A CONTROLLER DEGREE BOUND FOR $\mathcal{H}^\infty$-OPTIMAL CONTROL PROBLEMS OF THE SECOND KIND

D. J. N. LIMEBEER† AND G. D. HALIKIAS†

Abstract. This paper is a continuation of our work on $\mathcal{H}^\infty$-optimal control problems which may be embedded in the linear fractional configuration of Fig. 1. In two previous articles [19], [20], a controller degree bound was established for problems in which both $P_{12}(s)$ and $P_{21}(s)$ are square (problems of the first kind). If the McMillan degree of $P(s)$ is $n$, it was shown that there exist $\mathcal{H}^\infty$-optimal controllers with McMillan degree no greater than $n - 1$.

Here we switch our attention to problems of the second kind. That is, we allow $P_{12}(s)$ to have more rows than columns (with $P_{21}(s)$ square), or alternatively, we allow $P_{21}(s)$ to have more columns than rows (with $P_{12}(s)$ square). Our main result shows that the degree bound derived previously for problems of the first kind carries over to problems of the second kind without change. In addition to the controller degree bound, our analysis suggests a number of modifications which are easily made to currently available computer programs [7], [26]. Test calculations (for problems of the second kind) show that these improvements result in a marked reduction in computation time and also enhance the numerical robustness of the software.

Key words. $\mathcal{H}^\infty$-optimal control, approximation theory, cancellations, degree bound, Nehari's theorem

AMS (MOS) subject classification. 93C35

1. Introduction. The generalized regulator in Fig. 1 has been adopted as the standard configuration on which $\mathcal{H}^\infty$-optimal control studies are based. By appropriately choosing the four partitions of $P(s)$, most design examples of engineering interest may be embedded in this diagram. Early $\mathcal{H}^\infty$ studies were special in the sense that $P_{12}(s)$ and $P_{21}(s)$ could be chosen square. Examples of such problems (which we call problems of the first kind) are the optimal sensitivity problem [5], [11], [12], [25], [30], [31], and the robust stabilization problem [15], [16]. Problems of the first kind have now been fully analysed and a controller degree bound has also been found [19], [20]. Although this class of problems admits a particularly elegant and simple solution, it is too special for most practical engineering problems.

If we allow one of the off-diagonal blocks of $P(s)$ to be nonsquare, the range of problems which we may study becomes considerably larger. We call such problems, problems of the second kind. A popular example of which is the so-called mixed-sensitivity problem [8], [9], [10], [13], [27]. As one would expect, this enlarged class of problems is more difficult to analyse as well as being computationally more demanding. In this paper we carry out a detailed analysis of these problems and prove that

---

* Received by the editors February 17, 1987; accepted for publication (in revised form) August 12, 1987.
† Department of Electrical Engineering, Imperial College, Exhibition Road, London, U.K.
the transition from problems of the first kind to problems of the second kind is free in terms of the controller state dimension. We believe that this is a most surprising and encouraging development. Despite advances in microprocessor technology, the potentially high McMillan degree of $\mathcal{H}^\infty$-optimal controllers is an issue which worries engineers. The implementation of high order controllers is not necessarily straightforward. Computation time constraints and finite precision effects are the obvious bugbears.

In a conference paper [21] we derived a bound of $2n - 1$ (where $n = \deg(P)$) for the degree of a class of $\mathcal{H}^\infty$-optimal controllers for problems of the second kind. I. Postlethwaite and his colleagues inform us that several of their computer examples suggest the existence of controllers requiring only $n - 1$ states. Subsequent to their comments, we traced this discrepancy (of $n$ states) to a weak step in our original (unpublished) proof. This oversight has been rectified, and the present paper contains a cancellation analysis of the state-space algorithm described in [7], [26]. This work has not only lead to the degree bound, but it has also suggested a number of improvements to our current software [26].

A recent paper by Ball and Cohen [4] addresses the central model matching problem from a geometric viewpoint. Although their approach is still essentially iterative, it may lead to better computer algorithms and a shorter proof of the controller degree bound. This work may also be helpful in the case of problems of the third kind.

Section 2 contains the notation, a problem description and a brief review of the parametrization and optimization theory. In § 3 we use Riccati equation balancing techniques to derive minimal realizations for the transfer functions associated with the model-matching problem [7], [26]. Section 4 contains the cancellation analysis, the degree bound derivation and some computer time trials. The conclusions are in § 5. All the detailed calculations are contained in appendices at the end of the paper.

2. Notation and background theory.

2.1. Notation.

$\Re, \Re^+, \mathbb{C}$ real, nonnegative and complex numbers;

$\Re(s)$ field of rational functions in $s$ with real coefficients;

$\Gamma_{m \times l}$ set of $m \times l$ matrices with elements in $F (= \Re, \mathbb{C}, \Re(s)$ etc.);

$\mathbb{C}_+, \mathbb{C}_+$ open (respectively, closed) right half-plane;

$\mathbb{C}_-, \mathbb{C}_-$ open (respectively, closed) left half-plane;

$\lambda(A), \lambda_{\max}(A)$ eigenvalue of a square matrix $A$, largest eigenvalue of $A$;

$A^*$ complex conjugate transpose of $A \in \mathbb{C}^{m \times l}$ (transpose if $A \in \Re^{m \times l}$)! $A \succeq 0, A > 0$ $A$ is positive semidefinite (respectively, positive definite);

$A \preceq 0, A < 0$ $A$ is negative semidefinite (respectively, negative definite);

$\mathbb{R}\mathcal{L}^\infty$ space of matrices in $\Re(s)^{m \times l}$ which have no poles on the $j\omega$ axis (including the point at $\infty$);

$\| \cdot \|_\infty$ $\mathcal{L}^\infty$-norm of matrices in $\Re\mathcal{L}^\infty$;

$\mathbb{R}\mathcal{H}_{+}^\infty, \mathbb{R}\mathcal{H}_{-}^\infty$ subspaces of $\Re\mathcal{L}^\infty$; matrices which have no poles in $\mathbb{C}_+$ (respectively, $\mathbb{C}_-$);

$\Gamma_G$ Hankel operator associated with $G(s) \in \Re\mathcal{H}_{+}^\infty$;

$\sigma_i(G(s))$ $i$th Hankel singular value of $G(s)$ (i.e., of $\Gamma_G$) in decreasing order of magnitude;

$\|G(s)\|_H = \sigma_1(G(s))$, the Hankel norm of $G(s)$;

$\Re(s), \bar{s}, |s|$ the real part, complex conjugate and modulus of $s \in \mathbb{C}$;

$G^*(s) = G(-\bar{s})^*$, the para-Hermitian conjugate of $G(s)$;

$\Rightarrow, \Leftarrow, \Leftrightarrow$ implies, is implied by, if and only if.
Associated with a transfer function matrix \( G(s) \in \mathbb{R}(s)^{m \times l} \) of McMillan degree \( n \) is a state-space realisation

\[
G(s) = D + C(sI - A)^{-1}B
\]

where \( A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times l}, C \in \mathbb{R}^{m \times n} \) and \( D \in \mathbb{R}^{m \times l} \). We will use the alternative notation \( G(s) \models (A, B, C, D) \) or

\[
G(s) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.
\]

In the notation above, we have \( G^*(s) \models (-A^*, C^*, -B^*, D^*) \) and in the case that \( D \) is nonsingular, we also have \( G^{-1}(s) \models (A - BD^{-1}C, BD^{-1}, -D^{-1}C, D^{-1}) \). If \( G^{-1}(s) = G^*(s) \), then \( G(s) \) is all-pass. \( G(s) \) is called asymptotically stable if it has no poles in \( \mathbb{C}_+ \).

If \( G(s) \models (A, B, C, D) \), the system matrix of the given realisation is defined as [24]:

\[
\begin{bmatrix} sI - A & -B \\ C & D \end{bmatrix}
\]

and the system zeros are defined to be the points at which the system matrix loses normal rank. In the case when \( D \) is nonsingular, the system zeros are also given by \( \lambda (A - BD^{-1}C) \). The input decoupling zeros (uncontrollable modes) are points at which \( [sI - A|B] \) loses rank. The output decoupling zeros (unobservable modes) are the points at which \( [sI - A^*|C^*] \) loses rank. In the sequel, the term “zero” refers to “system zero” unless stated otherwise. Obviously, \{input decoupling zeros\} \cup \{output decoupling zeros\} are a subset of both \( \lambda (A) \) and the set of system zeros. The realisation \( (A, B, C, D) \) is minimal if it has no input/output decoupling zeros. A sufficient condition for this is that all system zeros are distinct from \( \lambda (A) \).

If \( G_1(s) \models (A_1, B_1, C_1, D_1) \) and \( G_2(s) \models (A_2, B_2, C_2, D_2) \) then the cascade system \( G_1G_2(s) \) has a realisation given by

\[
\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \times \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} = \begin{bmatrix} A_1B_1C_2 + B_2 & B_1D_2 \\ 0 & A_2B_2C_1 + B_1D_2 \end{bmatrix}
\]

where we have taken the “multiplication” of two realisations to mean cascading the two systems. This is not to be confused with ordinary matrix multiplication. The context will always make the distinction between these two possible interpretations clear.

If a basis change \( T \) is introduced into the state space of \( G(s) \), we will take this to mean \( G(s) \models (TAT^{-1}, TB, CT^{-1}, D) \). The McMillan degree of \( G(s) \) will be written as \( \text{deg}(G) \) and the set of poles (zeros) of \( G(s) \) will be denoted \{poles of \( G \}\} \{zeros of \( G \}\}.

Let \( P(s) \) be a partitioned matrix with a state space realisation given by

\[
P(s) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}(s) = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix};
\]

then

\[
P_0(s) = C_1(sI - A)^{-1}B_1 + D_0
\]
is a state-space realisation of $P_2(s)$. A linear fractional transformation for the partitioned matrix $P$ and a matrix $K$ is defined as

$$F_1(P, K) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21}$$

where $K$ is of dimension $l \times m$ if $P_{22}$ has dimension $m \times l$.

When proving the main theorem of this paper we will require a standard result describing the properties of the solutions of algebraic Riccati equations of the form

$$A^*P + PA + PBB^*P + C^*C = 0.$$  

The Hamiltonian matrix associated with (2.5) is

$$H = \begin{bmatrix} A & BB^* \\ -C^*C & -A^* \end{bmatrix}.$$  

**Lemma 2.1.** (i) If $(A, B)$ is stabilisable and $H$ is free of imaginary axis eigenvalues, there exists a unique stabilising solution $P = P^* \geq 0$ to (2.5).

(ii) If $(A, C)$ is observable, every solution $P$ to (2.5) is nonsingular.

**Proof.** The proof of the first part is essentially due to Kučera [17], although Doyle noted that changing the sign of quadratic term in (2.5) did not invalidate Kučera’s proof [7].

To prove the second part, we suppose for contradiction that $P$ is singular. This supposition means that there exists $v \neq 0$ such that $Pv = 0$. Now $v^*(2.5)v \Rightarrow Cv = 0$ and $(2.5)v \Rightarrow PAv = 0$. If $v' = Av$, then $v'^*(2.5)v' \Rightarrow Cv' = CAv = 0$ and $(2.5)v' \Rightarrow PA^2v = 0$. Continuing in this way gives $Cv = 0$, $CAv = 0$, $\cdots$, $CA^{n-1}v = 0$ or $v'[C^*A^*C^*\cdots |A^{(n-1)}|A^*C^*] = 0$ which contradicts the assumed observability of $(A, C)$. Since these arguments apply to any solution, every solution $P$ to (2.5) is nonsingular.

**2.2. Problem description.** The aim of our work is to analyse the cancellation phenomena which occur in the general class of $\mathcal{H}_\infty$ design problems characterised by the assumptions that $P_{21}(s)$ is square while $P_{12}(s)$ has more rows than columns. We also assume that $D_{21}$ is nonsingular, $D_{12}$ has full column rank and that $\text{Re}(\lambda(A - B_1D_{21}C_2)) \neq 0$ and $\text{Re}(\lambda(A - B_2(D_{22}D_{12})^{-1}D_{12}C_1)) \neq 0$. If we consider $\mathcal{R}(s)$, it is easy to see that an equivalent characterisation is given by the assumptions that $P_{12}(s)$ is square while $P_{21}(s)$ has more columns than rows. Any problem fitting either of these alternative descriptions will be called a problem of the second kind. Our analysis will show that certain cancellations are a direct consequence of $\mathcal{H}_\infty$ optimality.

The weighted sensitivity problem [5], [9], [10], [12], [13], [25], [30], [31] is given by

$$W_2(I + GK)^{-1}W_1 = \inf_{K \in \Xi} \|\mathcal{R}(s)\|_\infty$$

where $\Xi$ is the set of stabilizing compensators. By choosing

$$P_2(s) = \begin{bmatrix} W_2W_1 & W_2G \\ W_1 & G \end{bmatrix}$$

we may embed this problem in the generalized regulator configuration in Fig. 1. Since

$$\mathcal{R}(s) = F_1(P, -K) = P_{11} - P_{12}K(I + P_{22}K)^{-1}P_{21},$$

direct comparison with (2.9) reveals that

$$\mathcal{R}_s(s) = F_1(P_s(s), -K(s)).$$
In the case that $G(s)$ is square, this problem falls into the class of problems of the first kind which have been analysed elsewhere [19], [20]. If, on the other hand, $G(s)$ has more outputs than inputs and $W_2(s)$ is square, $P_{12}(s) = W_2 G(s)$ will have more rows than columns giving rise to a problem of the second kind.

