

State-space formulas for the computation of the gap

JESSE DE DOES†

We consider the computation of the gap between the graphs of rational transfer functions. The gap is related to the norm of a parallel projection, which can be computed as a regular indefinite linear quadratic optimal control problem with partially specified initial conditions. The use of coprime factorizations is avoided in this way, and Riccati equations are derived directly in terms of the state-space realizations. The computation needed to check whether the gap is smaller than some fixed number γ is shown to be an eigenvector/eigenvalue problem for a hamiltonian matrix in which the state-space parameters of minimal realizations of the transfer functions are stacked.

1. Introduction

Our computations are based on the geometrical relations between several subspaces of Hilbert spaces that one can associate with a linear system. Therefore, a few words about these different spaces are in order. For a transfer function G from U to Y , the graph is defined as a subspace of $H_2 := H_2(\mathbb{C}^+, Y \times U)$ by

$$\mathcal{G}(G) = \{(y, u) \in H_2 \mid y(s) = G(s)u(s)\}$$

The external behaviour $\mathcal{B}_W(\Sigma)$ of a linear system in a certain space W (which can be either $L_2(0, \infty)$ or $L_2(-\infty, \infty)$) is the set of its external solution trajectories in that space (cf. Willems 1991). The relation between the graph and the different behaviours can be sketched as follows: choose any state-space realization $\Sigma = \Sigma(A, B, C, D)$ of G , let \mathcal{L} be the Fourier-Laplace transform isomorphism from $L_2(-\infty, \infty)$ to $L_2(i\mathbb{R}) = H_2^- \oplus H_2$, and let $C(\mathbb{R}, X)$ be the space of continuous functions from \mathbb{R} to the state-space X . Then one defines $\mathcal{B}_{L_2(-\infty, \infty)}(G)$ as

$$\mathcal{B}_{L_2(-\infty, \infty)}(\Sigma) := \{(y, u) \in L_2(-\infty, \infty) \mid \exists x \in C(\mathbb{R}, X): (x, y, u) \in \mathbb{R}\}$$

The behaviour on the right half-line is obtained by applying the truncation operator Π_+ to the behaviour on the whole line

$$\mathcal{B}_{L_2^+}(G) := \Pi_+(\mathcal{B}_{L_2(-\infty, \infty)}(G))$$

In the following $\mathcal{B}(G)$ denotes the behaviour in $L_2(-\infty, \infty)$, $\mathcal{B}_+(G)$ is the behaviour in L_2^+ . Let

$$\mathcal{B}^0(\Sigma) := \{(y, u) \in L_2 \mid \exists x \in C(\mathbb{R}, X): (x, y, u) \in \Sigma \wedge x(0) = 0\}$$

and $\mathcal{B}_+^0(\Sigma) := \Pi_+(\mathcal{B}^0(\Sigma))$. Then the graph corresponds in the time domain to the solutions with initial state zero

$$\mathcal{G}(G) = \mathcal{L}(\mathcal{B}_+^0(G))$$

Received 6 June 1993. Revised 26 April 1994.

† CWI, PO Box 4079, 1009 AB Amsterdam, The Netherlands.

and an alternative useful description is

$$\mathcal{G}(G) = \mathcal{L}(\mathcal{B}(G) \cap L_2^+)$$

Here, L_2^+ is considered as embedded in the standard way in L_2 , $f \mapsto f'$ with $f'(t) = 0$ for $t < 0$.

The gap between two subspaces V_1 and V_2 of a Hilbert space is, by definition, $\delta(V_1, V_2) := \|\Pi_{V_1} - \Pi_{V_2}\|$, where Π_V is the orthogonal projection on V . Alternatively

$$\delta(V_1, V_2) = \max(\vec{\delta}(V_1, V_2), \vec{\delta}(V_2, V_1))$$

where $\vec{\delta}$ is the directed gap, given by

$$\vec{\delta}(V_1, V_2) = \sup_{x \in V_1, \|x\| \leq 1} d(x, V_2)$$

The second definition leads to an expression for the gap as the norm of a projection

$$\vec{\delta}(V_1, V_2) = \|\Pi_{V_2^\perp}|_{V_1}\|$$

In the following we refer to this projection as $\Pi_1 := \|\Pi_{V_2^\perp}\|$. It is on this expression as an operator norm that the computation of the gap in Georgiou (1988) is based: writing shift-invariant V_1 and V_2 as the images of isometric shift-invariant operators Θ_1 and Θ_2 respectively, we get

$$\|\Pi_1\| = \|\Pi_{\text{im}\Theta_2^\perp} \Theta_1\|$$

Now, by the commutant lifting theorem

$$\|\Pi_1\| = \|\Pi_{\text{im}\Theta_2^\perp}|_{\text{im}\Theta_1}\| = \|\Pi_{\text{im}\Theta_2^\perp} \Theta_1\| = \inf_{Q \in H_\infty} \|\Theta_1 + \Theta_2 Q\|_\infty$$

and the rightmost term in this equality defines a model-matching problem that can be solved by standard techniques of H_∞ -optimization. The initial idea behind this note was that the leftmost term 'looks easier'. Indeed, this can be shown directly, without using any H_∞ optimization theory, to define a linear quadratic optimal control problem (cf. Remark (2.5)). However, there are some difficulties in the case of non-strictly proper systems in this approach: the Riccati equations one obtains are of higher order than necessary.

