation $s = \frac{z-1}{z+1}$:

$$(\widetilde{E}, \widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) \leftarrow (\widetilde{E} - \widetilde{A}, \widetilde{A} + \widetilde{E}, \sqrt{2}\widetilde{B}, \sqrt{2}\widetilde{C}(\widetilde{E} - \widetilde{A})^{-1}\widetilde{E}, \widetilde{D} + \widetilde{C}(\widetilde{E} - \widetilde{A})^{-1}\widetilde{B})\,.$$

Remark 10.11 If $R(\lambda)$ is not square, then the **Procedure GNEHARI** can be applied to an augmented square $R_a(\lambda)$ formed by adding a sufficient number of zero rows or columns to $R(\lambda)$. From the resulting solution $Y_a(\lambda)$ we obtain the solution $Y(\lambda)$ of the original Nehari problem by removing the rows or columns corresponding to the added zero rows or columns in $R_a(\lambda)$. □

10.5 Numerical Software

Several basic requirements are desirable when implementing software tools for the numerical algorithms discussed in this book. These requirements are

- employing exclusively numerically stable or numerically reliable algorithms;
- ensuring high computational efficiency;
- enforcing robustness against numerical exceptions (overflows, underflows) and poorly scaled data;
- ensuring ease-of-use, high portability and high reusability.

The above requirements have been used for the development of high-performance linear algebra software libraries, such as BLAS, a collection of basic linear algebra subroutines and LAPACK, a comprehensive linear algebra package based on BLAS. These requirements have been also adopted to implement SLICOT, a subroutine library for control theory, based primarily on BLAS and LAPACK. The general-purpose library LAPACK contains over 1300 subroutines and covers most of the basic linear algebra computations for solving systems of linear equations and eigenvalue problems. The specialized library SLICOT[1] contains over 500 subroutines and covers the basic computational problems for the analysis and design of linear control systems. Among the covered problems we mention linear system analysis and synthesis, filtering, identification, solution of matrix equations, model reduction and system transformations. Of special interest for this book is the comprehensive collection of routines for handling descriptor systems and for solving generalized linear matrix equations, as well as the routines for computing Kronecker-like forms. The subroutine libraries BLAS, LAPACK and SLICOT have been originally implemented in the general-purpose language Fortran 77 and, therefore, provide a high level of reusability, which allows their easy incorporation in user-friendly software environments as—for example, MATLAB. In the case of MATLAB, selected LAPACK routines underlie the linear algebra functionalities, while the incorporation of selected SLICOT routines was possible via suitable gateways, as the provided *mex*-function interface.

In what follows, we succinctly describe available software tools in the MATLAB environment, which implement the numerically reliable algorithms discussed in this

[1] The SLICOT software library is freely available from http://www.slicot.org/. The version 4.5 is a free software distributed under the GNU General Public Licence (GPL), while the substantially enriched Release 5.0 is free for academic and non-commercial use.

Table 10.1 SLICOT-based *mex*-functions

Function	Description
sl_gstra	Generalized system similarity transformations
sl_klf	Kronecker-like forms
sl_glme	Generalized linear matrix equations
sl_gzero	Generalized system zeros and Kronecker structure
sl_gminr	Generalized minimal realization
sl_gsep	Generalized additive decompositions

book. First, we have to mention that the basic computational needs to implement the synthesis procedures presented in Chaps. 5 and 6 are covered by the functions of the DESCRIPTOR SYSTEMS Toolbox[2] for MATLAB. This toolbox is a proprietary software, developed by the author in the period 1997–2006 at the German Aerospace Center (DLR). The DESCRIPTOR SYSTEMS Toolbox underlies the implementation of the FAULT DETECTION Toolbox, developed by the author between 2005–2011 at DLR.[3]

To facilitate the implementation of the synthesis procedures described in this book, a new collection of freely available *m*-functions, called the DESCRIPTOR SYSTEM TOOLS, has been implemented by the author. The basic numerical linear algebra support for the implementation of this collection is provided by several LAPACK-based core functions of MATLAB, such as **svd**, **qr**, **schur**, **ordschur**, **qz**, **ordqz**, jointly with a set of *mex*-functions based on SLICOT subroutines. These *mex*-functions are listed in Table 10.1 and implement numerically reliable algorithms with special focus on descriptor system-related computations. These algorithms are described in this chapter and also underlie the implementations of the *m*-functions, which form the collection of DESCRIPTOR SYSTEM TOOLS. The functions of this collection, which are used in this book, are listed in Table 10.2.

The functions implemented in the collection DESCRIPTOR SYSTEM TOOLS use the object-oriented approach provided by the CONTROL Toolbox of MATLAB to handle LTI systems in descriptor system representation. Among the called computational functions, we mention **care** and **dare** for solving generalized continuous-time and discrete-time algebraic Riccati equations, respectively; **norm** for computing system norms; **minreal** to enforce pole–zero cancellations in TFMs; as well as functions for systems coupling, inversion, conjugation, etc.

