94-42a

A Note on Closed-Loop Balanced Truncation

Geert Schelfhout and Bart De Moor

Abstract—**Closed-loop balanced truncation is a model reduction method that aims at preserving the closed-loop properties of the original controller. We show that this method is equivalent to frequency-weighted balanced truncation with certain weightings and that for observer-based controllers, a significant amount of computation can be saved because of the observer-state feedback structure of the controller.**

I. Introduction

Many authors have contributed model reduction methods that take into account the preservation of certain desirable closed-loop properties. De Villemagne and Skelton [8] proposed an approach focused on putting the closed-loop controllability gramian in a form that enables one to see which combinations of states are insignificant, but does not employ balancing. This inspired the closed-loop balanced truncation (CLBT) approach, an intuitive but remarkably useful method of controller reduction proposed recently by Ceton *et al.* [3], [9], [10]. In Section II, we show that this method can also be viewed as a frequency-weighted model reduction method with certain weights chosen so as to retain closed-loop performance. Another contribution of this paper is found in Section III, where it is shown how significant savings can be made in the number of computations in the important special case of observer-based controllers. In the remainder of this section, we recapitulate the idea of CLBT in the framework of the standard plant configuration.

Below we employ the notations

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \triangleq C(sI-A)^{-1}B+D \triangleq [A,\, B,\, C,\, D].$$

Suppose the controller $C(s) \triangleq [A_c, B_c, C_c, D_c]$ stabilizes the standard plant

$$P(s) \triangleq \begin{bmatrix} P_{zw}(s) & P_{zu}(s) \\ P_{yw}(s) & P_{yu}(s) \end{bmatrix} \triangleq \left[\begin{array}{c|cc} A & B_w & B_u \\ \hline C_z & D_{zw} & D_{zu} \\ C_y & D_{yw} & D_{yu} \end{array}\right]$$

and yields a closed-loop transfer matrix $H_{zw} = P_{zw} + P_{zu}C(I - P_{yu}C)^{-1}P_{yw}$ which is optimal or at least satisfactory in some sense. Let n be the order of $P(s)$ and n_c the order of $C(s)$. Note that we may assume without loss of generality that $D_{yu} = 0$, since if $P_{yu}(s)$ is not strictly proper, the problem may always be transformed into another one where $P_{yu}(s)$ is strictly proper by means of a loop transformation. In such a case, one state-space representation for H_{zw} is

$$H_{zw}(s) \triangleq \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array}\right] = \left[\begin{array}{cc|c} A + B_u D_c C_y & B_u C_c & B_w + B_u D_c D_{yw} \\ B_c C_y & A_c & B_c D_{yw} \\ \hline C_z + D_{zu} D_c C_y & D_{zu} C_c & D_{zw} + D_{zu} D_c D_{yw} \end{array}\right]. \tag{1}$$

The controllability gramian

$$\tilde{P} \triangleq \begin{bmatrix} P_p & P_{pc} \\ P_{pc}^T & P_c \end{bmatrix}$$

and observability gramian

$$\tilde{Q} \triangleq \begin{bmatrix} Q_p & Q_{pc} \\ Q_{pc}^T & Q_c \end{bmatrix}$$

for this realization of the closed-loop system are computed from $\tilde{A}\tilde{P} + \tilde{P}\tilde{A}^T + \tilde{B}\tilde{B}^T = 0$ and $\tilde{A}^T\tilde{Q} + \tilde{Q}\tilde{A} + \tilde{C}^T\tilde{C} = 0$. To reduce the order of the controller, while changing the closed-loop transfer function as little as possible, CLBT performs truncation on a realization of the controller $C(s)$ in which P_c and Q_c, the parts of these gramians relevant to the controller states, are balanced [3].

This procedure may be rather intuitive and (like many other standard reduction techniques) cannot guarantee stability of the reduced controller or of the closed-loop system after reduction, but it compares very favorably with many well-known controller reduction methods. To provide some further motivation, we proceed to show that CLBT is equivalent to Enns's well-known frequency-weighted balanced truncation (FWBT) with certain performance-oriented weights.