These difficulties can be circumvented by using another way of relating the gap to the norm of a projection. Let $\Pi_{V_2}^{V_1}$ denote the parallel projection of $V_1 + V_2$ along V_1 on V_2 . We use the term parallel projection in a somewhat broader sense than usual, as we also consider it when $V_1 \cap V_2 \neq \{0\}$, in which case it is not a mapping, but it is still defined as the linear relation $\{(x, v_2) \in (V_1 + V_2) \times V_2 \mid \exists v_1 \in V_1: v_1 + v_2 = x\}$. Recall the definition of the minimal angle $\varphi(V_1, V_2)$ between subspaces by (we give the Banach space definition)

$$\sin \varphi(V_1, V_2) = \inf_{x \in V_1, \|x\| \leq 1} d(x, V_2)$$

Good references for the use of these notions in robustness analysis are Foias *et al.* (1991), Ober and Sefton (1991), Schumacher (1992). We now use the following simple geometric fact

$$(1 - \vec{\delta}(V_1, V_2)^2)^{1/2} = \sin \varphi(V_1, V_2^\perp) = \|\Pi_{V_2^\perp}^{V_1}\|^{-1}$$

For $\Pi_{V_1^\perp}$ we use the notation Π_2 . The norm of Π_2 turns out to be more convenient to compute.

In the remainder of this note, let $V = \mathcal{G}(G)$, $V_1 = \mathcal{G}(G_1)$.

2. State-space expression of $\Pi_{V_1^\perp}^V$

The first thing we need is a description of the orthogonal complement of the graph of a transfer function G . In $L_2(-\infty, \infty)$, the orthogonal complements of shift-invariant spaces are shift-invariant themselves, and the adjoint of a multiplication operator is also a multiplication operator, so the orthogonal complement of a behaviour in $L_2(-\infty, \infty)$ is easy to describe.

Lemma 2.1

$$\mathcal{B}(G)^\perp = \mathcal{B}(-\tilde{G})$$

Proof: Write

$$\mathcal{L}(\mathcal{B}(G)) = M \begin{bmatrix} N \\ M \end{bmatrix} L_2(i\mathbb{R})$$

where $G = NM^{-1}$ is a coprime factorization over L_∞ , and M_Θ is the multiplication operator by Θ . The adjoint of the multiplication operator by

$$\begin{bmatrix} N \\ M \end{bmatrix}$$

is of course the operator

$$M \begin{bmatrix} N \\ M \end{bmatrix}^* = M_{[\tilde{N} \quad \tilde{M}]}$$

so $\mathcal{L}(\mathcal{B}(G))^\perp = \ker M_{[\tilde{N} \quad \tilde{M}]}$, and the latter space is $\mathcal{L}(\mathcal{B}(-\tilde{G}))$. □

Lemma 2.2:

$$\mathcal{G}(G)^\perp = \mathcal{L}(\mathcal{B}_+(-\tilde{G}))$$

Proof: It is easy to verify that for V, W closed subspaces of a Hilbert space, one has, in general, $V \ominus (V \cap W) = \Pi_V(W^\perp)$. Apply this to $V = H_2$, $W = \mathcal{L}(\mathcal{B}(G))$, $W^\perp = \mathcal{L}(\mathcal{B}(-\tilde{G}))$, $V \cap W = \mathcal{G}(G)$, $\Pi_V(W^\perp) = \mathcal{L}(\mathcal{B}_+(-\tilde{G}))$. □

As a consequence, it is possible to derive state-space equations for the parallel projection along $\mathcal{G}(G)$ on $\mathcal{G}(G_1)^\perp = \mathcal{L}(\mathcal{B}_+(-\tilde{G}_1))$. Using the fact that the graph of a system corresponds in the time domain to solutions with initial state zero, and $(C(sI - A)^{-1}B + D)^\sim = -B^T(sI - (-A^T))^{-1}C^T + D^T$, we can render the situation

$$\begin{aligned} y_1 &= Gu_1 \\ (y_2, u_2) &\in \mathcal{L}(\mathcal{B}_+(-\tilde{G}_1)) \\ (y_3, u_3) &= (y_1, u_1) + (y_2, u_2) \end{aligned}$$

in the time domain as follows

$$\begin{aligned} \dot{x}_1 &= Ax_1 + Bu_1, \quad x_1(0) = 0 \\ y_1 &= Cx_1 + Du_1 \end{aligned}$$

$$\begin{aligned} \dot{x}_2 &= -A_1^T x_2 + C_1^T y_2 \\ u_2 &= B_1^T x_2 - D_1^T y_2 \\ \begin{bmatrix} y_3 \\ u_3 \end{bmatrix} &= \begin{bmatrix} y_1 \\ u_1 \end{bmatrix} + \begin{bmatrix} y_2 \\ u_2 \end{bmatrix} \end{aligned}$$