Another problem which has received attention is the mixed sensitivity problem [8], [9], [10], [13], [27]. In this case we seek

$$\inf_{K \in \mathcal{R}} \left\| \begin{bmatrix} W_2 G K (I + G K)^{-1} W_1 \\ W_2 (I + G K)^{-1} W_1 \end{bmatrix} \right\|_{\infty} = \inf_{K \in \mathcal{R}} \| R_{ms}(s) \|_{\infty}.$$  

Setting

$$P_{ms}(s) = \begin{bmatrix} 0 & W_2 G \\ W_3 W_1 & W_2 G \\ W_1 & G \end{bmatrix}$$

we have that

$$R_{ms}(s) = F_l(P_{ms}(s), -K(s))$$

which is also an example of a problem of the second kind.

Several other problems of the second kind may be found in the literature. See for example [9], [13] and the numerous references therein. Rather than analyse these problems individually, we have chosen to identify the common characteristics shared by all $\mathcal{H}_\infty$ control problems of the second kind.

2.3. Review of $\mathcal{H}_\infty$ optimisation theory. In this section we will briefly mention the $\mathcal{H}_\infty$ theory which is required in the later analysis. In the next subsection we summarize the Youla parametrisation [6], [29] which is used to characterise the class of all stabilising compensators and the corresponding closed-loop transfer functions $R(s)$ in Fig. 1. Following that, the closed loop transfer functions which have minimum $L^\infty$ norm are identified.

2.3.1. Parametrization of all stabilising controllers. Let $P(s)$ in Fig. 1 be given by

$$P(s) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

and suppose that $(A, B_2, C_2)$ is stabilisable and detectable. We remind the reader that there is no loss of generality in assuming that $D_{21} = I$ and that $D_{12}$ is part of an orthogonal matrix [20], [26]. It is always possible to achieve this by a constant rescaling of the problem. Under these assumptions there exist unique stabilising positive semidefinite solutions to the algebraic Riccati equations [7], [26]

$$X(A - B_2 D_{12}^* C_1) + (A - B_2 D_{12}^* C_1)^* X - X B_2 D_{12}^* X + C_1^* D_{12} D_{12}^* C_1 = 0$$

and

$$X(A - B_1 C_2)^* X + (A - B_1 C_2)^* X - X B_1 D_{21} D_{21}^* X + C_1^* D_{21} D_{21}^* C_1 = 0.$$  

Associated with these two Riccati equations we have the stabilizing matrices $F$ and $H$ given by [7], [20], [26],

$$F = D_{12}^* C_1 + B_2^* X$$

and

$$H = B_1 + X C_2^*.$$
The Youla parametrization theory allows us to write \([6], [7], [23], [26], [29]\)

\[
(2.19) \quad \mathcal{R}(s) = [T_{11} - T_{12}XT_{21}](s)
\]

and

\[
(2.20) \quad K(s) = F_i(K_0(s), X(s))
\]

in which

\[
(2.21) \quad K_0(s) = \begin{bmatrix}
A - B_2F - HC_2 + HD_{22}F & -H - B_2 - HD_{22} \\
0 & I \\
C_2 - D_{22}F & I & D_{22}
\end{bmatrix}
\]

and

\[
(2.22) \quad \begin{bmatrix}
T_{11} & T_{12} & T_{11} \\
T_{21} & 0 & 0
\end{bmatrix}
\begin{bmatrix}
A - B_2F & B_2F & 0 \\
0 & A - HC_2 & -\bar{\mathcal{X}}_0^*C_1^*D_\perp \\
C_1 - D_{12}F & D_{12}F & D_{11} & D_{12} & D_{\perp}
\end{bmatrix}
\begin{bmatrix}
B_1 \\
B_2 \\
-\bar{\mathcal{X}}_0^*C_1^*D_\perp
\end{bmatrix}
\begin{bmatrix}
B_1 \\
B_2 \\
-\bar{\mathcal{X}}_0^*C_1^*D_\perp
\end{bmatrix}
\]

\(\bar{\mathcal{X}}_0^*\) is the Moore-Penrose generalized inverse of \(\bar{\mathcal{X}}\). With the particular choice of the matrices \(F\) and \(H\) given in (2.17) and (2.18), \(T_{21}\) and \([T_{12}|T_{11}]\) are inner \([7], [26]\). We call \(T_{11}(s)\) an inner or all-pass extension of \(T_{12}(s)\).

\[2.3.2.\text{ The } \gamma\text{-iteration.}\] In this section we will outline an algorithm \([7], [8], [9], [10], [13], [26], [27]\) which reduces the problem of finding an upper bound for

\[
(2.23) \quad \inf_{X(s) \in \mathcal{R}^{\infty}} \|T_{11} - T_{12}XT(s)\|_\infty = \gamma_{\text{opt}}
\]

to the Nehari problem \([22]\) of identifying those matrices in \(\mathcal{R}^{\infty}\) which are closest (in the sense of the \(L^\infty\)-norm) to a given point (matrix) in \(\mathcal{R}^{\infty}\). The approach will be to generate a sequence \(\gamma_i = \|R_i(s)\|_\infty = \|T_{11} - T_{12}X_iT_{21}\|_\infty\) which converges on \(\gamma_{\text{opt}}\) (from above). In applications, the calculation of this sequence is terminated when an upper bound \(\alpha\) for \(\gamma_{\text{opt}}\) has been found such that \((\alpha - \gamma_{\text{opt}}) < \varepsilon\) for some sufficiently small \(\varepsilon > 0\).

It may be seen from (2.19) that

\[
(2.24) \quad \mathcal{R}(s) = \begin{bmatrix}
T_{11} & T_{12} & T_{11} \\
T_{21} & 0 & 0
\end{bmatrix}\begin{bmatrix}
X(s) \\
0
\end{bmatrix}
\begin{bmatrix}
T_{21} \\
0
\end{bmatrix}
\]

Since \([T_{12}|T_{11}]\) and \(T_{21}(s)\) are inner and thus norm-preserving, we can write

\[
(2.25) \quad \|\mathcal{R}(s)\|_\infty = \begin{bmatrix}
(T_{12}^*T_{11}^*T_{21}^* - X(s)) \\
T_{11}^*T_{21}^*(s)
\end{bmatrix}
\]

and a routine direct calculation from (2.22) shows that

\[
R(s) = \begin{bmatrix}
R_1 \\
R_2
\end{bmatrix}
\]

\[
(2.26) \quad \begin{bmatrix}
-A - B_2F & (C_1 - D_{12}F)^*D_{11} + \bar{\mathcal{X}}B_1C_2 \\
0 & -(A - HC_2)^* \\
\end{bmatrix}
\begin{bmatrix}
(C_1 - D_{12}F)^*D_{11} + \bar{\mathcal{X}}B_1 \\
-\bar{\mathcal{X}}_0^*C_1^*D_\perp \\
D_{12}^*C_1\bar{\mathcal{X}}_0^* \\
-\bar{\mathcal{X}}_0^*D_{11}C_2 \\
D_{12}^*D_{11}
\end{bmatrix}
\]

\(\Rightarrow R(s)\) is completely unstable.
We also observe that if
\[(2.27) \quad \|R(s)\|_\infty \leq \gamma\]
then
\[(2.28) \Rightarrow (R_1 - X)^*(s)(R_1 - X)(s) \leq \gamma^2 I - R^*_2(s)R_2(s) = M^*(s)M(s)\]
where \(M(s)\) is a stable and minimum-phase spectral factor of \(\gamma^2 I - R^*_2(s)R_2(s)\) (the existence of which requires \(\gamma \geq \|R_2(s)\|_\infty\)). Continuing, we deduce from (2.28) that
\[(2.29a) \quad \|R_1M^{-1}(s) - XM^{-1}(s)\|_\infty = \mu \leq 1.\]
Decomposing \(R_1M^{-1}(s)\) into stable and unstable parts gives
\[(2.29b) \quad \|[R_1M^{-1}(s)]_- - \tilde{X}(s)\|_\infty = \mu \leq 1\]
in which
\[(2.30) \quad \tilde{X}(s) = [R_1M^{-1}(s)]_+ + XM^{-1}(s).\]
Rearranging leads to
\[(2.31) \quad X(s) = (\tilde{X}(s) - [R_1M^{-1}(s)]_+)M(s).\]
We observe that (2.29) is a Nehari or optimal approximation problem [14], [22].

The specific algorithm for the \(\gamma\)-iteration is essentially a binary search procedure which is described elsewhere [7], [9], [10], [13], [26], [27].

2.3.3. Characterization of all solutions to the Nehari approximation problem. The purpose of the \(\gamma\)-iteration, which was mentioned in the last subsection, was to reduce the problem in (2.23) to a Nehari problem. That is, the problem of finding all \(X(s)\)s which achieve
\[(2.32) \quad \inf_{X \in \mathcal{N}_\infty} \|H(s) - X^*(s)\|_\infty\]
or else which satisfy
\[(2.33) \quad \|H(s) - X^*(s)\|_\infty \leq \rho > \|H(s)\|_H.\]
Glover [14] has shown that all the solutions to these problems may be characterised in terms of a balanced realisation of \(H(s)\). In [14], this characterisation is in terms of a linear fractional transformation which contains a free matrix contraction. In [20], we give a different version of these results which characterize all the solutions in terms of a bounded real type condition. We will use this characterisation to establish the main theorem in § 4. We refer the reader to Theorem 2.1 and Corollary 2.2, together with Remarks 2.1-2.4 in [20].

3. Balancing the Riccati equations. Our subsequent analysis is greatly simplified if the Riccati equations (2.15) and (2.16) have a diagonal balanced structure. As we have already established [20], this procedure allows us to remove the right half-plane zeros of the square off-diagonal block of \(P(s)\) in the early stages of the analysis. A similar dimension deflation corresponding to the nonsquare off-diagonal block is also possible. In addition, the balanced Riccati equations lead directly to a minimal realisation for
\[
\begin{bmatrix}
T_{11} & T_{12} & T_{13} \\
T_{21} & 0 & 0
\end{bmatrix}(s)
\]
(review (2.22)).
We begin by considering any basis change $T$ in the state-space of $P(s)$ in (2.14). In the new basis, the realization of $P(s)$ becomes

$$P(s) = \begin{bmatrix} TAT^{-1} & TB_1 & TB_2 \\ C_1T^{-1} & D_{11} & D_{12} \\ C_2T^{-1} & D_{21} & D_{22} \end{bmatrix}$$

and the algebraic Riccati equation (2.15) becomes

$$\xi(TAT^{-1} - TB_2D^*_2C_1T^{-1}) + (TAT^{-1} - TB_2D^*_2C_1T^{-1})*\xi - \xi TB_2B^*_2T^*\xi + T^{-*}C^*_1D_1D^*_1C_1T^{-1} = 0$$

where $T^{-*}$ denotes $(T^{-1})^*$. After pre-multiplication by $T^*$ and post-multiplication by $T$ we get

$$T^*\xi T(A - B_2D^*_2C_1) + (A - B_2D^*_2C_1)^*T^*\xi T$$

$\xi T^*T + C^*_1D_1D^*_1C_1 = 0$

which shows that $\xi$ undergoes the congruence transformation

$$\xi \rightarrow T^{-*}\xi T^{-1}.$$ 

Similarly, in the new basis, (2.16) becomes [20]:

$$T^{-1}y^*(A - B_1C_2)^* + (A - B_1C_2)T^{-1}y^*T^{-*} - T^{-1}y^*T^{-*}C^*_2C_2T^{-1}y^*T^{-*} = 0$$

and hence

$$y^* \rightarrow Ty^*T^*.$$ 

Formulae (3.3) and (3.5), together with $\xi = \xi^* \geq 0$ and $y^* = y^* \geq 0$, show that the construction in [14, App. B] may be used to select a $T$ so that

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix}$$

and

$$y = \begin{bmatrix} \xi_1 \\ 0 \\ \xi_3 \\ 0 \end{bmatrix}.$$ 

It is convenient to introduce a permutation matrix $J$ so that

$$JyJ^* = \begin{bmatrix} \xi_1 \\ \xi_3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \Sigma_2 \\ 0 \end{bmatrix}.$$ 

Clearly, $\Sigma_1 > 0$ and $\Sigma_2 > 0$.

