On minimum norm pole assignment with periodic state feedback

A. Varga

DLR - Oberpfaffenhofen
German Aerospace Research Establishment
Institute of Robotics and System Dynamics
P.O.B. 1116, D-82230 Wessling, Germany.
Fax: +49-8153-28-1441, E-mail: Andreas.Varga@dlr.de

Abstract

A computational approach based on a gradient search method is proposed to determine a minimum norm periodic state feedback assigning a specified set of eigenvalues. An explicit expression of the gradient is derived by using a periodic Sylvester equation based parametrization of the periodic pole assignment problem. The numerical evaluation of gradient is based on efficient numerically stable procedures.

1. Introduction

Consider the linear discrete-time periodic system of the form

$$x_{k+1} = A_k x_k + B_k u_k, \qquad (1)$$

where the matrices $A_k \in \mathbb{R}^{n\times n}$ and $B_k \in \mathbb{R}^{n\times m}$ are periodic with period $K \geq 1$. We consider the following *periodic eigenvalue assignment problem* (PEAP): given the uniformly controllable periodic matrix pair (A_k, B_k), determine the periodic feedback matrix $F_k \in \mathbb{R}^{m\times n}$ such that the eigenvalues of the closed-loop monodromy matrix $\Phi_{A+BF}(K,0) = (A_{K-1}+B_{K-1}F_{K-1})\cdots(A_0+B_0F_0)$ are at desired locations $\Gamma = \{\lambda_1,\ldots,\lambda_n\}$ in the complex plane. We assume that Γ is symmetric with respect to the real axis. This assumption guarantees that the resulting periodic matrix F_k is real.

In the multi-input case or when $K > 1$, the PEAP has nonunique solutions. So it is tempting to try to exploit the nonunicity by imposing additional conditions. One aspect which is desirable from both practical as well as numerical point of view is to determine feedback matrices with as small norm as possible. In this paper we address the problem to determine the minimum Frobenius-norm periodic state feedback which solves the PAEP. For standard systems the minimization of the norm of the feedback matrix assigning a set of eigenvalues has been considered in [1], where an explicit gradient search is performed. This approach has been improved in [2], by eliminating artificial constraints on the problem and by improving tremendously the computational efficiency. Alternative approaches based on search algorithms have been proposed in [2] and [3] in conjunction with parametric variants of the pole asignment algorithms of [4] and [5], respectively.

The proposed computational approach extends the method proposed in [1], by bringing all improvements of that method proposed in [2] into the periodic pole assignment. Explicit expression of the gradient is derived by using a convenient parametrization of the PEAP. The evaluation of the gradient involves the solution of two periodic Sylvester equations. An efficient algorithm based on the use of the periodic Schur form can be employed to solve these equations.

Notation and notational conventions. For a square time-varying matrix A_k, $k = 0,1,\ldots$, we denote $\Phi_A(j,i) = A_{j-1}A_{j-2}\cdots A_i$ for $j > i$ and $\Phi_A(i,i) := I$. If A_k is periodic with period K, then the *monodromy matrix* of the system (1) at time τ is $\Phi_A(\tau+K,\tau)$. Its eigeuvalues are independent of τ and are called *characteristic multipliers*. For an arbitrary periodic matrix X_k of period K we use alternatively the *script* not-

ation $\mathcal{X}$ which associates the block-diagonal matrix $\mathcal{X} = \mathrm{diag}\,(X_0, X_1, \ldots, X_{K-1})$ to the cyclic sequence of matrices X_k, $k = 0, \ldots, K-1$. This notation is consistent with the standard matrix operations. For example the operations with block-diagonal matrices $\mathcal{X} + \mathcal{Y}$, $\mathcal{X}\mathcal{Y}$, $\mathcal{X}^T$, or $\mathcal{X}^{-1}$ can be used to express the addition, multiplication, transposing, and inversion, respectively, performed simultaneously with all individual terms in a sequence of K matrices. We denote with $\sigma\mathcal{X}$ the K-cyclic shift $\sigma\mathcal{X} = \mathrm{diag}\,(X_1, \ldots, X_{K-1}, X_0)$ applied to the cyclic sequence X_k, $k = 0, \ldots, K-1$. The notation $\mathcal{X}_{ij}$ is used to refer simultaneously to *all* (i,j) elements or *all* (i,j) blocks in the cyclic sequence X_k, $k = 0, \ldots, K-1$. This notation also applies in the case of matrix partitioning. For instance the partitioning

$$\mathcal{X} = \begin{bmatrix} \mathcal{X}_{11} & \mathcal{X}_{12} \\ \mathcal{X}_{21} & \mathcal{X}_{22} \end{bmatrix}$$

refers to the same partitioning of all matrices of the cyclic sequence X_k, $k = 0, \ldots, K-1$.

