Equivalence of a behavioral distance and the gap metric

Joseph A. Ball
Department of Mathematics, Virginia Tech.,
Blacksburg, VA 24061, U.S.A.
E-mail: ball@math.vt.edu
and
Amol J. Sasane
Department of Mathematics, London School of Economics,
Houghton Street, London WC2A 2AE, U.K.
E-mail: A.J.Sasane@lse.ac.uk

CDAM Research Report CDAM-LSE-2004-17

November 2004

Abstract: In this note it is shown that the topology of the gap metric is precisely the topology induced by the behavioral distance introduced in [10], when restricted to the class of stable state space systems with a fixed MacMillan degree.

Keywords: Behaviors, distance, gap metric, state linear systems.

Mathematics Subject Classification numbers: 93A99

1 Introduction

The need to measure the distance between two systems is basic in control theory. Indeed, it arises naturally when one talks of approximating a system with another, for instance in the context of the important engineering problem of model reduction. In robust control theory, one investigates the uncertainties that can be tolerated in a system without loss of characteristics such as stability under the application of feedback.

In the classical Kalman finite dimensional state space theory, the gap metric (denoted throughout this note by δ) serves as a tool for the qualitative analysis and design of feedback systems (see for instance, Georgiou and Smith [4], Zames and El-Sakkary [13]). It is the weakest topology in which closed loop stability is a robust property, or in which the closed loop system varies continuously as a function of the open loop system.

In the behavioral setting of Willems (see Polderman and Willems [7] for an elementary introduction), a notion of distance was introduced in [10]. Here the set of controllable (distributional) behaviors ($\mathscr{L}_c^q$) was equipped with a metric d such that $(\mathscr{L}_c^q, d)$ is a metric space. In the case of state space systems, that is systems described by

$$\Sigma : \left\{ \begin{array}{rcl} \frac{dx}{dt}(t) & = & Ax(t) + Bu(t), \\ y(t) & = & Cx(t) + Du(t), \end{array} \right.$$

this behavioral distance d between two minimal stable systems of state space dimension n, number of inputs equal to m and number of outputs equal to p, was expressed in terms of the gap between a natural (closed) subspace of $H_2(\mathbb{C}_+, \mathbb{C}^{m+p})$ associated with the linear system, namely:

$$\mathscr{G}_e(\Sigma) = \left[\begin{array}{c} I \\ M_G \end{array} \right] H_2(\mathbb{C}_+, \mathbb{C}^m) + \left[\begin{array}{c} 0 \\ C(\cdot I - A)^{-1} \end{array} \right] \mathbb{C}^n,$$

where $M_G : H_2(\mathbb{C}_+, \mathbb{C}^m) \to H_2(\mathbb{C}_+, \mathbb{C}^p)$ denotes the multiplication map by the transfer function $G(\cdot) = D + C(\cdot I - A)^{-1}B \in H_\infty(\mathbb{C}_+, \mathbb{C}^{p\times m})$. This subspace $\mathscr{G}_e(\Sigma)$ differs from the classical graph of the system by

$$\left[\begin{array}{c} 0 \\ C(\cdot I - A)^{-1} \end{array} \right] \mathbb{C}^n. \tag{1}$$

The presence of such an extra subspace is natural, since in the behavioral framework, as opposed to the traditional transfer function set up, instead of the initial conditions being zero, one expects a term in the graph of the system which reflects all possible initial conditions: (1) is precisely that, since for given $u \in L_2((0,\infty), \mathbb{C}^p)$ and $x_0 \in \mathbb{C}^n$, the Laplace transform of the output with this input u and the initial state x_0 is given by $G(s)\hat{u}(s) + C(sI - A)^{-1}x_0$.

The question of whether these two metrics, namely the behavioral distance d and the gap metric δ, are equivalent, is a natural one. In this note we prove that in fact the topology induced by the gap metric is the same as the topology induced by the behavioral metric when restricted to the class of stable state space systems with a fixed MacMillan degree.