II. Equivalence with Enns's Scheme

Often the inputs w and outputs z of the standard plant are designed in such a way that good performance of the control system is expressed, as well as possible, by the smallness of $\|H_{zw}\|$ in some norm, e.g., the H_2 norm or the H_∞ norm. One may argue, therefore, that a performance-oriented reduced-order controller C_r should make $\|H_{zw}(C_r)\|$ small, where $H_{zw}(C_r) \triangleq P_{zw} + P_{zu}C_r(I - P_{yu}C_r)^{-1}P_{yw}$; this would result in a nonconvex optimization problem over the parameters describing C_r. However, there may be good reasons to resort to controller reduction, seeking to keep $H_{zw}(C_r) - H_{zw}(C)$ small; i.e., $H_{zw}(C_r)$ should approximate the optimal (or satisfactory) closed-loop behavior $H_{zw}(C)$ which contains a lot of information about the controller dynamics one should mimic, rather than approximate zero. One easily finds $H_{zw}(C_r) - H_{zw}(C) \approx P_{zu}(I - CP_{yu})^{-1}(C_r - C)(I - P_{yu}C)^{-1}P_{yw}$; this expression is correct up to first-order terms in $C_r - C$.

Enns's FWBT [1], [4] aims at making $\|W_o(s)[G(s) - G_r(s)]W_i(s)\|$ small, where W_i and W_o are user-specified frequency weights. For performance-oriented controller reduction, by the argument above, a natural choice for these weighting functions is $W_o(s) \triangleq P_{zu}(I - CP_{yu})^{-1}$ and $W_i(s) \triangleq (I - P_{yu}C)^{-1}P_{yw}$.

Manuscript received October 18, 1994; revised February 28, 1995. This work was supported in part by the Flemish Government (Concerted Research Action MIPS: Model-based Information Processing Systems), the National Fund for Scientific Research (G.0292.95 Matrix algorithms and differential geometry for adaptive signal processing, system identification and control), the NFWO Onderzoeksgemeenschappen (ICCoS, Identification and Control of Complex Systems and Advanced Numerical Methods for Mathematical Modeling), the Belgian Government (Interuniversity Attraction Poles IUAP-17: Modeling and Control of Dynamical Systems, IUAP-50: Automation in Design and Production) and the European Commission (Human Capital and Mobility Network SIMONET: System Identification and Modeling Network).

G. Schelfhout is with SIOMAR, B-9042, Ghent, Belgium.

B. De Moor is with ESAT-SISTA, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium. He is also with the NFWO (Belgian National Fund for Scientific Research).

Publisher Item Identifier S 0018-9286(96)06766-9.

The fact that these are derived from a first-order expression also indicates that this choice may be less appropriate for very small reduced orders, where C_r and C can be drastically different.

Enns's scheme with these weightings first computes the controllability gramian $\overline{P}$ of

$$C(s)W_i(s) = \left[\begin{array}{ccc|c} A+B_uD_cC_y & B_uC_c & 0 & B_w+B_uD_cD_{yw} \\ B_cC_y & A_c & 0 & B_cD_{yw} \\ B_cC_y & 0 & A_c & B_cD_{yw} \\ \hline D_cC_y & 0 & C_c & D_cD_{yw} \end{array}\right]$$

and the observability gramian $\overline{Q}$ of

$$W_o(s)C(s) = \left[\begin{array}{ccc|c} A+B_uD_cC_y & B_uC_c & B_uC_c & B_uD_c \\ B_cC_y & A_c & 0 & 0 \\ 0 & 0 & A_c & B_c \\ \hline C_z+D_{zu}D_cC_y & D_{zu}C_c & D_{zu}C_c & D_{zu}D_c \end{array}\right]$$

balances the lower right $n_c \times n_c$ blocks of $\overline{P}$ and $\overline{Q}$ and truncates the controller accordingly.

Replacing the state x_i of the realization of $C(s)W_i(s)$ by T_1x_i, where

$$T_1 \triangleq \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & I & I \end{bmatrix}$$

yields a new realization which is seen, by inspection of (1), to have $\left[\begin{smallmatrix}\tilde{P} & 0\\ 0 & 0\end{smallmatrix}\right]$ for controllability gramian; likewise, replacing the state x_o of the realization of $W_o(s)C(s)$ by $(T_1^{-1})^Tx_o$ yields a realization which has $\left[\begin{smallmatrix}\tilde{Q} & 0\\ 0 & 0\end{smallmatrix}\right]$ for observability gramian. Thus