2. Computation of Minimum Norm Feedback

We describe a gradient search based approach to compute a minimum Frobenius-norm periodic feedback F_k to assign a set of characteristic values Γ for the closed-loop monodromy matrix. The proposed approach parallels that described in [1] for standard systems. Several enhancements of this method proposed in [2] for the standard case are also extended to the periodic case.

The proposed approach relies on a straightforward parameterization of the PEAP. Let $G_k \in \mathbb{R}^{m\times n}$ be a given periodic parameter matrix and let $\widetilde{A}_k \in \mathbb{R}^{n\times n}$ be a periodic matrix such that $\Lambda(\Phi_{\widetilde{A}}(K,0)) = \Gamma$. If we determine F_k as $F_k = G_k X_k^{-1}$, where the periodic matrix X_k satisfies the *periodic Sylvester equation* (PSE)

$$A_k X_k - X_{k+1}\widetilde{A}_k + B_k G_k = 0,\ k = 0, \ldots, K\text{–}1, \quad (2)$$

then we have $X_0^{-1}\Phi_{A+BF}(K,0)X_0 = \Phi_{\widetilde{A}}(K,0)$, and thus, F_k is a periodic matrix which solves the PEAP. Usual restrictions on choosing $\widetilde{A}_k$ and G_k are similar to those in the standard case [1]: (1) the periodic pair $(\widetilde{A}_k, G_k)$ is uniformly observable; and (2) $\Lambda(\Phi_A(K,0)) \cap \Lambda(\Phi_{\widetilde{A}}(K,0)) = \emptyset$. If additionally the periodic pair (A_k, B_k) is uniformly controllable, then X_k satisfying (2) is generically nonsingular and for the above F_k, we have $X_{k+1}^{-1}(A_k + B_k F_k)X_k = \widetilde{A}_k$. Although highly redundant, the above parameterization has the advantage to allow the use of standard minimization procedures to compute the minimum norm feedback.

Consider the performance index

$$J = \frac{1}{2}\left(\sum_{k=0}^{K-1}\|F_k\|_F^2\right)^{1/2}$$

to be minimized. With the script notation, J can be expressed alternatively as

$$J = \frac{1}{2}\|\mathcal{F}\|_F = \frac{1}{2}\mathrm{tr}\,(\mathcal{F}^T\mathcal{F}). \quad (3)$$

Using the proposed parameterization, $\mathcal{F}$ is computed as

$$\mathcal{F} = \mathcal{G}\mathcal{X}^{-1}, \quad (4)$$

where $\mathcal{X}$ satisfies the PSE

$$\mathcal{A}\mathcal{X} - \sigma\mathcal{X}\widetilde{\mathcal{A}} + \mathcal{B}\mathcal{G} = 0. \quad (5)$$

The gradient of J with respect to $\mathcal{G}$ can be computed by employing the following result:

Proposition 1 *Let $\mathcal{F}$ be the periodic feedback computed as in (4), assigning the desired characteristic values Γ for given $\widetilde{\mathcal{A}}$ and $\mathcal{G}$. Then, the gradient of J with respect to $\mathcal{G}$ is given by*

$$\nabla_{\mathcal{G}} J = \mathcal{H}^T - \mathcal{B}^T\mathcal{U}^T, \quad (6)$$

where $\mathcal{H} = \mathcal{X}^{-1}\mathcal{F}^T$ and $\mathcal{U}$ satisfies the PSE

$$\widetilde{\mathcal{A}}\mathcal{U} - \sigma\mathcal{U}\sigma\mathcal{A} - \sigma\mathcal{H}\sigma\mathcal{F} = 0. \quad (7)$$

Proof. See Appendix A. □

Thus, to compute the function J and its gradient for a given $\mathcal{G}$, we have to solve two PSEs. Note that the PSE (7) has essentially the same form as equation (5).