The outline of this article is as follows. In §2, we recall a few facts about the gap metric. In the next section we prove that the topology induced by the behavioral distance is stronger/finer than that induced by the gap metric. Subsequently in §4, we prove the converse result: using a special realization of the transfer function obtained via an extremal factorization of the Hankel operator, we show that the topology induced by the behavioral distance is weaker/coarser than that induced by the gap metric. The results from §3 and §4 are summarized in the final §5, where we state our main theorem concerning the equivalence of the gap metric and the behavioral distance.

2 Preliminaries

In this section we quickly recall the definitions of the behavioral distance d and the gap metric δ.

Behaviors. We will use the following standard notation: $\mathscr{D}(\mathbb{R})$ denotes the space of test functions, that is the set of compactly supported infinitely many times differentiable functions; $\mathscr{D}'(\mathbb{R})$

denotes the space of distributions on $\mathbb{R}$; $L_2(\mathbb{R},\mathbb{C}^q)$ denotes the space of $\mathbb{C}^q$-valued square integrable functions on $\mathbb{R}$. Consider the polynomial matrix

$$R = \begin{bmatrix} r_{11} & \dots & r_{1q} \\ \vdots & & \vdots \\ r_{g1} & \dots & r_{gq} \end{bmatrix} \in \mathbb{C}[\xi]^{g\times q}.$$

The polynomial matrix R gives rise to a map $D_R : \mathscr{D}'(\mathbb{R})^q \to \mathscr{D}'(\mathbb{R})^g$, which acts as follows:

$$D_R \begin{bmatrix} w_1 \\ \vdots \\ w_q \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{q} r_{1k}\left(\frac{d}{dt}\right) w_k \\ \vdots \\ \sum_{k=1}^{q} r_{gk}\left(\frac{d}{dt}\right) w_k \end{bmatrix}.$$

The *behavior described by a mixed representation given by* (R,M), where $R \in \mathbb{C}[\xi]^{g\times q}$ and $M \in \mathbb{C}[\xi]^{g\times l}$ is defined to be

$$\mathfrak{B}(R,M) = \left\{ w \in \mathscr{D}'(\mathbb{R})^q \mid \text{there exists } \ l \in \mathscr{D}'(\mathbb{R})^l \ \text{ such that } \ D_R w = D_M l \right\}.$$

By a *behavior* we then mean a behavior described by a mixed representation given by (R,M), where R and M are some polynomial matrices in $\mathbb{C}[\xi]^{g\times q}$ and $\mathbb{C}[\xi]^{g\times l}$, respectively. The set of all behaviors in $\mathscr{D}'(\mathbb{R})^q$ that are given by a mixed representation given by some pair (R,M) will be denoted by $\mathscr{L}^q$.

A behavior $\mathfrak{B}$ is said to be *controllable* if for every $w_1, w_2 \in \mathfrak{B}$ and every pair of open sets O_1, O_2 in $\mathbb{R}$ with disjoint closures ($\overline{O_1} \cap \overline{O_2} = \emptyset$), there exists a $w \in \mathfrak{B}$ such that $w|_{O_1} = w_1|_{O_1}$ and $w|_{O_2} = w_2|_{O_2}$. The subset of $\mathscr{L}^q$ comprising controllable behaviors will be denoted by $\mathscr{L}^q_c$.

Distance between behaviors. Given a behavior $\mathfrak{B}$, let $\mathfrak{B}^0 := \{w \in \mathfrak{B} \cap \mathscr{D}(\mathbb{R})^q\}$. If A is a subset of $L_2(\mathbb{R},\mathbb{C}^q)$, then we denote the distance of a point $w \in L_2(\mathbb{R},\mathbb{C}^q)$ to A by $\text{dist}(w,A)$:

$$\text{dist}(w,A) = \inf\{\|w - \tilde{w}\|_{L_2} \mid \tilde{w} \in A\}.$$

If A is a subset of $L_2(\mathbb{R},\mathbb{C}^q)$, then S_A denotes the set $\{w \in A \mid \|w\|_{L_2} = 1\}$. Define

$$\delta(\mathfrak{B}_1,\mathfrak{B}_2) = \sup\nolimits_{w_1 \in S_{\mathfrak{B}_1^0}} \text{dist}(w_1,\mathfrak{B}_2^0), \tag{2}$$

$$d(\mathfrak{B}_1,\mathfrak{B}_2) = \max\{\delta(\mathfrak{B}_1,\mathfrak{B}_2), \delta(\mathfrak{B}_2,\mathfrak{B}_1)\}. \tag{3}$$