Several of implemented high-level descriptor systems *m*-functions can be seen as extensions of similar functions provided in the standard CONTROL SYSTEM Toolbox of MATLAB. These are **gpole**, to compute system poles; **gzero**, to compute system zeros; **gir**, to compute irreducible realizations; and **gminreal**, to compute minimal realizations. The functionality of these functions is however richer than that of their counterparts from the CONTROL SYSTEM Toolbox, such as **pole**, **zero**, or

[2]Software distributed by SYNOPTIO GmbH, http://synmath.synoptio.de/en/.

[3]This proprietary software is not distributed outside of DLR.

Table 10.2 Functions of the DESCRIPTOR SYSTEM TOOLS collection used in this book

Function	Description
gpole	System poles and infinite pole structure
gzero	System zeros and Kronecker structure of system pencil
gir	Generalized irreducible realization
gss2ss	Conversion to standard state-space representation
gminreal	Generalized minimal realization
gsorsf	Specially ordered generalized real Schur form
gklf	Generalized Kronecker-like form
glnull	Minimal rational left nullspace basis
gsdec	Generalized additive spectral decomposition
glcf	Generalized left coprime factorization
grcf	Generalized right coprime factorization
glcfid	Generalized left coprime factorization with inner denominator
grcfid	Generalized right coprime factorization with inner denominator
giofac	Generalized inner–outer factorization
goifac	Generalized co-outer–co-inner factorization
glsol	Solution of the linear rational equation $X(\lambda)G(\lambda) = F(\lambda)$
grsol	Solution of the linear rational equation $G(\lambda)X(\lambda) = F(\lambda)$
glmcover1	Left minimum dynamic cover of Type-1 based order reduction of proper systems
grmcover1	Right minimum dynamic cover of Type-1 based order reduction of proper systems
glmcover2	Left minimum dynamic cover of Type-2 based order reduction of proper systems
grmcover2	Right minimum dynamic cover of Type-2 based order reduction of proper systems
gbalmr	Balanced minimal realization of stable generalized systems
ghanorm	Hankel norm of a proper and stable generalized system
gnehari	Generalized optimal Nehari approximation
glsfg	Generalized left spectral factorization of $\gamma^2 I - G(\lambda)G^{\sim}(\lambda)$
glinfldp	Solution of the $\mathcal{L}_\infty$ least distance problem $\min \| F_1(\lambda) - X(\lambda) \;\; F_2(\lambda) \|_\infty$
gsfstab	Generalized state feedback stabilization

minreal. For example, **gpole** computes both the finite and infinite poles (counting multiplicities), while **pole** only computes the finite poles. The function **gzero** computes both the finite and infinite zeros (counting multiplicities) as well as the Kronecker structural invariants of the system pencil, while **zero** only computes the finite zeros. Finally, the functions **gir** and **gminreal** are applicable to a descriptor system model $(A - \lambda E, B, C, D)$ regardless E is singular or nonsingular. In contrast, the function **minreal** can be used only for systems with invertible E (because of the need to explicitly invert E).

Several functions implementing some of the analysis and synthesis procedures presented in Chap. 5 are provided as examples of prototype implementations of dedicated FDI-related software. The three functions listed in Table 10.3 are part of a

Table 10.3 Functions in the FDI TOOLS collection

Function	Description
genspec	Generation of achievable fault detection specifications
efdsyn	Exact synthesis of fault detection filters
efdisyn	Exact synthesis of fault detection and isolation filters

collection called FDI TOOLS (under development) and have been used in solving the case-study examples addressed in Chap. 8.

The collections DESCRIPTOR SYSTEM TOOLS and FDI TOOLS, together with the *m*-files of the synthesis examples presented in the Chaps. 5 and 6 of this book are available from the web address below.[4]

10.6 Notes and References

Section 10.1. The numerical linear algebra aspects related to the SVD, QR decomposition, the real Schur and generalized real Schur decompositions are covered in several textbooks, of which we mention the works of Stewart [112] and of Golub and Van Loan [55]. The latter work, which also contains an up to date list of further references, served for the estimation of the computational efforts in terms of the required number of flops for the basic decompositions considered in Sect. 10.1. The book [60] is a modern reference for roundoff error analysis of floating-point computations. The computation of the controllability and observability staircase forms for standard and descriptor systems using orthogonal similarity transformations is addressed in [116]. The detailed algorithm underlying **Procedure GCSF** has been proposed by the author in [125]. Algorithms for the computation of Kronecker-like forms of linear pencils, using SVD-based rank determinations, and SVD-based row and column compressions, have been proposed in [25, 115]. Albeit numerically reliable, these algorithms have a computational complexity $\mathcal{O}(n^4)$, where n is the minimum of row or column dimensions of the pencil. More efficient algorithms of complexity $\mathcal{O}(n^3)$ have been proposed in [9, 95, 128], which rely on using QR decompositions with column pivoting for rank determinations, and row and column compressions. The **Procedure PREDUCE** is based on the method proposed in [95].