Having explicit analytical expressions for the function and its gradient it is easy to employ any gradient based technique to minimize J. However, because the dimension of the minimization problem Knm is potentially large, a particularly well suited class of methods to solve our problem is the class of unconstrained descent methods, as for instance, the limited memory BFGS method [6] used in conjunction with a line search procedure with guaranteed decrease as that described by [7]. Both methods are implemented within the MINPACK-2 project (the successor of MINPACK-1 [8]) offering a convenient reverse communication interface which allows an easy implementation of function and gradient computations.

3. Numerical Aspects

To solve the PSEs (5) and (7) we can freely assume that the periodic matrix $\widetilde{A}_k$ is in a condensed form, as for instance, in *periodic Schur form* (PSF), where $\widetilde{A}_{K-1}$ is in real Schur form and the matrices $\widetilde{A}_k$ for $k = 0, \ldots, K-2$ are upper triangular. According to [9], given the periodic matrix A_k, there exists an orthogonal periodic matrix Z_k, such that the transformed periodic matrix $\widehat{A}_k = Z_{k+1}^T A_k Z_k$ is in PSF. By using this transformation we can simplify the solution of both PSEs. For instance, we determine first the orthogonal $\mathcal{Z}$ such that $\widehat{\mathcal{A}} = \sigma \mathcal{Z}^T \mathcal{A} \mathcal{Z}$ is in PSF. By multiplying the equation (5) with $\mathcal{Z}^T$ from left, one obtains a reduced PSE

$$\widehat{\mathcal{A}}\widehat{\mathcal{X}} - \sigma \widehat{\mathcal{X}} \widetilde{\mathcal{A}} + \widehat{\mathcal{B}}\mathcal{G} = 0, \qquad (8)$$

where $\widehat{\mathcal{X}} = \mathcal{Z}^T \mathcal{X}$ and $\widetilde{\mathcal{B}} = \mathcal{Z}^T \mathcal{B}$. Notice that by this transformation the resulted transformed PSE (8) has exactly the same form as the original one in (5). After solving this equation for $\widehat{\mathcal{X}}$, the solution of (8) results as $\mathcal{X} = \mathcal{Z}\widehat{\mathcal{X}}$. For the solution of the reduced PSE (8) an efficient method can be derived along the line of a more general procedure proposed in [10] (see Appendix B for details). This procedure requires about Kn^3 *floating-point operations* (flops). The whole computation to solve the PSE (5) requires about $N_{PSF} + 3Kn^3 + Kn^2m$ *flops*, where N_{PSF} is the number of *flops* necessary to determine the PSF and to accumulate the performed transformations. As a rough estimate of this value we can take $N_{PSF} = 10Kn^3$. If we count all necessary operations, then each evaluation of the the function and gradient (6) requires about $17Kn^3 + 4Kn^2m$ *flops*. Note that in all above evaluations we assumed $m \ll n$.

We can drastically reduce the cost of gradient evaluation by the following observation [2]. The performance index J is invariant to an orthogonal transformation, that is $J = \frac{1}{2}\mathrm{tr}\,(\widehat{\mathcal{F}}^T \widehat{\mathcal{F}})$, where $\widehat{\mathcal{F}} = \mathcal{F}\mathcal{Z}$ with $\mathcal{Z}$ an orthogonal matrix. If $\mathcal{F}$ is the minimum norm feedback for the pair $(\mathcal{A}, \mathcal{B})$, then $\widehat{\mathcal{F}}$ is the minimum norm feedback for the transformed pair $(\widehat{\mathcal{A}}, \widehat{\mathcal{B}}) = (\sigma \mathcal{Z}^T \mathcal{A} \mathcal{Z}, \sigma \mathcal{Z}^T \mathcal{B})$. Thus we can first reduce $\mathcal{A}$ to PSF which involves only once about $10Kn^3$ operations and then evaluate the gradient for the reduced pair $(\widehat{\mathcal{A}}, \widehat{\mathcal{B}})$ with $\widehat{\mathcal{A}}$ in PSF. In this way, each gradient evaluation involves the solution of two reduced PSEs and requires roughly only $3Kn^3 + 2Kn^2m$ *flops*.