In the rest of the analysis we will assume that the realization in (2.14) has been put into a basis corresponding to balanced Riccati equations. We continue by defining

$$M := J(A - B_1C_2)J^*,$$

$$Z := A - B_2D^*_2C_1$$
and introduce the partitioning

\begin{align}
C_1 &= [C_{11} | C_{12}], \\
[B_1 | B_2] &= \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}
\end{align}

which is consistent with that in (3.6). We will also make use of the partitioning

\begin{align}
C_2 J^* &= [\hat{C}_{21} | \hat{C}_{22}]
\end{align}

and

\begin{align}
F J^* &= [\hat{F}_1 | \hat{F}_2]
\end{align}

which is consistent with that in (3.8). Allowing (3.6) to induce a partitioning on \( Z \) and substituting into (2.15) gives

\begin{align}
\begin{bmatrix} 
\Sigma_1 & 0 \\
0 & \Sigma_2 \\
\end{bmatrix}
\begin{bmatrix} 
Z_{11} & Z_{12} \\
Z_{21} & Z_{22} \\
\end{bmatrix}
+ 
\begin{bmatrix} 
Z_{11}^* & Z_{21}^* \\
Z_{12}^* & Z_{22}^* \\
\end{bmatrix}
\begin{bmatrix} 
\Sigma_1 & 0 \\
0 & \Sigma_2 \\
\end{bmatrix}
\end{align}

\begin{align}
= & -\begin{bmatrix} 
\Sigma_1 & 0 \\
0 & \Sigma_2 \\
\end{bmatrix}
\begin{bmatrix} 
B_{12} \\
B_{22} \\
\end{bmatrix}
\begin{bmatrix} 
B_{12}^* & B_{22}^* \\
\end{bmatrix}
\begin{bmatrix} 
\Sigma_1 & 0 \\
0 & \Sigma_2 \\
\end{bmatrix}
\end{align}

\begin{align}
+ 
\begin{bmatrix} 
C_{11}^* \\
C_{12}^* \\
\end{bmatrix}
D_1 D_2^*[C_{11}, C_{12}] = 0.
\end{align}

From the (2, 2) block of (3.15) we get

\begin{align}
C_{12}^* D_1 D_2^* C_{12} = 0 \Rightarrow D_2^* C_{12} = 0 \Rightarrow C_{11}^* D_1 D_2^* C_{12} = 0.
\end{align}

Consequently, from the (1, 2) block of (3.15) we obtain

\begin{align}
\Sigma_1 Z_{12} = 0 \Rightarrow Z_{12} = 0 \quad \text{(since } \Sigma_1 > 0)\).
\end{align}

Finally, the (1, 1) block gives

\begin{align}
\Sigma_1 Z_{11} + Z_{11}^* \Sigma_1 - \Sigma_1 B_{12} B_{12}^* \Sigma_1 + C_{11}^* D_1 D_2^* C_{11} = 0
\end{align}

which is a deflated Riccati equation with a positive definite solution \( \Sigma_1 \). Since

\begin{align}
A - B_2 F = Z - B_2 B_2^* \tilde{x}
\end{align}

we get from (3.17) that

\begin{align}
A - B_2 F = \begin{bmatrix} 
Z_{11} & 0 \\
Z_{21} & Z_{22} \\
\end{bmatrix}
- \begin{bmatrix} 
B_{12} \\
B_{22} \\
\end{bmatrix}
\begin{bmatrix} 
B_{12}^* & B_{22}^* \\
\end{bmatrix}
\begin{bmatrix} 
\Sigma_1 & 0 \\
0 & \Sigma_2 \\
\end{bmatrix}
\end{align}

\begin{align}
= \begin{bmatrix} 
Z_{11} - B_{12} B_{12}^* \Sigma_1 & 0 \\
Z_{21} - B_{22} B_{12}^* \Sigma_1 & Z_{22} \\
\end{bmatrix}
\end{align}

Since \((A - B_2 F)\) is asymptotically stable, \( Z_{22} \) is also. We conclude also that each eigenvalue of \( Z_{22} \) corresponds to a stable mode of \( A - B_2 D_{22}^* C_1 \) which is undetectable through \( D_{22}^* C_1 \).

A similar procedure has already been applied to (2.16) \[20\] to obtain

\begin{align}
J(A - H C_2) J^* = \begin{bmatrix} 
-\Sigma_2 M_{11}^* \Sigma_2^{-1} & M_{12} - \Sigma_2 \hat{C}_{21} \hat{C}_{22} \\
0 & M_{22} \\
\end{bmatrix}
\end{align}

in which \( \{\lambda(M_{11})\} = \{\text{right half-plane zeros of } P_{21}\} \).
Making use of (3.12), (3.14), (3.16), (3.20), (3.21) and (2.15), we can rewrite (2.22) as

\[
\begin{bmatrix}
T_{11} & T_{12} & T_{13} \\
T_{21} & 0 & 0
\end{bmatrix}
\]

(3.22)

We also note that the change of basis

\[
T = \begin{bmatrix}
\Sigma_1^{1/2} & 0 \\
0 & \Sigma_1^{-1/2}
\end{bmatrix}
\]

will balance the realisation in (3.22). From the balanced version of (3.22) we obtain

\[
T_{21}(s) = \begin{bmatrix}
-\Sigma_1^{1/2}M_{11}\Sigma_1^{-1/2} & -\Sigma_1^{1/2}\hat{C}_{21} \\
\hat{C}_{21}\Sigma_1^{-1/2} & I
\end{bmatrix}
\]

(3.24)

which has the identity as both its controllability and observability gramians. We conclude, therefore, that (3.24) is both a minimal and a balanced realisation of \(T_{21}(s)\). We also deduce that

\[
\begin{bmatrix}
\Sigma_1^{1/2}(Z_{11} - B_{12}B_{12}^*\Sigma_1)\Sigma_1^{-1/2} \\
D_{12}^*D_{12}^*C_{11}^{-1} - D_{12}B_{12}^*\Sigma_1^{-1/2}
\end{bmatrix}
\]

is balanced with controllability and observability gramians the identity; this realisation is thus also minimal.

We conclude this section by pointing out that the realisation in (3.22) is also minimal. This may be established by proving that all the system zeros in (3.22) lie in the open right half-plane and consequently cannot cancel any of the poles of this realisation which lie in the left half-plane. An almost identical argument may be found in § 3 in [20].

As one would expect, replacing the realisation (2.22) with (3.22) allows the realisation in (2.26) to be reduced to

\[
R(s) = \begin{bmatrix}
-(Z_{11} - B_{12}B_{12}^*\Sigma_1)^* & A_R(1, 2) \\
-\Sigma_1^{1/2} & B_R(1, 1)
\end{bmatrix}
\]

(3.26)

where

\[
A_R(1, 2) = C_{11}^*D_{11}^*D_{11}^*\hat{C}_{21} + \Sigma_1B_{11}\hat{C}_{21} - \Sigma_1B_{12}D_{12}^*D_{11}\hat{C}_{21}
\]

and

\[
B_R(1, 1) = C_{11}^*D_{11}^*D_{11}^* - \Sigma_1B_{12}D_{12}^*D_{11} + \Sigma_1B_1.
\]

The realisation in (3.26) need not be minimal. The results of the analysis of this section are summarised in the following lemma.
LEMMA 3.1.

(i) (a) The number of zeros of $P_{21}(s)$ in $C_+ = \text{rank (}\mathcal{Y})$

(b) The number of stable modes of $A - B_2D^*_1C_1$ which are undetectable through $D^*_1C_1 = \text{rank (}\mathcal{X})$

(ii) The realization $(3.22)$ is minimal with McMillan degree $= \text{rank (}\mathcal{X}) + \text{rank (}\mathcal{Y})$

(iii) The realization $(3.24)$ is minimal and $\text{deg (}T_{21}\text{)} = \text{rank (}\mathcal{Y})$

(iv) The realization $(3.25)$ is minimal and $\text{deg ([}T_{12}\mid T_{21}\text{])} = \text{rank (}\mathcal{X})$

(v) $\text{deg (}R\text{)} = \text{rank (}\mathcal{X}) + \text{rank (}\mathcal{Y})$ (see 3.26).

4. Main results. In this section we combine the results we have already obtained with a new theorem to obtain the controller degree bound for all problems of the second kind. We will be treating both the optimal and suboptimal cases.

Suppose $n = \text{deg (}P\text{)}$, $t = \text{deg (}\mathcal{R}\text{)}$ and let $c = (\text{number of cancellations which occur between } P(s) \text{ and } K(s) \text{ as a result of closing the feedback loop in Fig. 1})$. Then

$$t = n + \text{deg (}K\text{)} - c,$$

that is,

$$\text{deg (}K\text{)} \leq t_b + c_b - n$$

where $t_b$ and $c_b$ are upper bounds on $t$ and $c$, respectively. We will derive the upper bound $t_b$ in §4.1 while $c_b$ will be found in §4.2. These results will be combined in §4.3 to give our main theorem.

4.1. An upper bound for the McMillan degree of all closed loop systems of the second kind. Our derivation of the bound $t_b$ for the degree of the closed loop requires several steps. Before stating and proving the main theorem of this section we will briefly sketch the route we intend to take:

(a) The reader will recall from Lemma 3.1(ii) that

$$\text{deg }\begin{bmatrix} T_{11} & T_{12} & T_{21} \\ T_{21} & 0 & 0 \end{bmatrix} = \text{rank (}\mathcal{X}) + \text{rank (}\mathcal{Y}).$$

(b) If $T_{11}$ and $T_{21}$ are all-pass right coprime, and $T_{11}$ and $[T_{12}\mid T_{21}]$ are all-pass left coprime, we have shown that [20, Thm. 4.1]

$$\text{deg }\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = \text{deg }\begin{bmatrix} T^*_1 & T^*_1 & T^*_1 \\ T^*_2 & T^*_1 & T^*_1 \end{bmatrix} = \text{rank (}\mathcal{X}) + \text{rank (}\mathcal{Y}).$$

$T_{11}$ and $T_{21}$ will be called all-pass right coprime if in

$$\begin{bmatrix} T_{11} \\ T_{21} \end{bmatrix}(s) = \begin{bmatrix} \hat{T}_{11} \\ \hat{T}_{21} \end{bmatrix}(s)A(s)$$

all-all pass common right divisors $A(s)$ of $T_{11}(s)$ and $T_{12}(s)$ are constant orthogonal matrices. All-pass left coprimeness is defined in a similar way.

(c) We will assume throughout that the all-pass coprimeness condition is satisfied. If this is not the case, there is always a factorisation

$$\begin{bmatrix} T_{11} & T_{12} & T_{21} \\ T_{21} & 0 & 0 \end{bmatrix}(s) = \begin{bmatrix} A_1(s) & 0 & 0 \\ 0 & I & \hat{T}_{11} \\ 0 & 0 & \hat{T}_{21} \end{bmatrix}\begin{bmatrix} A_r(s) & 0 \\ 0 & I \end{bmatrix}$$

in which $\hat{T}_{11}$ and $\hat{T}_{21}$ are all-pass right coprime, and $\hat{T}_{11}$ and $[\hat{T}_{12}\mid \hat{T}_{21}]$ are all-pass left coprime [20, Thm. 4.1]. For reasons which are almost identical to those given in [20], the existence or nonexistence of these all-pass common factors makes no difference to the bound $t_b$ that we seek.
(d) Next, it will be proved that

\[
\text{deg } ([R, M^{-1}]_{\ast}) \leq \text{rank } (\bar{x}) + \text{rank } (\bar{y})
\]

We remind the reader that the definition of $M(s)$ may be found in (2.28). In the hypothesis of Theorem 4.1 we will assume that (4.7) is met with equality since this assumption simplifies one of our later calculations.

(e) Section 4.1(d), together with the work of Glover [14], ensures that

\[
\text{deg } (\bar{X}) \leq \text{rank } (\bar{x}) + \text{rank } (\bar{y}) + \text{deg } (U) - 1
\]

in the optimal case, and

\[
\text{deg } (\bar{X}) \leq \text{rank } (\bar{x}) + \text{rank } (\bar{y}) + \text{deg } (U)
\]

in the suboptimal case. In the above $U(s) \in \mathcal{H}_\infty$ is a free matrix contraction to be chosen by the designer and $\bar{X}(s)$ is defined in (2.29b).

(f) We will prove that

\[
\text{deg } (X) \leq \text{deg } (\bar{X})
\]

and that

\[
\text{deg } (\tilde{X}) \leq \text{deg } (\bar{X}).
\]

$X(s)$ and $\bar{X}(s)$ are related in (2.31). Consequently,

\[
t_b = \text{rank } (\bar{x}) + \text{rank } (\bar{y}) + \text{deg } (U) - 1
\]

in the optimal case, or

\[
t_b = \text{rank } (\bar{x}) + \text{rank } (\bar{y}) + \text{deg } (U)
\]

in the suboptimal case. In the first instance the reader may wish to skip to §§ 4.2–4.4. In this way we may get an initial overview without getting swamped in the details surrounding the proofs of claims (d)-(f) above; these details are considerable.

The general form of the state space model for

\[
\begin{bmatrix}
T_{11} & T_{12} & T_{12} \\
T_{21} & 0 & 0
\end{bmatrix}(s)
\]

(in (3.22)) is the basis of the hypothesis for our next theorem.