We also consider the parallel projection as defined by these equations when it is not a mapping (that is when $\mathcal{G}(G) \cap \mathcal{G}(G_1)^\perp \neq \{0\}$). Perhaps it should be emphasized that the main idea of this section is a slight enhancement of the expressive power of differential equations with constant coefficients, obtained by specifying the initial conditions $x(0)$ partially. One can give state-space descriptions of mappings commuting with the forward shift (transfer functions) from L_2^+ to L_2^+ by fixing initial conditions $x(0) = 0$ for the whole state x . Backwards invariant subspaces of L_2^+ ('behaviours') are described by differential equations without restrictions on the initial state. By imposing partial constraints on the initial state, certain time-variant mappings that play a role in the geometric theory of robust stabilization can be provided with a state-space description (cf. Remark (2.5)).

The equations can be rewritten in state-space form.

Lemma 2.3: *Let V be a shift-invariant subspace in $L_2^+(Y \times U)$, with finite-dimensional state-space representation $V = \mathfrak{B}_+^0(\Sigma(A, B, C, D))$, where (A, B, C, D) is minimal. Let V_1 be analogously represented by $\Sigma_1 := (A_1, B_1, C_1, D_1)$. Assume $I + DD_1^T$ is non-singular. Then the parallel projection of $V + V_1^\perp$ along V on V_1^\perp is given by the system of equations*

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \mathcal{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \mathfrak{B}w, \quad x_1(0) = 0 \\ \Pi_{V_1^\perp}^V(w) &= \mathcal{C} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \mathfrak{D}w \\ w &= \begin{bmatrix} y(\cdot) \\ u(\cdot) \end{bmatrix} \in L_2^+ \\ x(\cdot) &\in C([0, \infty), X) \end{aligned}$$

where

$$\begin{aligned} \mathcal{A} &= \begin{bmatrix} A - BD_1^T(I + DD_1^T)^{-1}C & -BB_1^T + BD_1^T(I + DD_1^T)^{-1}DB_1^T \\ -C_1^T(I + DD_1^T)^{-1}C & -A_1^T + C_1^T(I + DD_1^T)^{-1}DB_1^T \end{bmatrix} \\ \mathfrak{B} &= \begin{bmatrix} BD_1^T(I + DD_1^T)^{-1} & -BD_1^T(I + DD_1^T)^{-1}D + B \\ C_1^T(I + DD_1^T)^{-1} & -C_1^T(I + DD_1^T)^{-1}D \end{bmatrix} \\ \mathcal{C} &= \begin{bmatrix} -(I + DD_1^T)^{-1}C & (I + DD_1^T)^{-1}DB_1^T \\ D_1^T(I + DD_1^T)^{-1}C & B_1^T - D_1^T(I + DD_1^T)^{-1}DB_1^T \end{bmatrix} \\ \mathfrak{D} &= \begin{bmatrix} (I + DD_1^T)^{-1} & -(I + DD_1^T)^{-1}D \\ -D_1^T(I + DD_1^T)^{-1} & D_1^T(I + DD_1^T)^{-1}D \end{bmatrix} \end{aligned}$$

Furthermore, the system $(\mathcal{A}, \mathfrak{B}, \mathcal{C}, \mathfrak{D})$ is minimal.

Proof: The formulas are obtained by straightforward manipulation. What remains to be shown is the minimality of the system $(\mathcal{A}, \mathfrak{B}, \mathcal{C}, \mathfrak{D})$. It is clear

that the order of the system is the sum $n + n_1$ of the McMillan degrees of G and $-\tilde{G}_1$. We use a relation between parallel projections and closed loop behaviours to see that the transfer function $\mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B} + \mathcal{D}$ also has degree $n + n_1$. For a feedback interconnection of a plant G and a controller K , put $P = \ker[I \ -G]$, $C = \ker[-K \ I]$. The closed loop transfer function (cf. Vidyasagar 1985) is defined by

$$H(P, C) = \begin{bmatrix} I & -G \\ -K & I \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} G \\ I \end{bmatrix} (I - KG)^{-1} [K \ I]$$

It is easy to verify that

$$\Pi_P^C = \begin{bmatrix} G \\ I \end{bmatrix} (I - KG)^{-1} [-K \ I], \quad H(P, C) = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + \Pi_P^C \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix}$$

so the McMillan degrees of the matrix expressions for Π_P^C and $H(P, C)$ are the same. But the McMillan degree of $H(P, C)$ is equal to that of its inverse, which equals the degree of

$$\begin{bmatrix} 0 & -G \\ -K & 0 \end{bmatrix}$$

which is obviously equal to the sum of the degrees of G and K . Apply this argument to $K = -\tilde{G}_1$, $\Pi_P^C = \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B} + \mathcal{D}$. \square

It must be noted that this lemma does not give a direct method to actually compute the parallel projection of a given function $w(\cdot) \in L_2$. A two-point boundary value would have to be solved to find an initial condition x_2 ; that is, such that the state-space trajectory $x(\cdot)$ of the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ remains L_2 -stable with initial state $(0, x_2)$ and input $w(\cdot)$.