By this approach we can even remove the restrictive condition $\Lambda(\Phi_A(K,0)) \cap \Lambda(\Phi_{\widetilde{A}}(K,0)) = \emptyset$ and allow for partial eigenvalue assignment with a simple computational trick. By using an orthogonal similarity transformation $\widehat{\mathcal{A}} = \sigma \mathcal{Z}^T \mathcal{A} \mathcal{Z}$, we can reduce the matrix $\mathcal{A}$ to an *ordered* PSF (see [9])

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

where the characteristic values of $\mathcal{A}_{11}$ are those corresponding to satisfactory poles (which will be kept unmodified), while the characteristic values of $\mathcal{A}_{22}$ will be moved to desired values. Let us partition

$$\sigma \mathcal{Z}^T \mathcal{B} = \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \end{bmatrix}$$

accordingly. Now we can compute by using gradient search techniques the minimum norm solution $\mathcal{F}_2$ of a reduced order PEAP such that the characteristic values of $\mathcal{A}_{22} + \mathcal{B}_2 \mathcal{F}_2$ are assigned to desired locations. The final minimum norm feedback results as $\mathcal{F} = [\,0 \;\; \mathcal{F}_2\,]\mathcal{Z}^T$.

Remark. The achieved minimum norm of $\mathcal{F}$ also depends on the choice of the matrix $\widetilde{\mathcal{A}}$. Thus, in general there is no guarantee that for arbitrary $\widetilde{\mathcal{A}}$, the resulting minimum norm feedback $\mathcal{F}$ has the least possible norm.

4. Conclusions

A numerical approach to solve the PEAP with minimum norm periodic state feedback has been proposed. By using a convenient parameterization of the PEAP, the computational problem has been formulated as an unconstrained minimization problem having as cost function the Frobenius-norm of the periodic state feedback matrix. Explicit expression for the gradient has been derived and computational issues involved in evaluating the function and gradient have been thoroughly discussed. Of independent importance is an efficient algorithm devised to solve periodic Sylvester equations.

References

[1] L. H. Keel, J. A. Fleming, and S. P. Bhattacharyya. Minimum norm pole assignment via Sylvester's equation. In *Linear Algebra and Its Role in Systems Theory*, volume 47 of *AMS Contemporary Mathematics*, pp. 265–272. 1985.

[2] A. Varga. Parametric methods for pole assignment. *Proc. ECC'97, Brussels*, 1997. (submitted).

[3] P. H. Petkov, M. M. Konstantinov, D. W. Gu, and I. Postletwaite. Optimal eigenstructure assignment of linear systems. Technical report, 93-64, Department of Engineering, Leicester University, UK, 1993.

[4] A. Varga. A Schur method for pole assignment. *IEEE Trans. Autom. Control*, AC-26:517–519, 1981.

[5] P. Hr. Petkov, N. D. Christov, and M. M. Konstantinov. A computational algorithm for pole assignment of linear multiinput systems. *IEEE Trans. Autom. Control*, AC-31:1755–1764, 1986.

[6] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45:503–528, 1989.

[7] J. J. Moré and D. J. Thuente. On line search algorithms with guaranteed sufficient decrease. *ACM Trans. Math. Software*, 20:286–307, 1994.

[8] J. J. Moré. User's Guide for MINPACK-1. Applied Mathematics Division Report ANL-80-74, Argonne National Laboratory, Argone,IL, 1980.

[9] A. W. Bojanczyk, G. Golub, and P. Van Dooren. The periodic Schur decomposition. Algorithms and applications. In F. T. Luk, editor, *Proceedings SPIE Conference*, vol. 1770, pp. 31–42, July 1992.

[10] R. Byers and N. Rhee. Cyclic Schur and Hessenberg-Schur numerical methods for solving periodic Lyapunov and Sylvester equations. Technical report, Dept. of Mathematics, Univ. of Missouri at Kansas City, June 1995.

[11] R. H. Bartels and G. W. Stewart. Algorithm 432: Solution of the matrix equation AX+XB=C. *Comm. ACM*, 15:820–826, 1972.