It is clear that if $\mathfrak{B}_1^0 \neq 0$, then $\delta(\mathfrak{B}_1, 0) = 0$. However, (2) has no meaning if $\mathfrak{B}_1^0 = 0$. In this case we simply set $\delta(0,\mathfrak{B}_2) = 0$. The function d is called the *behavioral distance*. In [10], it was shown that $(\mathscr{L}^q_c, d)$ is a metric space. Also in the case of stable state space systems, it was shown that the behavioral distance turns out to be equal to the gap between certain natural subspaces (of the Hardy space $H_2(\mathbb{C}_+,\mathbb{C}^{m+p})$) associated with the system. But before we quote this result from [10], we recall a few more definitions.

The gap between subspaces of a Hilbert space. Given two closed subspaces $\mathscr{V}_1$ and $\mathscr{V}_2$ of a Hilbert space $\mathscr{H}$, one defines the *gap*, denoted by g, between $\mathscr{V}_1$ and $\mathscr{V}_2$ as follows:

$$\mathrm{g}(\mathscr{V}_1, \mathscr{V}_2) = \left\| \Pi_{\mathscr{V}_1} - \Pi_{\mathscr{V}_2} \right\|,$$

$\Pi_{\mathscr{V}_i} : \mathscr{H} \to \mathscr{H}$ denote the projections onto $\mathscr{V}_i$, $i \in \{1, 2\}$. It can be verified that g makes the set of all closed linear subspaces of a Hilbert space into a (complete) metric space. Furthermore, it can also be shown that

$$\mathrm{g}(\mathscr{V}_1, \mathscr{V}_2) = \max\{\vec{\mathrm{g}}(\mathscr{V}_1, \mathscr{V}_2), \vec{\mathrm{g}}(\mathscr{V}_2, \mathscr{V}_1)\},$$

where

$$\vec{\mathrm{g}}(\mathscr{V}_1, \mathscr{V}_2) = \left\| (I - \Pi_{\mathscr{V}_2}) \Pi_{\mathscr{V}_1} \right\| = \sup_{v \in \mathscr{V}_1,\ \|v\|=1} \mathrm{dist}(v, \mathscr{V}_2)$$

is the *directed-gap*. For more details about the gap metric, we refer the reader to Kato [6] and the references therein.

Behavioral distance between state space systems. We denote the open right half complex plane by $\mathbb{C}_+$, that is, $\mathbb{C}_+ = \{s \in \mathbb{C} \mid \mathrm{Re}(s) > 0\}$. If $(\mathscr{E}, \|\cdot\|_{\mathscr{E}})$ is a Banach space, then

$$H_\infty(\mathbb{C}_+, \mathscr{E}) = \left\{ f : \mathbb{C}_+ \to \mathscr{E} \;\middle|\; f \text{ is analytic and } \sup_{\mathrm{Re}(s)>0} \|f(s)\|_{\mathscr{E}} < \infty \right\}.$$

If $(\mathscr{H}, \langle \cdot, \cdot \rangle_{\mathscr{H}})$ is a Hilbert space, then let

$$H_2(\mathbb{C}_+, \mathscr{H}) := \left\{ f : \mathbb{C}_+ \to \mathscr{H} \;\middle|\; f \text{ is analytic and } \|f\|_{H_2(\mathbb{C}_+, \mathscr{H})} = \sup_{\zeta>0} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \|f(\zeta + i\omega)\|_{\mathscr{H}}^2 d\omega \right) < \infty \right\}.$$

It can be shown that each $f \in H_2(\mathbb{C}_+, \mathscr{H})$, there exists a unique $\tilde{f} \in L_2(i\mathbb{R}, \mathscr{H})$ such that

$$\lim_{\zeta \downarrow 0} f(\zeta + i\omega) = \tilde{f}(i\omega) \text{ for almost all } \omega \in \mathbb{R} \text{ and } \lim_{\zeta \downarrow 0} \|f(\zeta + \cdot) - \tilde{f}(\cdot)\|_{L_2(i\mathbb{R}, \mathscr{H})} = 0.$$