Remark 2.4: For $V = V_1$, one gets the orthogonal projection on V^\perp . In this case, the equations are of the type that can arise from a variational approach to optimal control, and this is, of course, not accidental. To solve an optimal control problem with cost criterion $\int \|u\|^2 + \|y\|^2$, we can go about it as follows: among the external trajectories corresponding to the initial state x_0 , the one with minimal norm is the orthogonal projection of the origin on the set of trajectories with initial condition x_0 , so we must solve (for a strictly proper system)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A & -BB^T \\ -C^T C & -A^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & B \\ C^T & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad x_1(0) = x_0$$

and the problem of finding the optimal state-space trajectory becomes the problem of finding x_2 such that (x_0, x_2) is in the stable invariant subspace $\mathcal{X}_-(\mathcal{A})$ of the Hamiltonian matrix \mathcal{A} . This is, in turn, the problem of writing

$$\mathcal{X}_-(\mathcal{A}) = \text{im} \begin{bmatrix} I \\ X \end{bmatrix}$$

For a similar derivation of a solution to a smoothing problem see Weinert and Desai (1981). \square

Remark 2.5: Along the same lines, one can obtain state-space formulas for operators of the form

$$\Pi_{(\Theta H_2)^\perp} M_{\Theta_1}, \quad \Theta \in H_\infty, \quad \Theta_1 \in L_\infty$$

Our original gap computations were based on the state-space expression of

$$\Pi_{\mathcal{G}(G)^\perp} \begin{bmatrix} I \\ G_1 \end{bmatrix}$$

They were not entirely satisfactory because the system obtained has higher McMillan degree than the one in (2.3). However, it is suggested that the direct translation of this type of operator to state-space equations may be useful in other contexts, as it plays an important role in the branch of operator theory that is connected with H_∞ optimization (cf. for instance Nikol'skii 1986). \square

3. Computing $\|\Pi_2\| = \|\Pi_{V_1^\perp}^V\|$

Recall

$$(1 - \vec{\delta}(V, V_1)^2)^{1/2} = \sin \varphi(V, V_1^\perp) = \|\Pi_{V_1^\perp}^V\|^{-1}$$

So we must compute the norm of Π_2 as defined in § 1. Two special cases must be singled out.

- (i) The input–output behaviour determined by the equations of Lemma 2.3 is not a mapping. Then, $V \cap V_1^\perp \neq \{0\}$, and the gap is 1. This occurs when the stable subspace of the \mathcal{A} -matrix of the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ of Lemma 2.3 has non-trivial intersection with the space $\{x|x_1 = 0\}$.
- (ii) The system is not minimal, or $I + DD_1^\top$ is singular. We have already shown that the system is minimal if $I + DD_1^\top$ is non-singular. If it is singular the gap is 1.

Lemma 3.1: *Suppose $I + DD_1^\top$ is singular. Then $\delta(\mathcal{G}(G), \mathcal{G}(G_1)) = 1$.*

Proof: Using (2.1), it can be seen that the L_∞ -norm of the mapping

$$\mathcal{P} := \Pi_{\mathcal{G}(G)}^{\mathcal{G}(-\tilde{G}_1)} = \begin{bmatrix} G \\ I \end{bmatrix} (I + \tilde{G}_1 G)^{-1} [\tilde{G}_1 \quad I]$$

is related in the same way to the L_2 -gap (i.e. the gap between the behaviours of the systems Σ and Σ_1 in $L_2(-\infty, \infty)$) as the norm of Π_2 is to the H_2 -gap between the graphs

$$[1 - \vec{\delta}(\mathcal{B}(G), \mathcal{B}(G_1))^2]^{1/2} = \|\Pi_{\mathcal{G}(G)}^{\mathcal{G}(-\tilde{G}_1)}\|_\infty^{-1}$$

The graph of $\mathcal{P} = \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B} + \mathcal{D}$ is a subspace of the graph of Π_2 . This can be seen from the fact that the graph of \mathcal{P} corresponds in the time domain to solutions, with all initial conditions zero, of the system in Lemma 2.3. It follows that its norm is not greater than $\|\Pi_2\|$. This implies that the L_2 -gap is, in general, smaller than the H_2 -gap. When $I + DD_1^\top$ is singular, \mathcal{P} is not in L_∞ , so the L_2 -gap and the H_2 -gap are both 1 in this case. \square