**Theorem 4.1.**

Let

\[
\begin{bmatrix}
T_{11} & T_{12} & T_{12} \\
T_{21} & 0 & 0
\end{bmatrix}(s) = \begin{bmatrix}
A_{11} & A_{12} & B_{11} & B_{12} & B_{13} \\
0 & A_{22} & B_{21} & 0 & 0 \\
C_{11} & C_{12} & D_{11} & D_{12} & D_{13} \\
0 & C_{22} & I & 0 & 0
\end{bmatrix}
\]

be asymptotically stable and suppose also that

(i) $T_{21}(s) = \begin{bmatrix} A_{22} & B_{21} \\ C_{22} & I \end{bmatrix}$

is all-pass, minimal and balanced;

(ii) $[T_{12} T_{12}^\dagger](s) = \begin{bmatrix} A_{11} & B_{12} & B_{13} \\ C_{11} & D_{12} & D_{13} \end{bmatrix}$
is all-pass, minimal and balanced;

(iii) \( T_{11}(s) \) and \( T_{21}(s) \) are all-pass right coprime;

(iv) \( T_{11}(s) \) and \( [T_{12} \mid T_{12}] \) are all-pass left coprime;

(4.13) (v) \( D_{13}^* C_{12} = 0; \)

(4.14) (vi) \( A_{12} + C_{12}^* C_{12} = 0. \)

Then

(4.15) \[
\begin{bmatrix}
R_1 \\
R_2
\end{bmatrix} = \begin{bmatrix}
-A_{11}^* & -C_{11}^* D_{11} B_{21}^* - B_{11} B_{21}^*
0 & -A_{22}^*
-B_{12}^* & -D_{12}^* (D_{11} B_{21}^* + C_{12})
-B_{13}^* & -D_{13}^* D_{11} B_{21}^*
\end{bmatrix}
\begin{bmatrix}
C_{11}^* D_{11} + B_{11} & C_{22}^* \\
D_{13}^* D_{11} & D_{13}^* D_{11}
\end{bmatrix}
\]

is a minimal balanced realisation.

(4.16) (b) \( (R_1 M^{-1})_* = \begin{bmatrix} -\bar{A}^* & F^* E^{1/2} \\
-\bar{B}_1^* & 0 \end{bmatrix} \)

where the matrices \( E \) and \( F \) are defined in the proof; see (4.41) and (4.45), respectively.

(c) If \( (R_1 M^{-1})_* = \begin{bmatrix} -\bar{A}^* & F^* E^{1/2} \\
-\bar{B}_1^* & 0 \end{bmatrix} \)

is minimal and

\[
\begin{bmatrix}
\tilde{X}(s) \equiv \\
\hat{X}
\end{bmatrix} = \begin{bmatrix}
\hat{A} & \hat{B} \\
\hat{C} & \hat{D}
\end{bmatrix}
\]

is chosen so that

(4.17) \[
[(R_1 M^{-1})_* + \tilde{X}] (s) = \begin{bmatrix}
-\bar{A}^* & 0 \\
0 & \hat{A} \\
-\bar{B}_1^* & \hat{C} \\
\hat{B} & \hat{D}
\end{bmatrix} = \begin{bmatrix}
\tilde{A} & \tilde{B} \\
\tilde{C} & \tilde{D}
\end{bmatrix}
\]

satisfies the bounded real-type equations

(4.18) \[
\begin{bmatrix}
-(\bar{A} P + P \bar{A}^* + \bar{B} \bar{B}^*) & -(\bar{B} \bar{D}^* + P \bar{C}^*) \\
-(\bar{D} \bar{B}^* + \bar{C} P) & I - \bar{D} \bar{D}^*
\end{bmatrix} = \begin{bmatrix}
L & \begin{bmatrix} L^* & W^* \end{bmatrix}
\end{bmatrix}
\]

and their duals

(4.19) \[
\begin{bmatrix}
-(\bar{A}^* Q + Q \bar{A} + \bar{C}^* \bar{C}) & -(\bar{C}^* \bar{D} + Q \bar{B}) \\
-(\bar{D}^* \bar{C} + \bar{B}^* Q) & I - \bar{D}^* \bar{D}
\end{bmatrix} = \begin{bmatrix}
L^*_d & \begin{bmatrix} L_d & W_d \end{bmatrix}
\end{bmatrix}
\]

in which

(4.20) (i) \( Q P = I \)

and

(4.21) (ii) \( L^* = [0 \ L_{21}^*]; \quad L_d = [0 \ L_{21d}] \)
where the partitioning in (4.21) is conformal with that in (4.17).

Then

(4.22)  (1) \{\text{poles of } X\} \subseteq \{\text{poles of } \tilde{X}\},

(4.23)  (2) \{\text{poles of } T_{11} + T_{12}X T_{21}\} \subseteq \{\text{poles of } \tilde{X}\}.

Explicit formulae for \(X(s)\) and \([T_{11} + T_{12}X T_{21}](s)\) are given in (4.59) and (4.71) below.

**Remark 4.1.** The validity of the bounds in (4.11) is proven by applying Theorem 4.1 to a balanced version of (3.22). In this regard we should note the following:

(a) The general form of the realisations in (3.22) and (4.12) is the same.

(b) \(T_{21}\) and \([T_{12}|T_{13}]\) are both inner and their realizations are minimal and balanced.

(c) Equations (4.13) and (4.14) are easily seen to be satisfied after balancing (3.22). This will also be true should it be necessary to extract all-pass common factors.

(d) Theorem 2.1 and Cor. 2.2 in [20] ensure that the error systems corresponding to any Nehari or suboptimal extension of \([R_{1}M^{-1}]_{-}\) will satisfy the bounded real-type equations (4.18) and (4.19).

(e) Theorem 4.1 and Lemma 3.1(ii) thus establish the validity of (4.7)-(4.11).

**Remark 4.2.** In the case that the final value of \(\gamma > \gamma_{\text{opt}}\), the cancellation phenomena predicted by Theorem 4.1 can only be guaranteed if \(X(s)\) is a suboptimal extension of \((R_{1}M^{-1})_{-}(s)\) corresponding to an error system with an infinity norm of one.

**Proof.** In the interests of clarity, we have relegated long calculations and the treatment of certain technical, details to a sequence of appendices. The appendices and the main body of the proof will share common notation.

The assumed properties of the realisation of \(T_{21}(s)\) enforces

(4.24) \[ A_{22} + A_{22}^{*} + B_{21} B_{21}^{*} = 0 \]

and

(4.25) \[ C_{22} = -B_{21}^{*}. \]

Similarly, the realisation of \([T_{12}|T_{13}]\) must satisfy

(4.26) \[ A_{11} + A_{11}^{*} + B_{12} B_{12}^{*} + B_{13} B_{13}^{*} = 0, \]

(4.27) \[ A_{11} + A_{11}^{*} + C_{11} C_{11}^{*} = 0, \]

(4.28) \[ D_{12} C_{11} + B_{12}^{*} = 0, \]

(4.29) \[ D_{12}^{*} C_{11} + B_{12}^{*} = 0, \]

(4.30) \[ D_{12} B_{13} + D_{13} B_{13}^{*} + C_{11} = 0, \]

(4.31) \[ [D_{12}|D_{13}][D_{12}|D_{13}]^{*} = I. \]

The reader may wish to consult Glover [14, Thm. 5.1] for a state-space characterisation of all-pass matrices.

From (4.12) and (4.25) we get

\[
T_{11} T_{21}^{*} = \begin{bmatrix}
A_{11} & A_{12} & B_{11} \\
0 & A_{22} & B_{21} \\
C_{11} & C_{12} & D_{11}
\end{bmatrix}
\begin{bmatrix}
-A_{22} & B_{21} \\
B_{21} & I
\end{bmatrix}
\begin{bmatrix}
A_{11} & A_{12} & B_{12} B_{21}^{*} & B_{11} \\
0 & A_{22} & B_{21} B_{21}^{*} & B_{21} \\
0 & 0 & -A_{22} & B_{21} \\
C_{11} & C_{12} & D_{11} B_{21}^{*} & D_{11}
\end{bmatrix}
\]
Introducing the change of basis

\[ T = \begin{bmatrix} I & 0 & 0 \\ 0 & I & -I \\ 0 & 0 & I \end{bmatrix} \]

and making use of (4.24) gives

\[
T_{11} T_{21}^* = \begin{bmatrix}
A_{11} & A_{12} & A_{12} + B_{11} B_{21}^* & B_{11} \\
0 & A_{22} & 0 & 0 \\
0 & 0 & -A_{22}^* & B_{21} \\
C_{11} & C_{12} & D_{11} B_{21}^* + C_{12} & D_{11}
\end{bmatrix}
\]

(4.32)

Next,

\[
\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} (s) = \begin{bmatrix} T_{12}^* \\ T_{12} \end{bmatrix} T_{11} T_{21}^* (s) = \begin{bmatrix}
-A_{11}^* & C_{11}^* & C_{11}^* (D_{11} B_{21}^* + C_{12}) & C_{11}^* D_{11} \\
-B_{12}^* & D_{12}^* & D_{12}^* (D_{11} B_{21}^* + C_{12}) & D_{12}^* D_{11} \\
-A_{22}^* & D_{22}^* C_{11} & D_{22}^* (D_{11} B_{21}^* + C_{12}) & D_{22}^* D_{11} \\
-B_{13}^* & D_{13}^* C_{11} & D_{13}^* (D_{11} B_{21}^* + C_{12}) & D_{13}^* D_{11}
\end{bmatrix}
\]

by (4.13). Introducing the change of basis

\[ T = \begin{bmatrix} I & I & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \]

and invoking (4.14), (4.27)–(4.29) gives

\[
\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} (s) = \begin{bmatrix}
-A_{11}^* & -C_{11}^* D_{11} B_{21}^* - B_{11} B_{21}^* & C_{11}^* D_{11} + B_{11} \\
0 & -A_{22}^* & -B_{21} \\
-B_{12}^* & -D_{12}^* (D_{11} B_{21}^* + C_{12}) & D_{12}^* D_{11} \\
-B_{13}^* & -D_{13}^* D_{11} B_{21}^* & D_{13}^* D_{11}
\end{bmatrix}
\]

(4.33)

which is the same as (4.15). The minimality of this realisation follows from assumptions (i)–(iv) together with Limebeer and Hung [20, Thm. 4.1]. This completes the proof of (a).
We begin the proof of the remainder of the theorem by writing down the equations describing the spectral factorization of \((\gamma^2 I - R_2^* R_2)(s)\). Clearly,

\[
(\gamma^2 I - R_2^* R_2)(s) = \begin{bmatrix}
\bar{A} & \bar{B}_2 \bar{B}_2^* \\
0 & -\bar{A}^* \\
-\bar{C} & -\bar{D}_2 \bar{B}_2^*
\end{bmatrix} \gamma^2 I - \bar{D}_2 \bar{D}_2^*.
\]

Our first step is to carry out the decomposition

\[
(\gamma^2 I - R_2^* R_2)(s) = Z(s) + Z^*(s)
\]

in which \(Z(s)\) is positive real. Since \(\bar{A}\) is asymptotically stable, there exists a unique \(\Theta = \Theta^* \geq 0\) which satisfies the Lyapunov equation

\[
\bar{A}\Theta + \Theta \bar{A}^* + \bar{B}_2 \bar{B}_2^* = 0.
\]

For our later convenience we will also introduce the partitioning

\[
\Theta = \begin{bmatrix}
\Theta_{11} & \Theta_{12} \\
\Theta_{12}^* & \Theta_{22}
\end{bmatrix}
\]

which is conformable with that of \(\bar{A}\) in (4.33).

The change of basis in the state space of (4.34) gives

\[
Z(s) = \begin{bmatrix}
\bar{A} & N, -\bar{C}, \frac{1}{2}E
\end{bmatrix}
\]

in which

\[
N = \bar{B}_2 \bar{D}_2^* + \Theta \bar{C}^* = \begin{bmatrix}
N_1 \\
N_2
\end{bmatrix}
\]

where the partitioning is also induced by that in (4.33). Finally,

\[
E = \gamma^2 I - \bar{D}_2 \bar{D}_2^*.
\]

The spectral factor \(M(s)\) may now be expressed in terms of the unique positive definite solution of the Riccati equation [1]

\[
Y(\bar{A} + NE^{-1} \bar{C}) + (\bar{A} + NE^{-1} \bar{C})^* Y + YNE^{-1} N^* Y + \bar{C}^* E^{-1} \bar{C} = 0.
\]

The next lemma shows that the required solution to (4.42) always exists.

**Lemma A.** The Riccati equation (4.42) always has a unique positive definite stabilizing solution.

*Proof.* See Appendix A.