[12] A. Varga. Periodic Lyapunov equations: some applications and new algorithms. Technical Report TR R197-96, DLR-Oberpfaffenhofen, Institute for Robotics and System Dynamics, February 1996. (to appear in *Int. J. Control*).

Appendix A

To prove Proposition 1 we need the following result.

Lemma 1 *Let A_k, B_k, C_k, D_k be $n \times n$ periodic matrices of period $K \geq 1$. Let $\mathcal{X}$ be the solution of the PSE*

$$\mathcal{A}\mathcal{X} + \sigma\mathcal{X}\mathcal{B} = \mathcal{C}. \tag{9}$$

Then

$$\mathrm{tr}\,[\mathcal{D}\mathcal{X}] = \mathrm{tr}\,[\mathcal{U}\mathcal{C}]\,,$$

where $\mathcal{U}$ satisfies the PSE

$$\mathcal{B}\mathcal{U} + \sigma\mathcal{U}\sigma\mathcal{A} = \sigma\mathcal{D}. \tag{10}$$

Proof. To simplify the notation, let take $\mathcal{Q} = \mathcal{D}^T$. For an arbitrary periodic matrix Y_k of period K we shall also use the notation $\overline{\overline{y}}$ to denote

$$\overline{\overline{y}} = \begin{bmatrix} \mathrm{vec}\,(Y_0) \\ \vdots \\ \mathrm{vec}\,(Y_{K-1}) \end{bmatrix},$$

where the operator vec$(\cdot)$ generates a vector from the stacked columns of a matrix. By using this notation we can express the solution of the PSE (9) as

$$\overline{\overline{x}} = P^{-1}\overline{\overline{c}},$$

where, for $K = 3$, P has the form

$$P = \begin{bmatrix} I \otimes A_1 & B_1^T \otimes I & 0 \\ 0 & I \otimes A_2 & B_2^T \otimes I \\ B_3^T \otimes I & 0 & I \otimes A_3 \end{bmatrix}.$$

With the above expression for $\overline{\overline{x}}$, we obtain successively

$$\mathrm{tr}\,[\mathcal{D}\mathcal{X}] = \mathrm{tr}\,[\mathcal{Q}^T\mathcal{X}] = \overline{\overline{q}}^T\overline{\overline{x}} = \overline{\overline{q}}^T P^{-1}\overline{\overline{c}} = \overline{\overline{y}}^T\overline{\overline{c}},$$

where $\overline{\overline{y}}$ satisfies $P^T\overline{\overline{y}} = \overline{\overline{q}}$. Looking at

$$P^T = \begin{bmatrix} I \otimes A_1^T & 0 & B_3 \otimes I \\ B_1 \otimes I & I \otimes A_2^T & 0 \\ 0 & B_2 \otimes I & I \otimes A_3^T \end{bmatrix},$$

we observe that $P^T\overline{\overline{y}} = \overline{\overline{q}}$ corresponds to the PSE

$$\mathcal{Y}\mathcal{B}^T + \sigma\mathcal{A}^T\sigma\mathcal{Y} = \sigma\mathcal{Q}$$

or, after transposing,

$$\mathcal{B}\mathcal{Y}^T + \sigma\mathcal{Y}^T\sigma\mathcal{A} = \sigma\mathcal{Q}^T.$$

But this is precisely the PSE (10) with the obvious replacements $\mathcal{Y}^T = \mathcal{U}$ and $\mathcal{Q}^T = \mathcal{D}$. We further have

$$\mathrm{tr}\,[\mathcal{D}\mathcal{X}] = \overline{\overline{y}}^T\overline{\overline{c}} = \mathrm{tr}\,[\mathcal{Y}^T\mathcal{C}] = \mathrm{tr}\,[\mathcal{U}\mathcal{C}].$$

□

Proof of Proposition 1. To deduce the expression of the gradient, we compute the first order variation

of J in the form $\Delta J = \mathrm{tr}\,[(\nabla_{\mathcal{G}} J)^T \Delta \mathcal{G}]$. From (3) we have

$$\Delta J = \frac{1}{2}\mathrm{tr}\,[\mathcal{F}^T \Delta \mathcal{F} + \Delta \mathcal{F}^T \mathcal{F}] = \mathrm{tr}\,[\mathcal{F}^T \Delta \mathcal{F}].$$