So, we may assume $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is minimal. As is well-known (cf. for instance, Boyd *et al.* 1989 for the computation of the H_∞ norm), computing the norm of an operator is related to solving optimal control problems with the cost criterion

$$\omega(y, u) = \gamma^2 \|u\|^2 - \|y\|^2$$

In this case, it is more convenient to state the solution directly in terms of the

associated hamiltonian matrix than to solve the problem in terms of Riccati equations. The hamiltonian is

$$\mathcal{H}_\gamma = \begin{bmatrix} \mathcal{A} & 0 \\ 0 & -\mathcal{A}^\top \end{bmatrix} + \begin{bmatrix} \mathcal{B} & 0 \\ 0 & -\mathcal{C}^\top \end{bmatrix} \begin{bmatrix} -\mathcal{D} & \gamma I \\ \gamma I & -\mathcal{D}^\top \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{C} & 0 \\ 0 & \mathcal{B}^\top \end{bmatrix}$$

Let W be the subspace $\{x|x_1 = 0\}$ of the state space X , and let \mathcal{X}_γ be the stable subspace of \mathcal{H}_γ . We are now in a position to state the main result of this note. In the appendix we summarize its use for the computation of the gap by a γ -iteration procedure.

Proposition 3.2: $\|\Pi_2\| < \gamma \Leftrightarrow \mathcal{H}_\gamma$ has no imaginary eigenvalues, there is a symmetric matrix K such that

$$\mathcal{X}_\gamma = \text{im} \begin{bmatrix} I \\ K \end{bmatrix}$$

and K is negative definite on W .

Proof: Define the available storage $V_a(x_0)$ of a minimal state-space system Σ with respect to ω as

$$V_a(x_0) = \inf \left\{ \int_0^\infty \gamma^2 \|u(t)\|^2 - \|y(t)\|^2 dt \mid \exists x(\cdot) : (x(\cdot), y(\cdot), u(\cdot)) \in \mathcal{B}_+(\Sigma) \wedge x(0) = x_0 \right\}$$

Let the optimal cost $K^+(x) = -V_a(x)$. It is well-known how to compute the optimal cost from a Riccati equation (cf., for instance, Trentelman and Willems 1991): if K is such that

$$\mathcal{X}_\gamma = \text{im} \begin{bmatrix} I \\ K \end{bmatrix}$$

then $\langle Kx, x \rangle = K^+(x)$. It is clear that optimal cost must be negative definite on the subspace $\{x|x_1 = 0\}$ for the parallel projection norm to be less than γ . \square

As already pointed out, the L_∞ norm of the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is related to the L_2 -gap

$$[1 - \bar{\delta}(\mathcal{B}(G), \mathcal{B}(G_1))^2]^{1/2} = \|\mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B} + \mathcal{D}\|_\infty^{-1}$$

So the smallest value of γ for which the hamiltonian has no imaginary eigenvalues corresponds to the L_2 -gap. As we already noted, the L_2 gap is, in general, smaller than the H_2 -gap; the topology it induces is weaker. It is the lower bound for the gap given by Zhu *et al.* (1989).

4. Criterion for $\delta(V, V_1) = 1$

After the computation of one directed gap, it must be checked whether the other directed gap is equal to it. It is known that the gap is 1 when the two directed gaps are not equal, so one needs to do just the one step in the γ -iteration for the computation of $\|\Pi_{V_1}^{V_1}\|$ to verify this. We give yet another criterion.

Lemma 4.1: Let Y, V_1 be closed subspaces of a Hilbert space. Assume $\bar{\delta}(V, V_1) < 1$, and $\bar{\delta}(V, V_1) \neq \bar{\delta}(V_1, V)$. Then $V^\perp \cap V_1 \neq \{0\}$.

Proof: This is not new (cf. Georgiou and Smith 1990, Krasnosel'skiĭ *et al.* 1972). It is known that the two directed gaps are equal unless the restricted orthogonal projection $\Pi_{V_1}|_V$ is not bijective. This is the same as $V^\perp \cap V_1 \neq \{0\}$. \square

Proposition 4.2: *Suppose V, V_1 are graphs of transfer functions with minimal state-space realizations (A, B, C, D) and (A_1, B_1, C_1, D_1) respectively. Then*

$$\vec{\delta}(V, V_1) < 1, \quad \vec{\delta}(V, V_1) \neq \vec{\delta}(V_1, V)$$

implies that the stable subspace of the matrix

$$\mathcal{A}' = \begin{bmatrix} -A^T + C^T(I + D_1 D^T)^{-1} D_1 B^T & -C^T(I + D_1 D^T)^{-1} C_1 \\ -B_1 B^T + B_1 D^T(I + D_1 D^T)^{-1} D_1 B^T & A_1 - B_1 D^T(I + D_1 D^T)^{-1} C_1 \end{bmatrix}$$

has non-trivial intersection with the subspace $\{x|x_2 = 0\}$.

Proof: Combine remark (i) of the previous section with Lemma 4.1. \square

To give a complete classification of the cases in which the gap is 1, we need the robustness margin for Fredholm indices of pairs of closed subspaces of a Banach space as given by Kato (1982)

$$\gamma(U, V) := \inf_{v \in V \setminus U} \frac{d(v, U)}{d(v, U \cap V)}$$

and the result

$$\gamma(U, V) > 0 \Leftrightarrow U + V \text{ is closed}$$

When $U \cap V = \{0\}$, $\gamma(U, V)$ is equal to $\sin \varphi(U, V) = \|\Pi_V^U\|^{-1}$.