As with (4.37), the solution to (4.42) has a partitioning induced by (4.33):

\[
Y = \begin{bmatrix}
Y_{11} & Y_{12} \\
Y_{12}^* & Y_{22}
\end{bmatrix}
\]

Since \(Y\) is stabilising,

\[
M(s) = \begin{bmatrix}
\bar{A}, N, -E^{1/2} F, E^{1/2}
\end{bmatrix}
\]

in which

\[
F = E^{-1}(\bar{C} + N^* Y) = [F_1 | F_2]
\]
is a minimum phase spectral factor (it is easy to check that \( M^*(s)M(s) = Z(s) + Z^*(s) \)).
\( \gamma > \| R_2 \|_\infty \) ensures the positive definiteness of \( E \) and therefore that

\[
M^{-1}(s) = \left[ \begin{array}{cc}
\bar{A} + NF & NE^{-1/2} \\
F & E^{-1/2}
\end{array} \right]
\]

is proper.

We will now carry out the calculations which lead to state space realisations of \( R_1M^{-1}(s) \), \( [R_1M^{-1}(s)]_+ \), and \( [R_1M^{-1}(s)]_- \).

\[
R_1M^{-1}(s) = \left[ \begin{array}{cc}
-\bar{A}^* & \bar{C}^* \\
-\bar{B}^*_1 & \bar{D}^*_1
\end{array} \right] \left[ \begin{array}{cc}
\bar{A} + NF & NE^{-1/2} \\
F & E^{-1/2}
\end{array} \right]
\]

The change of basis

\[
T = \left[ \begin{array}{cc}
I & Y \\
0 & I
\end{array} \right]
\]

in the state space of (4.47) together with (4.42) and (4.45) yields

\[
R_1M^{-1}(s) = \left[ \begin{array}{cc}
-\bar{A}^* & 0 \\
0 & \bar{A} + NF \\
-\bar{B}^*_1 & \bar{D}^*_1F + \bar{B}^*_1Y
\end{array} \right] \left[ \begin{array}{cc}
F & E^{1/2} \\
NE^{-1/2} & E^{-1/2}
\end{array} \right]
\]

In (4.48) we note that \(-\bar{A}^*\) is completely unstable while \(\bar{A} + NF\) is asymptotically stable. Thus

\[
R_1M^{-1}(s)_- = \left[ \begin{array}{cc}
-\bar{A}^* & F^{*E^{1/2}} \\
0 & \bar{A} + NF
\end{array} \right] \left[ \begin{array}{cc}
F & \bar{E}^{1/2} \\
NE^{-1/2} & E^{-1/2}
\end{array} \right]
\]

and

\[
[R_1M^{-1}(s)]_+ = \left[ \begin{array}{cc}
\bar{A} + NF & \bar{D}^*_1F + \bar{B}^*_1Y, \bar{D}^*_1E^{-1/2}
\end{array} \right]
\]

The remainder of the proof is based on detailed manipulations requiring various partitions of the bounded real-type equations (4.18) and (4.19). The \((1,1)\) and \((2,1)\) blocks of (4.18) may be written out in full as

\[
\left[ \begin{array}{ccc}
-A_{11}^* & -(C_{11}^*D_{11} + B_{11})B_{21}^* & 0 \\
0 & -A_{22} & 0 \\
0 & 0 & \hat{A}
\end{array} \right]
\left[ \begin{array}{ccc}
P_{11} & P_{12} & P_{13} \\
P_{12} & P_{22} & P_{23} \\
P_{13} & P_{23} & P_{33}
\end{array} \right]
\]

\[
\left[ \begin{array}{ccc}
P_{11} & P_{12} & P_{13} \\
P_{12} & P_{22} & P_{23} \\
P_{13}^* & P_{23} & P_{33}^*
\end{array} \right]
\left[ \begin{array}{ccc}
-A_{11} & 0 & 0 \\
0 & -B_{21}(D_{11}^*C_{11} + B_{11}^*) & -A_{22} \\
0 & 0 & \hat{A}^*
\end{array} \right]
\]

\[
\left[ \begin{array}{ccc}
F_{1}^{*E^{1/2}} \\
F_{2}^{*E^{1/2}} \\
\hat{B}
\end{array} \right]
\left[ \begin{array}{ccc}
E^{1/2}F_1 & E^{1/2}F_2 & \hat{B}^* \\
E^{1/2}F_2 & E^{1/2}F_1 & \hat{B}
\end{array} \right] = \left[ \begin{array}{ccc}
0 & 0 & 0 \\
0 & 0 & L_{21}^*
\end{array} \right] = 0
\]
and

\[ (4.52) \quad \hat{D}[E^{1/2}F_1, E^{1/2}F_2, \hat{B}^{*}] + [-\hat{B}_{11}^{*}, -\hat{B}_{21}^{*}, \hat{C}] \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{12} & P_{22} & P_{23} \\ P_{13} & P_{23} & P_{33} \end{bmatrix} \]

\[ + \begin{pmatrix} 0 \\ 0 \end{pmatrix} L_{21}^{*} = 0. \]

In the same way the (1, 1) and (2, 1) blocks of (4.19) may be written out as

\[ (4.53) \quad \begin{bmatrix} \begin{bmatrix} -A_{11} & 0 & 0 \\ -B_{21}(D_{11}^{*}C_{11} + B_{11}^{*}) & -A_{22} & 0 \\ 0 & 0 & \hat{A}^{*} \end{bmatrix} & \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{bmatrix} \end{bmatrix} \]

\[ + \begin{bmatrix} -\hat{B}_{11}^{*} \\ -\hat{B}_{21}^{*} \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 0 \end{bmatrix} L_{21}^{*} = 0. \]

(4.54)

\[ \hat{D}^{*}[-\hat{B}_{11}^{*}, -\hat{B}_{21}^{*}, \hat{C}] + [E^{1/2}F_1, E^{1/2}F_2, \hat{B}^{*}] \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{bmatrix} \]

\[ + \begin{pmatrix} 0 \\ 0 \end{pmatrix} L_{21}^{*} = 0. \]

An easy rearrangement shows that (4.42) may be written in the alternative form

\[ (4.55) \quad Y\hat{A} + \hat{A}^{*}Y + F^{*}EF = 0 \]

which together with (4.51) yields

\[ (4.56) \quad \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{12}^{*} & Y_{22} \end{bmatrix} = -\begin{bmatrix} P_{11} & P_{12} \\ P_{12}^{*} & P_{22} \end{bmatrix}. \]

The (1, 1) block of (4.53) together with (4.26), (4.36) and (4.37) yields

\[ (4.57) \quad I + Q_{11} = \Theta_{11}. \]

Next, we have

\[ X(s) = (\hat{\times} - (R_{1}M^{-1})_{s})M(s) \]

\[
\begin{bmatrix}
\hat{A} & 0 \\
0 & \hat{A} + NF \\
\hat{C} & \hat{D}^{*}F + \hat{B}^{*}Y \\
\end{bmatrix}
\begin{bmatrix}
\hat{B} \\
-N E^{-1/2} \\
-\hat{D}^{*}E^{-1/2} + \hat{D} \\
\end{bmatrix}
\begin{bmatrix}
\hat{A} & N \\
\hat{A} + NF & \hat{A} \\
\hat{A} + NF & \hat{A} \\
\end{bmatrix}
\begin{bmatrix}
\hat{B}E^{1/2} \\
-N \\
N \\
\end{bmatrix}
\begin{bmatrix}
\hat{B}E^{1/2} \\
-\hat{B}E^{1/2}F \\
-N \\
\end{bmatrix}
\begin{bmatrix}
\hat{B}E^{1/2} \\
-N \\
N \\
\end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\hat{B}E^{1/2} \\
-N \\
N \\
\end{bmatrix}
\]
which after the change of basis
\[
T = \begin{bmatrix} I & 0 & 0 \\ 0 & I & I \\ 0 & 0 & I \end{bmatrix}
\]
becomes
\[
(4.58) \quad X(s) = \begin{bmatrix} \hat{A} & -\hat{B}E^{1/2}F & \hat{B}E^{1/2} \\ 0 & \hat{A} & N \\ \hat{C} & -\hat{B}E^{1/2}F - \hat{B}^{*}Y & \hat{D}E^{1/2} - \hat{D}^{*} \end{bmatrix}.
\]
The change of basis
\[
T = \begin{bmatrix} I & P^{*}_{13} & P^{*}_{23} \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix},
\]
together with the (3, 1) and (3, 2) blocks of (4.51), and the (1, 1) and (1, 2) blocks of (4.52), gives
\[
(4.59) \quad X(s) = \begin{bmatrix} \hat{A} & \hat{B}E^{1/2} + P^{*}_{13}N_{1} + P^{*}_{23}N_{2} \\ \hat{C} & \hat{D}E^{1/2} - \hat{D}^{*} \end{bmatrix}
\]
thereby proving (4.22).

The last part of the proof is concerned with showing that (4.23) is true. We will establish this by a chain of intricate manipulations of the state-space realization of \(T_{12}XY_{21}(s)\). From (4.12) and (4.59) we have
\[
T_{12}XT_{21}(s) = \begin{bmatrix} A_{11} & B_{12} \\ C_{11} & D_{12} \end{bmatrix} \begin{bmatrix} \hat{A} & \hat{B}E^{1/2} + P^{*}_{13}N_{1} + P^{*}_{23}N_{2} \\ \hat{C} & \hat{D}E^{1/2} - \hat{D}^{*} \end{bmatrix} \begin{bmatrix} A_{22} & B_{21} \\ C_{22} & I \end{bmatrix}
\]
\[
(4.60) \quad \begin{bmatrix} A_{11} & B_{12}(\hat{D}E^{1/2} - \hat{D}^{*})C_{22} & B_{12}\hat{C} & B_{12}(\hat{D}E^{1/2} - \hat{D}^{*}) \\ 0 & A_{22} & 0 & B_{21} \\ 0 & (\hat{B}E^{1/2} + P^{*}_{13}N_{1} + P^{*}_{23}N_{2})C_{22} & \hat{A} & \hat{B}E^{1/2} + P^{*}_{13}N_{1} + P^{*}_{23}N_{2} \\ C_{11} & D_{12}(\hat{D}E^{1/2} - \hat{D}^{*})C_{22} & D_{12}\hat{C} & D_{12}(\hat{D}E^{1/2} - \hat{D}^{*}) \end{bmatrix}.
\]
The first change of basis in the state space of (4.60) is given by
\[
(4.61) \quad T = \begin{bmatrix} I & 0 & -Q_{13} \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}.
\]
This together with the (1, 3) block of (4.53) gives
\[
(4.62) \quad \begin{bmatrix} A_{11} & B_{12}(\hat{D}E^{1/2} - \hat{D}^{*})C_{22} - Q_{13}(\hat{B}E^{1/2} + P^{*}_{13}N_{1} + P^{*}_{23}N_{2})C_{22} & 0 & \Pi \\ 0 & A_{22} & 0 & B_{21} \\ 0 & (\hat{B}E^{1/2} + P^{*}_{13}N_{1} + P^{*}_{23}N_{2})C_{22} & \hat{A} & \hat{B}E^{1/2} + P^{*}_{13}N_{1} + P^{*}_{23}N_{2} \\ C_{11} & D_{12}(\hat{D}E^{1/2} - \hat{D}^{*})C_{22} & D_{12}\hat{C} + C_{11}Q_{13} & D_{12}(\hat{D}E^{1/2} - \hat{D}^{*}) \end{bmatrix}.
\]

Our next lemma gives a simplified expression for \(\Pi\) in (4.62).
**Lemma B.**

\[ \Pi = (\Theta_{12} - Q_{12})B_{21} - B_{11} \]

**Proof.** See Appendix B.

A second change of basis

\[
T = \begin{bmatrix}
I & Q_{12} - \Theta_{12} & 0 \\
0 & I & 0 \\
0 & 0 & I
\end{bmatrix}
\]

(4.63)

gives

\[
T_{12}X_{21}(s) = \begin{bmatrix}
A_{11} & \Sigma & 0 \\
0 & A_{22} & 0 \\
0 & (\hat{B}E^{1/2} + p_{13}^*N_1 + p_{23}^*N_2)C_{22} & \hat{A}
\end{bmatrix}
\begin{bmatrix}
-B_{11} \\
-B_{21}
\end{bmatrix}
\]

**Lemma C.**

\[ \Sigma = -A_{12}. \]

**Proof.** See Appendix C.

Making use of Lemma C yields

(4.64)

Before continuing further, we need to link \( \Theta \) and \( Y \) (the solutions to (4.36) and (4.42), respectively). This connection is provided next.

**Lemma D.** (a) There exists a unique “largest” solution \( Z_0 \) to the Riccati equation

\[
\{A_{11} + B_{13}D_{13}^*D_{11}E^{-1}(D_{11}^*C_{11} + B_{11}^*)\}Z + Z\{A_{11} + B_{13}D_{13}^*D_{11}E^{-1}(D_{11}^*C_{11} + B_{11}^*)\}^* + Z(C_{11}^*D_{11} + B_{11})E^{-1}(D_{11}^*C_{11} + B_{11}^*)Z + B_{13}[I + D_{13}^*D_{11}E^{-1}D_{11}^*D_{13}]B_{13} = 0.
\]

Further,

\[
H_0 = (B_{13}D_{13}^*D_{11} + Z_0(C_{11}^*D_{11} + B_{11}))E^{-1}
\]

is a destabilizing output injection for \( [A_{11}, (D_{11}^*C_{11} + B_{11}^*)] \). In other words, \( \text{Re} [\lambda (A_{11} + H_0(D_{11}^*C_{11} + B_{11}^*))] > 0 \).