$$\overline{P} = T_1^T \begin{bmatrix} \tilde{P} & 0 \\ 0 & 0 \end{bmatrix} T_1 = \begin{bmatrix} P_p & P_{pc} & P_{pc} \\ P_{pc}^T & P_c & P_c \\ P_{pc}^T & P_c & P_c \end{bmatrix}$$

and

$$\overline{Q} = T_1^T \begin{bmatrix} \tilde{Q} & 0 \\ 0 & 0 \end{bmatrix} T_1 = \begin{bmatrix} Q_p & Q_{pc} & Q_{pc} \\ Q_{pc}^T & Q_c & Q_c \\ Q_{pc}^T & Q_c & Q_c \end{bmatrix}.$$

Hence, both in FWBT with the weightings above and in CLBT, P_c and Q_c are balanced, and the methods are equivalent; but as the size of the gramians to be computed in CLBT is only $n+n_c$ compared to $n+2n_c$ in FWBT, the former is more efficient. Moreover, with CLBT there are no restrictions on the poles of the controller; these may be unstable or even lie on the imaginary axis. Even more savings can be made when the controller is observer-based, as will be shown in the next section.

III. Observer-Based Controllers

In many control design techniques, a controller is found as the combination of a state feedback controller and a state observer, e.g., when using pole placement or in linear quadratic Gaussian (LQG) design. For this type of controller, the order of the controller equals the order of the plant n; hence, the procedure above would normally require the solution of two $2n \times 2n$ Lyapunov equations. However, in this case, the computation of the gramians P_c and Q_c may be simplified. Suppose the estimator state equation is given by $\dot{x}_c = Ax_c + B_uu + K_e(C_yx_c - y)$ and the state feedback gain is K_r, then the controller has state-space representation $C(s) = [A + B_uK_r + K_eC_y, -K_e, K_r, 0]$, and for the closed-loop system one finds

$$H_{zw}(s) = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & D_{zw} \end{array}\right] = \left[\begin{array}{cc|c} A & B_uK_r & B_w \\ -K_eC_y & A+B_uK_r+K_eC_y & -K_eD_{yw} \\ \hline C_z & D_{zu}K_r & D_{zw} \end{array}\right].$$

With the definitions

$$\begin{aligned} T &\triangleq \begin{bmatrix} I & I \\ 0 & I \end{bmatrix} \\ \tilde{A}_1 &\triangleq T^{-1}\tilde{A}T \\ \tilde{A}_2 &\triangleq T^{-T}\tilde{A}T^T \\ \tilde{B}_1 &\triangleq T^{-1}\tilde{B} \\ \tilde{B}_2 &\triangleq T^{-T}\tilde{B} \\ \tilde{C}_1 &\triangleq \tilde{C}T \\ \tilde{C}_2 &\triangleq \tilde{C}T^T \end{aligned}$$

one has $H_{zw}(s) = [\tilde{A}_1, \tilde{B}_1, \tilde{C}_1, D_{zw}] = [\tilde{A}_2, \tilde{B}_2, \tilde{C}_2, D_{zw}]$. Now let $\tilde{A}_1\tilde{P}_1 + \tilde{P}_1\tilde{A}_1^T + \tilde{B}_1\tilde{B}_1^T = 0$ and $\tilde{A}_2^T\tilde{Q}_2 + \tilde{Q}_2\tilde{A}_2 + \tilde{C}_2^T\tilde{C}_2 = 0$, where $\tilde{P}_1$ and $\tilde{Q}_2$ are partitioned as before

$$\tilde{P}_1 \triangleq \begin{bmatrix} P_{p1} & P_{pc1} \\ P_{pc1}^T & P_{c1} \end{bmatrix}$$

and

$$\tilde{Q}_2 \triangleq \begin{bmatrix} Q_{p2} & Q_{pc2} \\ Q_{pc2}^T & Q_{c2} \end{bmatrix}.$$