From (4) we get

$$\Delta \mathcal{F} = \Delta \mathcal{G} \mathcal{X}^{-1} - \mathcal{G}\mathcal{X}^{-1} \Delta \mathcal{X} \mathcal{X}^{-1}$$

and thus

$$\begin{aligned} \Delta J &= \mathrm{tr}\,[\mathcal{F}^T (\Delta \mathcal{G} - \mathcal{F} \Delta \mathcal{X}) \mathcal{X}^{-1}] \\ &= \mathrm{tr}\,[\mathcal{X}^{-1} \mathcal{F}^T (\Delta \mathcal{G} - \mathcal{F} \Delta \mathcal{X})] \\ &= \mathrm{tr}\,[\mathcal{X}^{-1} \mathcal{F}^T \Delta \mathcal{G}] - \mathrm{tr}\,[\mathcal{X}^{-1} \mathcal{F}^T \mathcal{F} \Delta \mathcal{X})]\,. \end{aligned}$$

From (5) follows that $\Delta \mathcal{X}$ satisfies the PSE

$$\mathcal{A} \Delta \mathcal{X} - \sigma \Delta \mathcal{X} \tilde{\mathcal{A}} + \mathcal{B} \Delta \mathcal{G} = 0.$$

By using Lemma 1, we can write

$$\mathrm{tr}\,[\mathcal{X}^{-1} \mathcal{F}^T \mathcal{F} \Delta \mathcal{X})] = \mathrm{tr}\,[\mathcal{U} \mathcal{B} \Delta \mathcal{G}],$$

where $\mathcal{U}$ satisfies the PSE

$$\tilde{\mathcal{A}} \mathcal{U} - \sigma \mathcal{U} \sigma \mathcal{A} - \sigma(\mathcal{X}^{-1} \mathcal{F}^T \mathcal{F}) = 0,$$

which, with $\mathcal{H} = \mathcal{X}^{-1} \mathcal{F}^T$, is in fact the equation (7). We further obtain

$$\Delta J = \mathrm{tr}\,[(\mathcal{H} - \mathcal{U}\mathcal{B}) \Delta \mathcal{G}]$$

from which the expression of the gradient (6) follows. □

Appendix B

Let $A_k \in \mathbb{R}^{n \times n}$, $B_k \in \mathbb{R}^{m \times m}$, $C_k \in \mathbb{R}^{n \times m}$ be periodic matrices of period $K \geq 1$. In this appendix we consider the solution of PSE of the form

$$\mathcal{A}\mathcal{X} + \sigma \mathcal{X} \mathcal{B} = \mathcal{C}, \tag{11}$$

where $\mathcal{X}$ is the periodic solution. The solution method presented here is a direct generalization of the well-known Bartels-Stewart method [11] and has been discussed in a more general setting in [10]. The method presented here can be seen as specialization of the procedure of [10], for the case when both $\mathcal{A}$ and $\mathcal{B}$ are reduced to PSF. Assume $\mathcal{A}$ and $\mathcal{B}$ are already in PSF and partitioned according to their PSF

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} & \cdots & \mathcal{A}_{1\bar{n}} \\ 0 & \mathcal{A}_{22} & \cdots & \mathcal{A}_{2\bar{n}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{A}_{\bar{n}\bar{n}} \end{bmatrix},$$

$$\mathcal{B} = \begin{bmatrix} \mathcal{B}_{11} & \mathcal{B}_{12} & \cdots & \mathcal{B}_{1\bar{m}} \\ 0 & \mathcal{B}_{22} & \cdots & \mathcal{B}_{2\bar{m}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{B}_{\bar{m}\bar{m}} \end{bmatrix}.$$

Let us partition analogously the matrix $\mathcal{X}$

$$\mathcal{X} = \begin{bmatrix} \mathcal{X}_{11} & \mathcal{X}_{12} & \cdots & \mathcal{X}_{1\bar{m}} \\ \mathcal{X}_{21} & \mathcal{X}_{22} & \cdots & \mathcal{X}_{2\bar{m}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}_{\bar{n}1} & \mathcal{X}_{\bar{n}2} & \cdots & \mathcal{X}_{\bar{n}\bar{m}} \end{bmatrix}.$$