(b) The stabilizing solution to (4.42) is related to \( Z_0 \) and \( \Theta \) by

\[
Y^{-1} + \Theta = \begin{bmatrix} Z_0 & 0 \\ 0 & \gamma^2 I \end{bmatrix}
\]

where the partitioning is induced by that in (4.37) and (4.43).

**Proof.** See Appendix D.

For convenience we introduce the notation

(4.68)

\[
Y^{-1} = \begin{bmatrix}
\hat{Y}_{11} & \hat{Y}_{12} \\
\hat{Y}_{12}^* & \hat{Y}_{22}
\end{bmatrix}
\]
and the last transformation we require is

\[ T = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & P_{13}^* \hat{Y}_{12} + P_{23}^* \hat{Y}_{22} & I \end{bmatrix}. \]

Applying this to (4.64) gives

\[ T_{12} X_{T_2} = \begin{bmatrix} A_{11} & A_{12} & 0 & -B_{11} \\ 0 & A_{22} & 0 & -B_{21} \\ 0 & \Phi & \hat{A} & \Delta \\ C_{11} & \Psi & D_{12} \hat{C} + C_{11} Q_{13} & D_{12} (\hat{D}E^{1/2} - \hat{D} \hat{K}) \end{bmatrix}. \]

**Lemma E.**

\[ \Psi = C_{12}, \]
\[ \Phi = 0, \]
\[ \Delta = \hat{B} E^{1/2} + P_{13}^* B_{13} D_{13}^* D_{11} + P_{23}^* C_{23}^* E + (P_{13}^* \Theta_{11} + P_{23}^* \Theta_{12}^*) (C_{11}^* D_{11} + B_{11}). \]

**Proof.** See Appendix E.

Invoking Lemma E gives

\[ (T_{11} + T_{12} X_{T_2})(s) = \begin{bmatrix} \hat{A} & \Delta \\ D_{12} \hat{C} + C_{11} Q_{13} & D_{12} (\hat{D}E^{1/2} - \hat{D} \hat{K}) \end{bmatrix} \]

which verifies (4.23) and concludes our proof. \( \Box \)

**4.2. The bound \( c_b. \)** We will establish \( c_b \) by counting those points (including multiplicities) at which cancellations may occur in \( F_l(P(s), -K(s)) \) in the case that \( K(s) \) is stabilizing.

Theorem 4.2 below shows that every uncontrollable mode in \( F_l(P(s), -K(s)) \) is due to a cancellation at a zero of \( P_{21}(s) \). In the case that \( K(s) \) is (internally) stabilizing, the number of uncontrollable modes is bounded above by the number of zeros of \( P_{21}(s) \) in \( C_- \) (counting multiplicities). In the same way, the number of unobservable modes is bounded above by the number of zeros of \( P_{12}(s) \) in \( C_- \) (counting multiplicities).

**Theorem 4.2** [2], [19], [20]. Let

\[ \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} A & B_1 \\ C_1 & D_{11} \\ C_2 & D_{21} \end{bmatrix} \]

in which \( P_{12}(s) \in \mathbb{R}^{p_1 \times m_2}(s) \) with \( p_1 \geq m_2 \) and \( P_{21}(s) \in \mathbb{R}^{p_2 \times m_1}(s) \) with \( m_1 \geq p_2 \). Suppose also that

\[ K(s) = \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} \]

is a minimal realization and that the well-posedness condition \( \det(I - D_{22} \hat{D}) \neq 0 \) is satisfied. Then, in the closed loop of Fig. 1,

(a) Every unobservable mode (from \( y_1 \)) is a Smith zero of

\[ sI - A \begin{bmatrix} B_2 \\ C_1 \end{bmatrix}. \]
(b) Every uncontrollable mode (from \( u_i \)) is a Smith zero of

\[
\begin{bmatrix}
  sI - A & B_1 \\
  C_2 & D_{21}
\end{bmatrix}.
\]

**Proof.** See [19], [20].

Theorem 4.2 thus allows us to write

\[
c_b = \{\text{number of zeros of } P_{12} \text{ in } \mathbb{C}_-\} + \{\text{number of zeros of zeros of } P_{21} \text{ in } \mathbb{C}_-\}
\]

and, as a result of Lemma 3.1(i) and Lemma 4.3 below, this becomes

\[
c_b = \{n - \text{rank } (\mathcal{F})\} + \{n - \text{rank } (\mathcal{Y})\}
\]

(4.76)

\[= 2n - \text{rank } (\mathcal{F}) - \text{rank } (\mathcal{Y}).\]

Lemma 4.3 provides a link between the Smith zeros of \( P_{12} = (A, B_2, C_1, D_{12}) \) and the modes of \((A - B_2 D_{12}^\perp C_1)\) which are undetectable through \( D_{12}^\perp C_1 \).

**Lemma 4.3.** Suppose that

\[
P(s) = \begin{bmatrix}
  sI - A & -B_2 \\
  C_1 & D_{12}
\end{bmatrix}
\]

is a polynomial matrix of dimension \((n + p) \times (n + m)\) with \(p > m\). If \(D_{12}\) is part of an orthogonal matrix and \(D_\perp\) is its orthogonal completion, then every Smith zero of \( P(s) \) is an unobservable mode of \([A - B_2 D_{12}^\perp C_1, D_{12}^\perp C_1]\) and vice versa.

**Proof.** If \(s_0\) is a Smith zero of \( P(s) \), there exists a vector \([w^* | v^*] \neq 0\) such that

\[
\begin{bmatrix}
  s_0 I - A & -B_2 \\
  C_1 & D_{12}
\end{bmatrix} \begin{bmatrix}
  w \\
  v
\end{bmatrix} = 0.
\]

(4.78)

We note also that \(w \neq 0\) since if this were not the case, we would have \(D_{12} v = 0\) which is impossible. From (4.78) we get

\[
(s_0 I - A) w - B_2 v = 0
\]

(4.79)

and

\[
C_1 w + [D_{12} | D_\perp] \begin{bmatrix}
  v \\
  0
\end{bmatrix} = 0.
\]

(4.80)

Multiplying (4.80) on the left by \([D_{12} | D_\perp]^*\) gives

\[
D_{12}^\perp C_1 w + v = 0,
\]

(4.81a)

\[
D_{12}^\perp C_1 w = 0.
\]

(4.81b)

Substituting (4.81a) into (4.79) and combining the result with (4.81b) yields

\[
\begin{bmatrix}
  s_0 I - A + B_2 D_{12}^\perp C_1 \\
  D_\perp C_1
\end{bmatrix} w = 0
\]

(4.82)

which completes the proof in one direction.

If (4.82) is satisfied, we may write

\[
(s_0 I - A) w - B_2 v = 0
\]

(4.83)

where

\[
v := -D_{12}^\perp C_1 w.
\]

(4.84)
Combining (4.84) with the (2, 1) block of (4.82) gives
\[
\begin{bmatrix}
D_1^+ \\
D_1^*
\end{bmatrix} C_1 w + \begin{bmatrix} v \\ 0 \end{bmatrix} = 0
\]
and hence also
\[(4.85) \quad C_1 w + D_{12} v = 0.\]
Finally, we note that (4.85) combined with (4.83) gives (4.78) thereby establishing the result. \(\square\)

4.3. The controller degree bound. The main theorem is proved by substituting (4.11) and (4.76) into (4.2).

**Theorem 4.4.** For any \(\mathcal{H}^\infty\)-optimal control problem of the second kind, every \(\mathcal{H}^\infty\)-optimal controller satisfies
\[(4.86) \quad (1) \ deg(K) \leq n + \deg(U) - 1\]
and every suboptimal controller \((\| R(s) \|_\infty > \gamma_{\text{opt}})\) satisfies
\[(4.87) \quad (2) \ deg(K) \leq n + \deg(U).\]
In (4.86) and (4.87), \(U(s) \in \mathcal{RH}_\infty\) is an arbitrary matrix contraction of specified dimensions, which may be chosen constant (or even zero). \(\square\)

4.4. \(\mathcal{H}^\infty\)-optimal control problems of the third kind. The \(n - 1\) degree bound has now been proved in the case of problems of the first and second kind. With this background it is natural to ask: "Does this bound carry over to problems of the third kind?" Problems of the third kind are characterized by the assumption that \(P_{12}(s)\) in (2.14) has more rows than columns while \(P_{21}(s)\) has more columns than rows. Several small computer examples \((n = 1, 2, 3, 4, 5)\) indicate that the answer to this question is indeed "yes." In the case of larger problems, finite precision effects make cancellation phenomena increasingly difficult to detect. On the basis of these experimental observations, we offer the following conjecture.

**Conjecture.** For any \(\mathcal{H}^\infty\)-optimal control problem of the third kind, every \(\mathcal{H}^\infty\)-optimal controller satisfies
\[
(i) \quad \deg(K) \leq n + \deg(U) - 1
\]
and every suboptimal controller \((\| R(s) \|_\infty > \gamma_{\text{opt}})\) satisfies
\[
(ii) \quad \deg(K) \leq n + \deg(U).
\]
As before, \(U(s) \in \mathcal{RH}_\infty\) is an arbitrary matrix contraction which may be chosen constant.

4.5. Computation time trials. In this subsection we present quantitative data which substantiate our claims regarding the importance of removing cancellation phenomena from \(\mathcal{H}^\infty\) computer software. We performed a number of test computations on both the original software, and an improved program which takes into account many of the cancellation phenomena predicted by the results in this paper. Both programs were run on a VAX 750 computer under UNIX. The timing data was obtained using the UNIX routine dtme. In every case we allowed the \(\gamma\)-iteration to run until the solution was almost optimal: The computation exited from the iterative loop when \(\mu\) (see again (2.29)) was in the interval \(1 - 0.5 \times 10^{-5} \leq \mu \leq 1.\)
Table 4.1

<table>
<thead>
<tr>
<th>Number of states</th>
<th>Modified program</th>
<th>Original program</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.517 s</td>
<td>18.23 s</td>
</tr>
<tr>
<td>2</td>
<td>11.617 s</td>
<td>67.95 s</td>
</tr>
<tr>
<td>3</td>
<td>53.417 s</td>
<td>406.98 s</td>
</tr>
<tr>
<td>4</td>
<td>102.47 s</td>
<td>442.32 s</td>
</tr>
<tr>
<td>6</td>
<td>141.43 s</td>
<td>692.98 s</td>
</tr>
<tr>
<td>8</td>
<td>215.77 s</td>
<td>1611.42 s</td>
</tr>
<tr>
<td>14</td>
<td>1128.48 s</td>
<td>25273.70 s</td>
</tr>
<tr>
<td>17</td>
<td>1703.47 s</td>
<td>terminated after 12 hours of processor time</td>
</tr>
</tbody>
</table>

The results given in Table 4.1 show the state dimension of $P(s)$, the execution time of the original program and the execution time of the improved program (both in seconds). Apart from a marked reduction in computation time, the improved program demonstrated improved robustness properties.

5. Conclusions. The purpose of this paper has been to generalise the analysis in [20] to problems of the second kind. If in Fig. 1 $\deg(P) = n$, we have shown that any $H^\infty$-optimal control problem of the second kind has an associated controller which requires no more than $n - 1$ states. In the case that a suboptimal value of $\gamma (> \gamma_{opt})$ is chosen, there is a continuum of controllers with $\deg(K) \approx n$. These results are stated formally in Theorem 4.4.

Our experience has been that the solution of “large” (big $n$) $H^\infty$ control problems is time consuming, especially when several iterations corresponding to various weight selections are required. Further, long calculations of this type are susceptible to severe numerical difficulties. The work in this paper has shown that there is considerable scope for reducing these problems by using cancellation theory to remove state inflation effects from computer code. Although the level of benefit varies from problem to problem, a cpu calculation time reduction of between five and ten times is easy to achieve. The modified code is also considerably more robust from a numerical point of view.

Appendix A.

Proof of Lemma A. Since $\bar{A}$ is asymptotically stable, $[\bar{A}, N]$ is stabilisable. $\gamma > \| R_2(s) \|_\infty \Rightarrow \gamma^2 I - D_2 \bar{D}_2^* = E$ is positive definite and

$$
\begin{bmatrix}
I & 0 \\
-sI - \bar{A}|N| & -E^{-1/2}
\end{bmatrix}
= \begin{bmatrix}
-sI - \bar{A} - NE^{-1}\bar{C} & |NE^{-1/2}|
\end{bmatrix}
$$

ensures that $[\bar{A} + NE^{-1}\bar{C}, NE^{-1/2}]$ is also stabilisable.