Lemma 4.3: *Let $V = \mathcal{G}(G), V_1 = \mathcal{G}(G_1)$. Suppose $\delta(\mathcal{B}(G), \mathcal{B}(G_1)) < 1$. Then $\mathcal{G}(G) + \mathcal{G}(G_1)^\perp$ is closed, and $\vec{\delta}(V, V_1) = 1 \Rightarrow V \cap V_1^\perp \neq \{0\}$.*

Proof: It follows from $\delta_{L_2}(G, G_1) < 1$ that the mapping $\mathcal{P} = \Pi_{\mathcal{G}(G)}^{(-\tilde{G}_1)}$ is in L_∞ . This implies that $\mathcal{G}(G) + \mathcal{G}(-\tilde{G}_1)$ is closed, as $\gamma(\mathcal{G}(G), \mathcal{G}(-\tilde{G}_1)) = \|\mathcal{P}\|^{-1}$ when the intersection is trivial. But it follows from the relation $\mathcal{G}(F) = \mathcal{L}(B_+^0(F))$ that the co-dimension $[\mathcal{L}(\mathcal{B}_+(-\tilde{G}_1)) : \mathcal{G}(-\tilde{G}_1)] = [\mathcal{B}_+(-\tilde{G}_1) : B_+^0(-\tilde{G}_1)]$ is finite (in fact it is equal to the McMillan degree of G_1). This implies that $\mathcal{L}(\mathcal{B}_+(-\tilde{G}_1)) + \mathcal{G}(G)$ is closed. Hence, $\gamma(V, V_1) > 0$, so $(V \cap V_1^\perp = \{0\}) \Rightarrow \vec{\delta}(V, V_1) < 1$. \square

Lemma 4.4: *Let $V = \mathcal{G}(G), V_1 = \mathcal{G}(G_1)$, let \mathcal{A} and \mathcal{A}' be as in Lemmas 2.3 and 4.2 respectively. Then we have $\delta(V, V_1) = 1$ iff one of the following possibilities holds:*

- (i) *The stable space of \mathcal{A} has non-trivial intersection with $\{x|x_1 = 0\}$.*
- (ii) *The stable space of \mathcal{A}' has non-trivial intersection with $\{x|x_2 = 0\}$.*
- (iii) *\mathcal{A} has eigenvalues on the imaginary axis.*
- (iv) *\mathcal{A}' has eigenvalues on the imaginary axis.*
- (v) *$I + DD_1^T$ is singular.*

Proof: We proceed by elimination. It is clear that (iii)–(v) correspond to the L_2 gap being 1, and that (i), (ii) correspond respectively to $V \cap V_1^\perp \neq \{0\}, V_1 \cap V^\perp \neq \{0\}$. So all of (i)–(v) imply that the gap is 1. If none of (iii)–(v) holds, $\delta(\mathcal{B}(G), \mathcal{B}(G_1)) < 1$. By the previous lemma this implies that if neither (i) nor (ii) holds, then $\delta(V, V_1) < 1$. \square

Note that it suffices to calculate the eigenvalues of \mathcal{A} , as $-\lambda$ is an eigenvalue of \mathcal{A}' for each $\lambda \in \sigma(\mathcal{A})$, and vice versa. In fact, we can simplify the calculations a bit further. Let $\mathcal{X}_-(A)$ be the maximal stable invariant subspace of a matrix, with $\mathcal{X}_+(A)$ its antistable space.

Proposition 4.5: *We have $\delta(V, V_1) < 1$ iff $I + DD_1^T$ is non-singular, \mathcal{A} has no eigenvalues on the imaginary axis, and $\mathcal{X}_-(\mathcal{A})$ is complementary to $\{x|x_1 = 0\}$.*

Proof: It is easily verified that $\mathcal{A}' = -\mathcal{A}^T$. Now, since for matrices without eigenvalues on the imaginary axis we have $\mathcal{X}_-(X^T) = \mathcal{X}_+(X)^\perp$, it follows that

$$\mathcal{X}_-(\mathcal{A}') = (\mathcal{X}_-(\mathcal{A}))^\perp$$

Hence $\mathcal{X}_-(\mathcal{A}') \cap \{x|x_2 = 0\} \Leftrightarrow \mathcal{X}_-(\mathcal{A})^\perp \cap \{x|x_2 = 0\} = \{0\} \Leftrightarrow \mathcal{X}_-(\mathcal{A}) + \{x|x_1 = 0\}$ is the whole space. \square

An alternative to the gap that has the advantage of giving a less conservative estimate of robustness is the following (proposed by Vinnicombe 1993). For completeness we sketch the relation with the formulas of this section.