From (11) follows that the (r,l)-th block $\mathcal{X}_{rl}$ satisfies the PSE

$$\mathcal{A}_{rr} \mathcal{X}_{rl} + \sigma \mathcal{X}_{rl} \mathcal{B}_{ll} = \mathcal{M}_{rl}, \tag{12}$$

where

$$\mathcal{M}_{rl} = \mathcal{C}_{rl} - \sum_{i=r+1}^{\bar{n}} \mathcal{A}_{ri} \mathcal{X}_{il} - \sum_{j=1}^{l-1} \sigma \mathcal{X}_{rj} \mathcal{B}_{jl}.$$

The above equations can be solved successively for $\mathcal{X}_{\bar{n}1}, \ldots, \mathcal{X}_{21}, \mathcal{X}_{11}, \mathcal{X}_{\bar{n}2}, \ldots, \mathcal{X}_{12}, \ldots, \mathcal{X}_{\bar{n}\bar{m}}, \ldots, \mathcal{X}_{1\bar{m}}$ and this leads to the following procedure:

Algorithm: *Periodic Schur method to solve PSE.*

Compute the orthogonal $\mathcal{U}$ and $\mathcal{V}$ to reduce $\mathcal{A}$ and $\mathcal{B}$, respectively, to a PSF.

$\mathcal{A} \leftarrow \sigma \mathcal{U}^T \mathcal{A} \mathcal{U}$, $\mathcal{B} \leftarrow \sigma \mathcal{V}^T \mathcal{B} \mathcal{V}$, $\mathcal{C} \leftarrow \sigma \mathcal{U}^T \mathcal{C} \mathcal{V}$.

for $l = 1, \ldots, \bar{m}$

  for $r = \bar{n}, \ldots, 1$

    $\mathcal{M} = \mathcal{C}_{rl} - \sum_{i=r+1}^{\bar{n}} \mathcal{A}_{ri} \mathcal{X}_{il} - \sum_{j=1}^{l-1} \sigma \mathcal{X}_{rj} \mathcal{B}_{jl}$

    Solve $\mathcal{A}_{rr} \mathcal{X}_{rl} + \sigma \mathcal{X}_{rl} \mathcal{B}_{ll} = \mathcal{M}$

  end

end

$\mathcal{X} \leftarrow \mathcal{U} \mathcal{X} \mathcal{V}^T$.

This algorithm allows to overwrite $\mathcal{C}$ with the computed solution $\mathcal{X}$. Thus the additional storage necessary to implement this algorithm is $K(n^2 + m^2)$ locations. If we neglect the effort to solve the low order PSEs, then the core algorithm performs about $0.5K(m^2 n + m n^2)$ *flops*. The total number of operations to solve a PSE is about $K(10n^3 + 10m^3 + 2.5n^2 m + 2.5nm^2)$ *flops*.

The computation of the solution $\mathcal{X}_{rl}$ of (12) requires the efficient solution of low order PSEs of the form

$$E_k Y_k + Y_{k+1} F_k = G_k, k = 0, \ldots, K-1;\ Y_0 = Y_K$$

where $E_k \in \mathrm{R}^{n_1 \times n_1}$, $F_k \in \mathrm{R}^{n_2 \times n_2}$ and $G_k \in \mathrm{R}^{n_1 \times n_2}$ with $1 \leq n_1, n_2 \leq 2$. An efficient method to solve such equations is discussed in [12]. The method relies essentially on rewriting the above equations with the help of Kronecker products as a system of $n_1 n_2 K$ simultaneous linear equations $Hy = g$, where the coefficient matrix H is a structured sparse matrix. Ignoring the sparse structure of H in solving $Hy = g$ leads, even for moderate values of K, to rather expensive computations. To exploit the structure of H, we can arrange, by an appropriate grouping of unknowns in the vector y and by a suitable ordering of the equations, to obtain the coefficient matrix H in a block-Hessenberg form. Then a specialized block variant of the Gaussian elimination algorithm is used to solve $Hy = g$. For details see [12].