Section 10.2. For a complete coverage of the topic of this section see [110]. The algorithms for the solution of linear matrix equations can be seen as extensions of the Bartels–Stewart method proposed for the solution of the Sylvester equation $AX + BX = C$ in [5]. This algorithm employs the reduction of A and B to RSFs and is considered a numerically reliable method. Further enhancements of this method and extensions to generalized Sylvester equations have been proposed in [54], where one of matrices (that with larger size) is reduced to a Hessenberg form, while the other

[4] https://sites.google.com/site/andreasvargacontact/home/book/matlab.

is reduced to the RSF. Detailed algorithms for the solution of generalized Sylvester matrix equation are described in [47]. Similar algorithms with obvious simplifications can be employed to solve standard and generalized Lyapunov equations. An important algorithm for the solution of Lyapunov equations having positive semi-definite solutions has been proposed in [59], where the solution $X \geq 0$ is directly determined in a Cholesky-factored form $X = SS^T$. The extension of this algorithm to solve generalized Lyapunov equations has been proposed in [102]. The first numerically reliable algorithm to solve standard Riccati equations is the Schur method proposed in [74]. Enhancements of this method to cover discrete-time problems with singular state matrix followed in [100] and to address nearly singular problems in [85, 117]. In all these methods, however, the underlying Hamiltonian or symplectic structure of intervening matrix pencils is not exploited. Therefore, a new direction in developing algorithms for solving GCAREs and GDAREs are the structure exploiting and structure preserving methods to compute eigendecompositions of the Hamiltonian and symplectic pencils (see the book [85] and the recent survey [11]).

Section 10.3. The reliable numerical computation of irreducible realizations of descriptor systems has been considered in [116]. The orthogonal reduction-based algorithm to compute generalized controllability staircase forms, which underlies **Procedure GIR**, has been proposed in [125]. The algorithm to compute a rational nullspace basis of a rational matrix has been proposed in [132] and is related to the approach proposed in [8] to compute polynomial basis using pencil reduction techniques. For the computation of system zeros, an algorithm based on the Kronecker-like form has been proposed in [86]. The approach for the computation of the additive spectral decomposition employed in **Procedure GSDEC** has been proposed in [67]. The iterative pole dislocation techniques underlying the **Procedure GRCF** and **Procedure GRCFID** have been developed in the spirit of the approach described in [118] (see also [129]). Alternative, non-iterative approaches to compute coprime factorizations with inner denominators have been proposed in [94, 96]. The methods presented in Sect. 10.3.6 to compute inner–outer factorizations of full column rank rational matrices are particular versions of the general methods for continuous-time systems proposed in [97] and for discrete-time systems proposed in [94]. The formulas for the complementary inner factors have been derived in [164]. The numerically reliable computational approach for solving linear rational equations, presented in Sect. 10.3.7, has been proposed in [134].

Section 10.4. The algorithm underlying **Procedure GSCSF** to compute the special controllability staircase form, employed in the methods to determine minimum dynamic covers, is a particular instance of the descriptor controllability staircase algorithm of [125]. This algorithm and the computational methods of minimal dynamic covers have been developed in [136]. The minimal realization procedure, based on balancing techniques, has been proposed in [126] for standard systems. The extension of these techniques to descriptor systems has been proposed in [113] and is the basis of **Procedure GBALMR**. The state-space method for the solution of the Nehari problem for continuous-time systems has been developed in [51].

Section 10.5. BLAS is a set of specifications for standard vector and matrix operations, which form the core of implementing numerical algebra algorithms. Three levels of abstraction served to define the functionality of BLAS. Level-1 BLAS basically covers operations with and on vectors [75] and served for the implementation of the widely used linear algebra package LINPACK [35]. Level 2 BLAS for matrix–vector operations [34] and Level-3 BLAS for matrix–matrix operations [33] formed the basic layer for implementing the high-performance linear algebra package LAPACK [3]. This package, originally written in Fortran 77, has been designed to run efficiently on a wide range of high-performance machines using the BLAS, which can be optimized for each computing environment. Moreover, the use of BLAS makes the subroutines portable and efficient across a wide range of computers. The technology for developing, testing and documenting LAPACK has been adopted by the developers of SLICOT [12, 120]. The initial version of the DESCRIPTOR SYSTEMS Toolbox for MATLAB is described in [130] (see also [120]). The first version of the FAULT DETECTION Toolbox is described in [138], while the last version of this toolbox is described in [148].