Definition 4.6: *The Vinnicombe gap δ_V between $V = \mathcal{G}(G)$ and $V_1 = \mathcal{G}(G_1)$ is given by the following definition: If $\dim V \cap V_1^\perp = \dim V_1 \cap V^\perp$ then $\delta_V(V, V_1) = \delta_{L_2}(V, V_1)$; else $\delta_V(V, V_1) = 1$.*

It can be shown that this distance measure is a metric equivalent to the gap, and that the robustness margin for the two metrics is identical. The next proposition shows how it can be computed in terms of the formulas of this paper.

Proposition 4.7: *We have $\delta_V(V, V_1) = 1$ iff*

- (i) $\delta_{L_2}(V, V_1) = 1$ or
- (ii) $\dim \mathcal{X}_-(\mathcal{A}) \cap \{x|x_1 = 0\} \neq \dim (X_1 + X_2) \ominus (\mathcal{X}_-(\mathcal{A}) + \{x|x_1 = 0\})$

Proof: Using the proof of the previous proposition, we can see that it is sufficient to see that

$$\dim V \cap V_1^\perp = \dim \mathcal{X}_-(\mathcal{A}) \cap \{x|x_1 = 0\}$$

This is not difficult to see from Lemma 2.3, for the space of stable outputs with input 0 of the system of equations given there has dimension equal to $\dim \mathcal{X}_-(\mathcal{A}) \cap \{x|x_1 = 0\}$. \square

So the Vinnicombe gap is somewhat easier to compute than the standard one, as the calculation consists of one test for the position of $\mathcal{X}_-(\mathcal{A})$ and the computation of an L_∞ norm, for which fast optimized routines are available.

Appendix

Below, we summarize the computation of the gap with accuracy ε in the form of a pseudo-Algol procedure. By $\mathcal{H}_\gamma \in \text{Dom}(\text{Ric})$ it is meant that the stable invariant subspace of \mathcal{H}_γ can be written in the form

$$\text{im} \begin{bmatrix} I \\ X \end{bmatrix}$$

and in this case $X = \text{Ric}(\mathcal{H}_\gamma)$.

The test for definiteness is likely to be numerically unreliable when X is

close to being singular, i.e. when we are close to the L_2 gap. We briefly compare this computation to the one that follows from Georgiou's (1988) reduction to a model matching problem. To do a computation in Georgiou's approach one needs to solve a Riccati equation in order to obtain the normalized coprime factorization, and then the resulting model matching problem can be solved using state-space formulas. An implementation of such a procedure is available, for instance, in the MATLAB package 'mu-tools'. The amount of work that needs to be done to solve a model matching problem in such a way is comparable with the running time of our procedure.

We have three sources for differences in efficiency.

- (1) We have a more efficient test for $\delta = 1$ that involves calculating the stable subspace of a matrix of dimension $n + n_1$ only (where n, n_1 are the degrees of the two original systems) and not the full size $2(n + n_1)$ dimension of the hamiltonian used in the iteration. In addition, we do not need any normalized or other factorizations.
- (2) The normalized coprime factorization does not have to be performed. A normalized coprime factorization is less expensive than one step in the bisection procedure; so, on the whole, this makes for some difference in efficiency in cases when the gap turns out to be 1, or no high accuracy is required and not many iteration steps are performed.
- (3) The problem $X = \text{Ric}(\mathcal{H})$ that has to be solved in both procedures is of the same dimension, but we only have to test the definiteness of one block of X .

We have compared a MATLAB procedure of our own to the one in 'mu-tools'. For fairness, our procedure relies on the same 'mu-tools' routine `ric_schr`. Sample systems $[A, B, C, D]$, $[A_1, B_1, C_1, D_1]$ were chosen according to the following procedure (not that it matters much): first, A, B, C, D were chosen randomly with parameters between 0 and 1. Then, a random weight $\rho \in [0, 1]$ was chosen, and a random perturbation $\Delta A, \Delta B, \Delta C, \Delta D$, again with parameters in $[0, 1]$. Our A_1, B_1, C_1, D_1 were then $A + \rho\Delta A, B + \rho\Delta B, C + \rho\Delta C, D + \rho\Delta D$. On a sample of 100 randomly chosen gap problems between systems of order 3, with accuracy 1/100, we had the following results.

Mutools. Average number of floating point operations: 737 350. The gap was computed as one 25 times; the number of operations needed to find out about this was, on average, 197 656. With the gap not equal to one, mutools needed on the average 917 248 flops.

Our own routine. It needed on average 590 422 operations. It decided 116 times that the gap was equal to one, and needed, on average, 67 184 flops to find out. The average number of operations needed for gaps not equal to 1 was 690 086.