The Hamiltonian matrix associated with (4.42) is

$$
H = \begin{bmatrix}
\bar{A} + NE^{-1}\bar{C} & NE^{-1}N^* \\
-\bar{C}^*E^{-1}\bar{C} & -(\bar{A} + NE^{-1}\bar{C})^*
\end{bmatrix}
$$

and we note that

(A.1) $\{\lambda(H)\} \subseteq \{$Smith–McMillan zeros of $(\gamma^2 I - R^*_2 R_2)(s)$\} \cup \{\lambda_{(\bar{A})}\} \cup \{\lambda_{(-\bar{A})}\}$. Since $\gamma > \| R_2(s) \|_\infty$ no Smith–McMillan zero of $(\gamma^2 I - R^*_2 R_2)(s)$ lies on the imaginary axis. This together with the asymptotic stability of $\bar{A}$ ensures that $H$ is free of imaginary axis eigenvalues. Finally, the stabilisability of $[\bar{A} + NE^{-1}\bar{C}, NE^{-1/2}]$ and Lemma 1(i)
ensure the existence of a unique stabilising solution $Y = Y^* \equiv 0$ to (4.42). Part (a) in the theorem statement ensures that $[\bar{A}, E^{-1/2} \bar{C}]$ is observable and Lemma 1(ii) $\Rightarrow Y$ is nonsingular or else that the unique stabilising solution satisfies $Y = Y^* > 0$.

**Appendix B.**

*Proof of Lemma B.* The change of basis (4.61) in the state space of (4.60) gives

$$\Pi = B_{12}(\hat{D}E^{1/2} - \tilde{D}^*) - Q_{13}(\hat{B}E^{1/2} + P_{13}^* N_1 + P_{23}^* N_2).$$

Making use of the (1, 1) block of (4.54) yields

$$\Pi = Q_{11}' F_1^* E + Q_{12}' F_2^* E - B_{13} D_{12}^* D_{11} - Q_{13} P_{13}^* N_1 - Q_{13} P_{23}^* N_2.$$

From the (1, 1) and (1, 2) blocks of (4.20), and from (4.56), we get

$$\Pi = -N_1 + Q_{11}(C_{11}^* D_{11} + B_{11}) + Q_{12} C_{22}^* - B_{12} D_{12}^* D_{11}.$$

Equation (4.57) gives

$$\Pi = -C_{11}^* D_{11} - B_{11} + (Q_{12} - \Theta_{12}) C_{22}^* - B_{13} D_{13}^* D_{11} - B_{12} D_{12}^* D_{11}.$$

Finally, from (4.30) we get

$$\Pi = -B_{11} + (Q_{12} - \Theta_{12}) C_{22}^*.$$

**Appendix C.**

*Proof of Lemma C.* After the coordinate change (4.61) in (4.60) we get

$$\Sigma = (Q_{12} - \Theta_{12}) A_{22} - A_{11}(Q_{12} - \Theta_{12}) + B_{12}(\hat{D}E^{1/2} - \tilde{D}^*) C_{22}$$

$$\quad - Q_{13}(\hat{B}E^{1/2} + P_{13}^* N_1 + P_{23}^* N_2) C_{22}.$$

Making use of (4.15) and the (1, 1) block of (4.54) gives

$$\Sigma = (Q_{12} - \Theta_{12}) A_{22} - A_{11}(Q_{12} - \Theta_{12}) - B_{13} D_{12}^* D_{11} C_{22}$$

$$\quad - Q_{13}(P_{13}^* N_1 + P_{23}^* N_2) C_{22} + (Q_{11}' F_1^* E + Q_{12}' F_2^* E) C_{22}.$$

Substituting from the (1, 1) and (1, 2) blocks of (4.20), (4.40), (4.45), (4.24), (4.25) and (4.15) yields

$$\Sigma = (Q_{12} - \Theta_{12}) A_{22} - A_{11}(Q_{12} - \Theta_{12}) - (B_{13} D_{13}^* D_{11} + \Theta_{11}(C_{11}^* D_{11} + B_{11}) + \Theta_{12} C_{22}^*) C_{22}$$

$$\quad + Q_{13} (B_{11} + C_{11}^* D_{11}) C_{22} - B_{12} D_{12}^* D_{11} C_{22} + Q_{12}(-A_{22} - A_{22}^*).$$

Invoking (4.24), (4.25), (4.30) and (4.57) gives

$$\Sigma = A_{11}(\Theta_{12} - Q_{12}) - (Q_{12} - \Theta_{12}) A_{22}^* - B_{11} C_{22}.$$

Finally, the (1, 2) block of (4.36), the (1, 2) block of (4.53) and (4.25) lead to the required result

$$\Sigma = -A_{12}.$$

**Appendix D.**

*Proof of Lemma D.* A minor variant of Lemma 2.1(i) shows that the existence of the destabilising solution $Z_0$ is established by proving

(i) $[A_{11}^*, C_{11}^* D_{11} + B_{11}]$ is controllable;

(ii) The Hamiltonian matrix corresponding to (4.65) is free of imaginary axis eigenvalues.

The fact that $Z_0$ is the largest solution (i.e., $Z_0 - Z \equiv 0$ for all other solutions) is only of peripheral interest and consequently will not be proved here. In fact, all that is needed is a minor modification of an argument in [28, Lemma 3].
We begin by showing that (i) is true. We know from part (a) of Theorem 4.1 and (4.25) that

\[
\begin{bmatrix}
-A_{11}^* & (C_{11}^* D_{11} + B_{11}) C_{22} \\
0 & -A_{22}^*
\end{bmatrix}
\begin{bmatrix}
C_{11}^* D_{11} + B_{11} \\
C_{22}
\end{bmatrix}
\]

is controllable. Next, we suppose for contradiction that (i) is not satisfied. That is, there exists a vector \( w \neq 0 \) such that

\[
w^* A_{11}^* = \lambda w^*,
\]

\[
w^* (C_{11}^* D_{11} + B_{11}) = 0.
\]

From this we have

\[
(D.1) \begin{bmatrix} w^* | 0 \end{bmatrix} \begin{bmatrix}
-A_{11}^* & (C_{11}^* D_{11} + B_{11}) C_{22} \\
0 & -A_{22}^*
\end{bmatrix} \begin{bmatrix} -\lambda w^* | 0 \end{bmatrix}
\]

and

\[
(D.2) \begin{bmatrix} w^* | 0 \end{bmatrix} \begin{bmatrix}
C_{11}^* D_{11} + B_{11} \\
C_{22}^*
\end{bmatrix} = [0 | 0]
\]

which contradicts part (a) of Theorem 4.1 (which is already proved). This contradiction establishes (i).

As we will now show, (ii) is in fact a consequence of Lemma A. Direct substitution into the Hamiltonian in Lemma A gives

\[
H = (C_{11}^* D_{11} + B_{11}) E^{-1} (D_{11}^* C_{11} + B_{11}^*) - A_{11}^* E^{-1} C_{22} - N_1 E^{-1} N_1^* - N_2 E^{-1} N_2^* - A_{11}^* - (C_{11}^* D_{11} + B_{11}) E^{-1} N_1^* - C_{22} E^{-1} N_1^* - (B_{11} + C_{11}^* D_{11}) (E^{-1} N_2^* - C_{22})
\]

A tortuous but routine computation based on (4.13), (4.24), (2.25), (4.33), (4.40) and (4.41) shows that

\[
(D.4) \begin{bmatrix} A_{11}^* + (C_{11}^* D_{11} + B_{11}) E^{-1} D_{11}^* D_{13} B_{13}^* \\
-B_{13} D_{13}^*[I + D_{11} E^{-1} D_{11}^*] D_{13} B_{13}^* \\
0 \\
C_{22}^* E^{-1} D_{11}^* D_{13} B_{13}^*
\end{bmatrix}
\]

\[
-(C_{11}^* D_{11} + B_{11}) E^{-1} (D_{11}^* C_{11} + B_{11}^*) - A_{11}^* - B_{13} D_{13}^* D_{11} E^{-1} (D_{11}^* C_{11} + B_{11}^*) - B_{13} D_{13}^* D_{11} E^{-1} C_{22} - 0
\]

\[
0 \begin{bmatrix} A_{22}^* \\
C_{22}^* E^{-1} D_{11}^* C_{11} + B_{11}^* \\
C_{22}^* E^{-1} C_{22} \end{bmatrix} - A_{22}
\]
where

\[
T = \begin{bmatrix}
0 & 0 & I & 0 \\
I & 0 & \Theta_{11} & \Theta_{12} \\
0 & I & \Theta_{12}^* & \Theta_{22} - \gamma^2 I \\
0 & 0 & 0 & I
\end{bmatrix}.
\] 

Since \( H \) in (D.3) is free of imaginary axis eigenvalues (Lemma A), so too is the matrix in (D.4). Suppose

\[
S = \begin{bmatrix}
(1, 1) & (1, 2) \\
(2, 1) & (2, 2)
\end{bmatrix}
\]

where \((i, j)\) represents the \((i, j)\)th block of (D.4). Then it is easy to see that

(a) \( S \) is the Hamiltonian associated with (4.65);

(b) \( \lambda(S) \subset \lambda(H) \Rightarrow (ii) \).

This concludes the proof of the existence part of (a).

Let us suppose that the orthogonal matrix \( W \) transforms \( S \) into an ordered upper Schur form [18], specifically

\[
S = \begin{bmatrix}
W_{11} & W_{12} \\
W_{21} & W_{22}
\end{bmatrix} = \begin{bmatrix}
T_{11} & T_{12} \\
T_{21} & T_{22}
\end{bmatrix}
\]

in which \( T_{11} \) is completely unstable and \( T_{22} \) is asymptotically stable. Substituting from (D.4) into (D.6) gives

\[
\begin{bmatrix}
S_{11} & S_{12} \\
S_{21} & S_{22}
\end{bmatrix} \begin{bmatrix}
W_{11} \\
W_{21}
\end{bmatrix} = \begin{bmatrix}
W_{11} \\
W_{21}
\end{bmatrix} \begin{bmatrix}
T_{11} \\
0
\end{bmatrix}
\]

where

\[
S_{11} = (A_{11} + B_{13} D_{13}^* D_{11} E^{-1}(D_{11}^* C_{11} + B_{11})),
\]

\[
S_{12} = (C_{11}^* D_{11} + B_{11}) E^{-1}(D_{11}^* C_{11} + B_{11}^*),
\]

\[
S_{21} = -B_{13} (I + D_{13}^* D_{11} E^{-1} D_{11}^* D_{13}) B_{13}^*,
\]

\[
S_{22} = -S_{11}^*.
\]

and the Riccati equation solution is [18]

\[
Z_0 = W_{21} W_{11}^{-1}.
\]

Conjugating the \((1, 1)\) block of (D.8) and multiplying on the left by \( W_{11}^* \) gives

\[
A_{11} + (B_{13} D_{13}^* D_{11} + Z_0 (C_{11}^* D_{11} + B_{11})) E^{-1}(D_{11}^* C_{11} + B_{11}) = W_{11}^* T_{11}^* W_{11}^*
\]

which proves that \( H_0 \) in (4.66) is a destabilising output injection. Thus (a) is proved.

The (b) part of the lemma will be established in two steps. First, we will prove that any solution to (4.65) will generate a solution to (4.42) via (4.67). Following that, we show that the largest (destabilising) solution to (4.65) generates the stabilising solution to (4.42).