Since $\tilde{P} = T\tilde{P}_1T^T$ and $\tilde{Q} = T^{-1}\tilde{Q}_2T^{-T}$, it is seen that $P_c = P_{c1}$ and $Q_c = Q_{c2}$. Therefore CLBT is performed by solving consecutively the six Sylvester equations below

$$(A + K_eC_y)P_{p1} + P_{p1}(A + K_eC_y)^T + (B_w + K_eD_{yw})(B_w + K_eD_{yw})^T = 0 \quad (2)$$

$$(A + K_eC_y)P_{pc1} + P_{pc1}(A + B_uK_r)^T - P_{p1}C_y^TK_e^T - (B_w + K_eD_{yw})(K_eD_{yw})^T = 0 \quad (3)$$

$$(A + B_uK_r)P_{c1} + P_{c1}(A + B_uK_r)^T - K_eC_yP_{pc1} - P_{pc1}^TC_y^TK_e^T + K_eD_{yw}(K_eD_{yw})^T = 0 \quad (4)$$

and

$$(A + B_uK_r)^TQ_{p2} + Q_{p2}(A + B_uK_r) + (C_z + D_{zu}K_r)^T(C_z + D_{zu}K_r) = 0 \quad (5)$$

$$(A + B_uK_r)^TQ_{pc2} + Q_{pc2}(A + K_eC_y) + Q_{p2}B_uK_r + (C_z + D_{zu}K_r)^TD_{zu}K_r = 0 \quad (6)$$

$$(A + K_eC_y)^TQ_{c2} + Q_{c2}(A + K_eC_y) + K_r^TB_u^TQ_{pc2} + Q_{pc2}^TB_uK_r + K_r^TD_{zu}^TD_{zu}K_r = 0 \quad (7)$$

which will yield $P_{c1} = P_c$ and $Q_{c2} = Q_c$.

These computations may be done most efficiently if $A+B_uK_r$ and $A+K_eC_y$ are first put in Schur form, then solving (2)–(7) requires only two Schur decompositions instead of six. The operation counts below are based on the assumption that the computation of a Schur decomposition costs $25n^3$ flops; this value is of course only approximate [5, p. 380]. $2n^3$ flops are needed to solve a Sylvester equation once the coefficient matrices have been put in Schur form (n^3 for a Lyapunov equation). Counting this way, P_{c1} and Q_{c2} may be computed using about $2\times 25n^3+2\times 2n^3+4n^3 = 58n^3$ flops, whereas computing $\tilde{P}$ and $\tilde{Q}$ require $25(2n)^3 + 2(2n)^3 = 216n^3$ flops.

A word of caution is in order w.r.t. the fact that only two Schur decompositions are needed to solve (2)–(7). In principle any Schur decomposition of $A+B_uK_r$ and $A+K_cC_y$ may be used for solving these equations, but for large Sylvester or Lyapunov equations the eigenvalue ordering of the Schur matrices may be important. The QR algorithm naturally yields an ordering with the eigenvalues of the Schur matrix going from large absolute values in the left upper corner to small ones in the right lower corner; this is numerically beneficial for solving (5)–(7), but in equal degree numerically adverse for (2)–(4) if the same Schur decompositions are used without any reordering. Therefore, an inverse ordering, which is done using orthogonal transformations only, may be advisable for (2)–(4) but requires about $6n^3$ extra operations [5], [7], taking the cost to compute P_c and Q_c to $62n^3$ operations. Note, however, that the same remark applies to the computation of $\tilde{P}$ and $\tilde{Q}$; a reordering is equally advisable here, but requires $6(2n)^3 = 48n^3$ operations, adding up to a total cost of $264n^3$ operations.

To avoid the trouble of reordering the Schur decompositions, one might also opt to recompute the Schur decomposition of transposes; in this case our scheme has an operation count of $108n^3$, compared to $416n^3$ for the approach computing $\tilde{P}$ and $\tilde{Q}$.

It should be kept in mind that control design is a highly iterative process so that this controller reduction may have to be repeated tens of times over; accordingly, the savings gain importance.

If the eigenvalue decompositions $A+B_uK_r = T_r^{-1}\Lambda_rT_r$ and $A+K_cC_y = T_c^{-1}\Lambda_cT_c$ are available (e.g., when using a pole placement technique), solving each of the six equations (2)–(7) is done in $O(n^2)$ flops; indeed, these represent similarity transformations that bring $A + B_uK_r$ and $A + K_cC_y$ into diagonal form. The price to be paid for this, however, is a loss of numerical reliability, since these transformations are no longer orthogonal. Note that it is trivial to bring $A+B_uK_r$ and $A+K_cC_y$ into real quasidiagonal form if one wishes to stick to real computations.