So, we may conclude there is some increase in efficiency, although it is not exactly sensational.

real $\delta(A, B, C, D, A_1, B_1, C_1, D_1, \varepsilon)$ /* Assumes minimality of both systems */
begin

if Singular($I + DD_1^T$) **then return** 1 **fi**

($\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$) := SkewProjection($A, B, C, D, A_1, B_1, C_1, D_1$)/* Form the

```

system of Lemma 2.3 */
n := dim  $\mathcal{A}$ 
N := dim  $\mathcal{A}$ 

/* First check whether  $\delta = 1$  as in § 4 */

Perform an ordered (complex) Schur decomposition  $\mathcal{A} = UTU^*$ 
i := 1
while  $i \leq N \wedge \operatorname{Re} T_{ii} < 0$  do  $i := i + 1$  od
if  $i \neq m + 1 \vee \operatorname{Re} T_{ii} = 0 \vee \operatorname{Singular}(U_{1:m,1:m})$  return 1 fi

/* Now do the  $\gamma$ -iteration of § 3 */

 $\delta_- := \left(1 - \frac{1}{\|\mathcal{D}\|^2}\right)^{1/2}$ 
 $\delta_+ := 1$ 
Error :=  $\delta_+ - \delta_-$ 
 $\delta := 1$ 
while Error >  $\varepsilon$  do
   $\delta := (\delta_- + \delta_+)/2$ 
   $\gamma := \frac{1}{(1 - \delta^2)^{1/2}}$ 
   $\mathcal{H}_\gamma := \begin{bmatrix} \mathcal{A} & 0 \\ 0 & -\mathcal{A}^T \end{bmatrix} + \begin{bmatrix} \mathcal{B} & 0 \\ 0 & -\mathcal{C}^T \end{bmatrix} \begin{bmatrix} -\mathcal{Q} & \gamma I \\ \gamma I & -\mathcal{Q}^T \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{C} & 0 \\ 0 & \mathcal{B}^T \end{bmatrix}$ 
  if  $\mathcal{H}_\gamma \in \operatorname{Dom}(\operatorname{Ric})$  then
    SolutionExists := true
     $X := \operatorname{Ric}(\mathcal{H}_\gamma)$ 
    SolutionDefinite := NegativeDefinite( $X_{m+1:n, m+1:n}$ )
  else
    SolutionDefinite := SolutionExists := false
  fi
  if SolutionExists  $\wedge$  SolutionDefinite then  $\delta_+ := \delta$  else  $\delta_- := \delta$  fi
  Error :=  $\delta_+ - \delta_-$ 
od
return  $\delta$ 
end

```

REFERENCES

- BOYD, S., BALAKRISHNAN, V., and KABAMBA, P., 1989, A bisection method for computing the H_∞ norm of a transfer matrix and related problems. *Mathematics of Control, Signals and Systems*, **2**, 207–219.
- FOIAS, C., GEORGIU, T. T., and SMITH, M. C., 1991, Robust stabilization in the gap metric: a geometric approach. *1991 International Symposium on the Mathematical Theory of Networks and Systems*, Kobe, Japan.
- GEORGIU, T. T., 1988, On the computation of the gap metric. *Systems & Control Letters*, **11**, 253–257.
- GEORGIU, T. T., and SMITH, M. C., 1990, Optimal robustness in the gap metric. *IEEE Transactions on Automatic Control*, **35**, 673–687.
- KATO, T., 1982, *Perturbation Theory for Linear Operators* (New York: Springer-Verlag).
- KRASNOSEL'SKIĬ, M. A., VAINIKKO, G. M., ZABREIKO, P. P., RUTITSKIĬ, YA. B., STETSSENKO, V. YA., 1972, *Approximate Solution of Operator Equations*. Groningen.

- OBER, R. J., and SEFTON, J. A., 1991, Stability of control systems and graphs of linear systems, *Systems & Control Letters*, **17**, 265–280.
- NIKOL'SKII, N., 1986, *Treatise on the Shift Operator*. Grundlehren der Mathematischen Wissenschaft (Berlin: Springer-Verlag).
- SCHUMACHER, J. M., 1992, A pointwise criterion for controller robustness. *Systems & Control Letters*, **18**, 1–8.
- TRENTELMAN, H. L., and WILLEMS, J. C., 1991, The dissipation inequality and the algebraic Riccati equation. *The Riccati Equation*, edited by S. Bittanti, A. J. Laub, and J. C. Willems (Berlin: Springer-Verlag), pp. 197–242.
- VIDYASAGAR, M., 1985, *Control System Synthesis, a Factorization Approach* (New York: MIT Press).
- VINNCOMBE, G., 1993, Frequency domain uncertainty and the graph topology. *IEEE Transactions on Automatic Control*, **38**, 1371–1383.
- WEINERT, H. L., and DESAI, U. B., 1981, On complementary models and fixed-interval smoothing. *IEEE Transactions on Automatic Control*, **26**, 863–867.
- WILLEMS, J. C., 1991, Paradigms and puzzles in the theory of dynamical systems. *IEEE Transactions on Automatic Control*, **36**, 259–294.
- ZHU, S. Q., HAUTUS, M. L. J., PRAAGMAN, C., 1989, A lower and an upper bound for the gap metric. *Proceedings of the 28th Conference on Decision and Control*, pp. 2337–2341.