From (4.55) we get

\[
Y^{-1} \tilde{A}^* + \tilde{A} Y^{-1} + Y^{-1} F^* EFY^{-1} = 0;
\]

adding this to (4.36) gives

\[
(Y^{-1} + \Theta) \tilde{A}^* + \tilde{A} (Y^{-1} + \Theta) + Y^{-1} F^* EFY^{-1} + \tilde{B}_2 \tilde{B}_2^* = 0.
\]
Substituting from (4.67) and writing (D.12) out in full gives

\[
\begin{bmatrix}
Z \\
0
\end{bmatrix}
\begin{bmatrix}
A_{11}^* & A_{12}^* \\
0 & A_{22}^*
\end{bmatrix}
+ \begin{bmatrix}
A_{11} & 0 \\
A_{12} & A_{22}
\end{bmatrix}
\begin{bmatrix}
Z \\
0
\end{bmatrix}
\gamma^2 I
+ \begin{bmatrix}
Z - \Theta_{11} & -\Theta_{12} \\
-\Theta_{12}^* & \gamma^2 I - \Theta_{22}
\end{bmatrix}
\begin{bmatrix}
C_{11}^* D_{11} + B_{11} \\
C_{22}^*
\end{bmatrix}
E^{-1}[D_{11}^* C_{11} + B_{11}^* | C_{22}]
+ \begin{bmatrix}
Z - \Theta_{11} & -\Theta_{12} \\
-\Theta_{12}^* & \gamma^2 I - \Theta_{22}
\end{bmatrix}
\begin{bmatrix}
N_1 \\
N_2
\end{bmatrix}
E^{-1}[D_{11}^* C_{11} | C_{22}]
+ \begin{bmatrix}
Z - \Theta_{11} & -\Theta_{12} \\
-\Theta_{12}^* & \gamma^2 I - \Theta_{22}
\end{bmatrix}
\begin{bmatrix}
B_{13} \\
B_{23} D_{11} D_{13}
\end{bmatrix}
[B_{13}^* | D_{11}^* D_{11} B_{23}^{**}]
+ \begin{bmatrix}
Z - \Theta_{11} & -\Theta_{12} \\
-\Theta_{12}^* & \gamma^2 I - \Theta_{22}
\end{bmatrix}
\begin{bmatrix}
C_{11}^* D_{11} + B_{11} \\
C_{22}^*
\end{bmatrix}
E^{-1}[N_1 | N_2]
+ \begin{bmatrix}
N_1 \\
N_2
\end{bmatrix}
E^{-1}[N_1^* | N_2^*] = 0
\]

in which

(D.14) \[\bar{A}_{12} = B_{21}(B_{11}^* + D_{11}^* C_{11}).\]

Equations (4.40) and (4.25) allow the (2, 2) block (denoted (2, 2)) of (D.13) to be written out in full as

\[
(2, 2) = \gamma^2 A_{22} + \gamma^2 A_{22}^* + [B_{21} E + N_2] E^{-1}[E B_{21}^* + N_2^*] - N_2 [B_{21}^* + E^{-1} N_2^*]
+ N_2 E^{-1} N_2^* - [B_{21} + N_2 E^{-1}] N_2^* + B_{21}(\gamma^2 I - E) B_{21}^*.
\]

After we cancel terms this becomes

(D.15) \[\gamma^2 (A_{22} + A_{22}^* + B_{21} B_{21}^*) = 0 \text{ by (4.24)}.\]

The definitions of N and E in (4.40) and (4.41) allow the (1, 2) block of (D.13) (denoted (1, 2)) to be written out in full as

\[
(1, 2) = Z \bar{A}_{12}^* + [Z (C_{11}^* D_{11} + B_{11}) + B_{13} D_{11}^* D_{11} - N_1] E^{-1}[-E B_{21}^* - N_2^*]
+ N_1 [B_{21}^* + E^{-1} N_2^*] + [Z (C_{11}^* D_{11} + B_{11}) + B_{13} D_{11}^* D_{11} - N_1] E^{-1} N_2^*
+ N_1 E^{-1} N_2^* + B_{13} D_{11}^* D_{11} B_{21}^*.
\]

After we cancel terms this becomes

(D.16) \[Z (\bar{A}_{12}^* - (C_{11}^* D_{11} + B_{11}) B_{21}^*) = 0 \text{ by (D.14)}.\]

Clearly, (2, 1) = 0 follows by symmetry.

As with (D.15) and (D.16), it is easy to show that the (1, 1) block of (D.13) is zero provided Z is a solution of (4.65). This verification only requires the definitions of N₁ and N₂ in (4.40).
We now begin a sequence of arguments which prove that the largest solution \( Z_0 \) to (4.65) generates the stabilizing solution to (4.42). Since \( H_0 \) in (4.66) is a destabilizing output injection, and since \( A_{22} \) is stable, the matrix

\[
\begin{bmatrix}
A^*_1 + (C^*_1 D_{11} + B_{11}) H^*_0 & 0 \\
C^*_2 E^{-}[D^*_1 D_{13} B_{13}^* + (D^*_1 C_{11} + B_{11}^*) Z_0] & -A_{22}
\end{bmatrix}
\]

is completely unstable,

\[
A^*_1 + (C^*_1 D_{11} + B_{11}) H^*_0 \\ C^*_2 E^{-}[D^*_1 D_{13} B_{13}^* + (D^*_1 C_{11} + B_{11}^*) Z_0] & A^*_22 + C^*_22 C^*_22
\]

is completely unstable by (D.14), (4.24) and (4.25), and

\[
\begin{bmatrix}
A^*_1 & \tilde{A}^*_2 \\
0 & A^*_22
\end{bmatrix} + \begin{bmatrix}
C^*_1 D_{11} + B_{11} \\
C^*_22
\end{bmatrix} E^{-}[D^*_1 D_{13} B_{13}^* + (D^*_1 C_{11} + B_{11}^*) Z_0] E_{22}
\end{bmatrix}
\]

is completely unstable by (4.66). Using the notation

\[
Y^{-1} = \begin{bmatrix}
\hat{Y}_{11} & \hat{Y}_{12} \\
\hat{Y}_{12} & \hat{Y}_{22}
\end{bmatrix}
\]

together with (4.67) and (4.40) establishes that

\[
\begin{bmatrix}
A^*_1 & \tilde{A}^*_2 \\
0 & A^*_22
\end{bmatrix} + \begin{bmatrix}
C^*_1 D_{11} + B_{11} \\
C^*_22
\end{bmatrix} E^{-}[A | B]
\]

where

\[
\begin{align*}
\mathbb{A} &= N_1^* + (D^*_1 C_{11} + B^*_1) \hat{Y}_{11} + C_{22} \hat{Y}_{12}, \\
\mathbb{B} &= N_2^* + (D^*_1 C_{11} + B^*_1) \hat{Y}_{12} + C_{22} \hat{Y}_{22}
\end{align*}
\]

is completely unstable. Substituting from

\[
(4.15) \Rightarrow \tilde{A}^* + \tilde{C}^* E^{-1}(N^* + \tilde{C}^* \hat{Y})
\]

is completely unstable. From (4.42) we get

\[
\tilde{A}^* + \tilde{C}^* E^{-1}(N^* + \tilde{C}^* \hat{Y}) = -Y \{ \tilde{A} + NE^{-1} \tilde{C} + N^* \hat{Y} \} \hat{Y},
\]

whence by (4.45)

\[
(4.22) \Rightarrow \tilde{A} + NE^{-1}(\tilde{C} + N^* \hat{Y}) = \tilde{A} + NF
\]

is completely stable. This thus proves that \( Z_0 \) generates the required solution to (4.42). □

**Appendix E.**

*Proof of Lemma E.* We begin by proving that

\[
\Psi = C_{11}(Q_{12} - \Theta_{12}) - D_{12}(\hat{D} E^{1/2} - D_{12}^* D_{11}) C_{22}
\]

\[
- (D_{12} \hat{C} + C_{11} Q_{13})(P_{13} \hat{Y}_{12} + P_{23} \hat{Y}_{22}) = C_{12}.
\]

From the (1, 1) and (1, 2) blocks of (4.52), and from (4.56) and (4.15), we get

\[
\begin{align*}
-\hat{C} P_{13} &= \hat{D} E^{1/2} F_{11} + B_{12} Y_{11} + D_{12}^* (D_{12} B_{21} + C_{12}) Y_{12}, \\
-\hat{C} P_{23} &= \hat{D} E^{1/2} F_{12} + B_{12} Y_{12} + D_{12}^* (D_{12} B_{21} + C_{12}) Y_{22},
\end{align*}
\]
We will also use the three equations from (see (4.68))

\[
\begin{bmatrix}
  Y_{11} & Y_{12} \\
  Y_{12}^* & Y_{22}
\end{bmatrix}
\begin{bmatrix}
  \hat{Y}_{11} \\
  \hat{Y}_{12}
\end{bmatrix}
= \begin{bmatrix}
  I & 0 \\
  0 & I
\end{bmatrix}.
\]

Substituting (E.2), (E.3) and (E.4) into (E.1) gives

\[
\Psi = C_{11}(Q_{12} - \Theta_{12}) - D_{12}\hat{D}E^{1/2}C_{22} + D_{12}\hat{D}E^{1/2}F_1\hat{Y}_{12} + D_{12}D_{12}^*C_{12} \\
- C_{11}Q_{13}P_{13}^*\hat{Y}_{12} + D_{12}\hat{D}E^{1/2}F_2\hat{Y}_{22} - C_{11}Q_{13}P_{23}^*\hat{Y}_{22}.
\]

Using (4.13), (4.31), (4.45), (4.25) and (E.4) gives

\[
\Psi = C_{11}(Q_{12} - \Theta_{12} - Q_{13}P_{13}^*\hat{Y}_{12} - Q_{13}P_{23}^*\hat{Y}_{22}) + C_{12} \\
+ D_{12}\hat{D}E^{-1/2}[(D_{11}^*C_{11} + B_{11}^*)\hat{Y}_{12} + C_{22}\hat{Y}_{22} + N_2^* + EC_{22}].
\]

Using (4.13), (4.25), (4.67) and (4.41) we get from (E.6) that

\[
\Psi = C_{11}(Q_{12} - \Theta_{12} - Q_{13}P_{13}^*\hat{Y}_{12} - Q_{13}P_{23}^*\hat{Y}_{22}) + C_{12}.
\]

By (4.20) and

\[
\hat{Y}_{22} = (Y_{22} - Y_{12}^*Y_{12}^{-1}Y_{12})^{-1} \quad (Y > 0 \Rightarrow Y_{11} > 0 \Rightarrow Y_{11}^{-1} \text{ exists}),
\]

\[
\hat{Y}_{12} = Y_{11}^{-1}Y_{12}(Y_{22} - Y_{12}^*Y_{12}^{-1}Y_{12})^{-1}
\]

gives (E.1) as required.

We now establish that

\[
\Phi = (P_{13}^*\hat{Y}_{12} + P_{23}^*\hat{Y}_{22})A_{22} - \hat{\Theta}(P_{13}^*\hat{Y}_{12} + P_{23}^*\hat{Y}_{22}) \\
- (\hat{D}E^{1/2} + P_{13}^*N_1 + P_{23}^*N_2)C_{22} = 0.
\]

Substituting the (3, 1) and (3, 2) blocks of (4.51), (4.45) and (E.4) gives

\[
\Phi = \hat{D}E^{-1/2}\{D_{11}^*C_{11}\hat{Y}_{12} + B_{11}^*\hat{Y}_{12} + N_2^* + C_{22}\hat{Y}_{22} - EC_{22}\} \\
- P_{23}^*\{A_{12}^*\hat{Y}_{12} + A_{22}\hat{Y}_{22} + N_2C_{22} - \hat{Y}_{22}A_{22}\} \\
- P_{13}^*\{A_{11}\hat{Y}_{12} + N_1C_{22} - \hat{Y}_{12}A_{22}\}.
\]

By (4.40), (4.41) and (4.67) we get

\[
\Phi = -P_{23}^*\{A_{12}^*\hat{Y}_{12} + A_{22}\hat{Y}_{22} + N_2C_{22} - \hat{Y}_{22}A_{22}\} - P_{13}^*\{A_{11}\hat{Y}_{12} + N_1C_{22} - \hat{Y}_{12}A_{22}\}.
\]

Making use of (D.14), (4.13), (4.24), (4.40), (4.67), and the (2, 2) block of (4.36) gives

\[
\Phi = -P_{13}^*\{A_{11}\hat{Y}_{12} + B_{13}\hat{D}_{13}^*D_{13}C_{22} + \Theta_{11}(C_{11}^*D_{11} + B_{11})C_{22} - \Theta_{12}A_{22}\^2\}.
\]

From the (1, 2) block of (4.36) we obtain

\[
A_{11}\Theta_{12} - \Theta_{11}(C_{11}^*D_{11} + B_{11})C_{22} + \Theta_{12}A_{22}^2 = B_{13}\hat{D}_{13}^*D_{11}C_{22}
\]

and this together with the (1, 2) block of (4.67) gives

\[
\Phi = 0
\]
as required.

Finally, we have from (4.64), (4.69) and (4.70) that

\[
\Delta = \hat{D}E^{1/2} + P_{13}^*N_1 + P_{23}^*N_2 - (P_{13}^*\hat{Y}_{12} + P_{23}^*\hat{Y}_{22})B_{21}.
\]

From (4.40) this becomes

\[
\Delta = \hat{D}E^{1/2} + P_{13}^*B_{13}\hat{D}_{13}^*D_{11} + \Theta_{11}(C_{11}^*D_{11} + B_{11}) + \Theta_{12}C_{22}^* - \hat{Y}_{12}B_{21} \\
+ P_{23}^*B_{21}\hat{D}_{11}^*D_{11}\hat{D}_{13}^*D_{11} + \Theta_{12}(C_{11}^*D_{11}B_{11}) + \Theta_{22}C_{22}^* - \hat{Y}_{22}B_{21}.
\]
Using (4.25), (4.41) and (4.67) to cancel terms gives
\[
\Delta = \hat{\beta} E^{1/2} + P_{13}^*(B_{13} D_{13} D_{11} + \Theta_{11}(C_{11}^* D_{11} + B_{11})) + P_{23}^*(C_{22}^* E + \Theta_{12}^*(C_{11}^* D_{11} + B_{11}))
\]
\[
= \hat{\beta} E^{1/2} + P_{13}^* B_{13} D_{13} D_{11} + P_{23}^* C_{22}^* E + (P_{11}^* \Theta_{11} + P_{23}^* \Theta_{12}^*)(C_{11}^* D_{11} + B_{11})
\]
as required.

6. Acknowledgments. It is our pleasure to thank E. Kasenally and H. Wang for their assistance with the computer programming and for running numerous examples.

REFERENCES