The usual way of solving a Lyapunov equation $AP + PA^T + BB^T = 0$ is to bring A into (quasi-)triangular form by means of a Schur decomposition and to solve the transformed Lyapunov equation by means of forward or backward substitution [2], [5]. The splitting of the Lyapunov equations $\tilde{A}_1\tilde{P}_1 + \tilde{P}_1\tilde{A}_1^T + \tilde{B}_1\tilde{B}_1^T = 0$ and $\tilde{A}_2^T\tilde{Q}_2 + \tilde{Q}_2\tilde{A}_2 + \tilde{C}_2^T\tilde{C}_2 = 0$, into three Sylvester equations each, amounts to a block version of this substitution process. We remark that this artifice is not really necessary. If $A + B_uK_r$ and $A + K_cC_y$ are put in real Schur form, $\tilde{A}_1$ and $\tilde{A}_2$ will also have the quasitriangular form suitable for the substitution process. The only advantage of this splitting into three equations is the fact that (2)–(7) may be solved consecutively so that one needs to store at most four $n \times n$ matrices into memory at one time, compared to three $2n \times 2n$ matrices without the splitting. The operation count is identical for both ways, and w.r.t. accuracy, no remarkable differences were recorded.

Once P_{c1} and Q_{c2} are computed, the reduced-order controller may be found. For this purpose, it is not necessary to actually balance P_{c1} and Q_{c2}; instead, one may use, e.g., Safonov's Schur method [6].

For easy reference, the overall procedure is recapitulated here.

- Consecutively solve (2)–(4) and (5)–(7).
- Compute matrices V_r and V_l, whose n_r columns form bases for the right and left eigenspaces of $P_{c1}Q_{c2}$ associated with the r biggest eigenvalues.
- Compute a singular value decomposition $U_e\Sigma_eV_e^T = V_l^TV_r$.
- Defining $L \triangleq \Sigma_e^{-1/2}U_e^TV_l^T$ and $R \triangleq V_rV_e\ \Sigma_e^{-1/2}$, $A_r \triangleq L(A+BK_r+K_eC)R$, $B_r \triangleq -LK_e$, $C_r \triangleq K_rR$, the reduced-order controller is $C_r(s) \triangleq [A_r, B_r, C_r, 0]$.

Evidently this reduced-order controller may be considered as the series connection of a reduced-order observer $[L(A + BK_r + K_eC)T, -LK_e, T, 0]$ and the state feedback gain K_r.

Acknowledgment

The authors wish to thank Prof. P. Van Dooren for some helpful suggestions.

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Control-Linear Quadratic Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1989, ch. 10.
[2] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation $AX + XB = C$," *Comm. ACM*, vol. 15, no. 9, 1972.
[3] C. Ceton, P. M. R. Wortelboer, and O. H. Bosgra, "Frequency-weighted closed-loop balanced reduction," in *Proc. 2nd Eur. Contr. Conf.*, Groningen, The Netherlands, 1993, pp. 697–701.
[4] D. Enns, "Model reduction for control system design," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1984.
[5] G. H. Golub and C. Van Loan, *Matrix Computations*. John Hopkins Univ. Press, 1989.
[6] M. G. Safonov and R. Y. Chiang, "A Schur method for balanced model reduction," in *Amer. Contr. Conf.*, 1988, pp. 1036–1040.
[7] G. W. Stewart, "Algorithm 406: HQR3 and EXCHNG: Fortran subroutines for calculating and ordering the eigenvalues of a real upper Hessenberg matrix," *ACM Trans. Math. Soft.*, vol. 2, pp. 275–280, 1976.
[8] C. De Villemagne and R. E. Skelton, "Controller reduction using canonical interactions," *IEEE Trans. Automat. Contr.*, vol. 33, no. 8, pp. 740–750, 1988.
[9] P. M. R. Wortelboer, "Frequency-weighted balanced reduction of closed-loop mechanical servo-systems: Theory and tools," Ph.D. dissertation, Techn. Univ. Delft, 1994.
[10] P. M. R. Wortelboer and O. H. Bosgra, "Frequency weighted closed-loop order reduction in the control design configuration," in *Proc. 33rd IEEE Conf. Dec. Contr.*, Lake Buena Vista, 1994, pp. 2714–2719.