The Hardy space $H_2(\mathbb{C}_+, \mathscr{H})$ is a Hilbert space with the inner product defined by

$$\langle f, g \rangle_{H_2(\mathbb{C}_+, \mathscr{H})} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle \tilde{f}(i\omega), \tilde{g}(i\omega) \rangle_{\mathscr{H}} d\omega.$$

Given a linear system

$$\Sigma : \left\{ \begin{array}{rcl} \frac{dx}{dt}(t) & = & Ax(t) + Bu(t) \\ y(t) & = & Cx(t) + Du(t) \end{array} \right.$$

where $A \in \mathbb{C}^{n\times n}$, $\sigma(A) \subset \{s \in \mathbb{C} \mid \mathrm{Re}(s) < 0\}$, $B \in \mathbb{C}^{n\times m}$, $C \in \mathbb{C}^{p\times n}$ and $D \in \mathbb{C}^{p\times m}$, let its transfer function be denoted by G:

$$G(s) = D + C(sI - A)^{-1}B, \;\; s \in \mathbb{C}_+.$$

We define the *graph* of the system Σ by

$$\mathscr{G}(\Sigma)=\left\{\left[\begin{array}{c} I \\ M_G \end{array}\right]u \,\middle|\, u\in H_2(\mathbb{C}_+,\mathbb{C}^m)\right\},$$

where $M_G : H_2(\mathbb{C}_+,\mathbb{C}^m)\to H_2(\mathbb{C}_+,\mathbb{C}^p)$ denotes the multiplication map by $G\in H_\infty(\mathbb{C}_+,\mathbb{C}^{p\times m})$:

$$(M_G u)(s)=G(s)u(s),\ \ s\in\mathbb{C}_+,\ \text{for } u\in H_2(\mathbb{C}_+,\mathbb{C}^m).$$

This is a closed subspace of $H_2(\mathbb{C}_+,\mathbb{C}^{m+p})$. The *extended graph* of the system Σ is defined as follows:

$$\mathscr{G}_e(\Sigma)=\left\{\left[\begin{array}{c} I \\ M_G \end{array}\right]u+\left[\begin{array}{c} 0 \\ C(sI-A)^{-1} \end{array}\right]x \,\middle|\, u\in H_2(\mathbb{C}_+,\mathbb{C}^m),\ x\in\mathbb{C}^n\right\}.$$

It is easy to see that this is a closed subspace of $H_2(\mathbb{C}_+,\mathbb{C}^{m+p})$. We recall the following from [10] (see Theorem 7 on page 1221):

Theorem 2.1 *Let Σ_1 and Σ_2 be the following two stable, minimal state space systems with state space dimension n, number of inputs equal to m and number of outputs equal to p:*

$$\Sigma_1:\left\{\begin{array}{rcl}\frac{dx_1}{dt}(t) &=& A_1x_1(t)+B_1u_1(t)\\ y_1(t) &=& C_1x_1(t)+D_1u_1(t)\end{array}\right. \ \ \textit{and}\ \ \Sigma_2:\left\{\begin{array}{rcl}\frac{dx_2}{dt}(t) &=& A_2x_2(t)+B_2u_2(t)\\ y_2(t) &=& C_2x_2(t)+D_2u_2(t)\end{array}\right. .$$

Let $\mathfrak{B}_1$ and $\mathfrak{B}_2$ denote the behaviors described by mixed representation given by (R_1,M_1), (R_2,M_2), where

$$R_k=\left[\begin{array}{cc} I & -D_k \\ 0 & B_k \end{array}\right] \textit{ and } M_k=\left[\begin{array}{c} C_k \\ \xi I-A_k \end{array}\right],\ \ k\in\{1,2\}.$$

Then

$$d(\mathfrak{B}_1,\mathfrak{B}_2)=\mathrm{g}(\mathscr{G}_e(\Sigma_1),\mathscr{G}_e(\Sigma_2)).$$

The classical gap between systems and the set S. In the classical state space theory, the *gap metric* between two systems of the type in the above Theorem 2.1 is defined to be the gap between the corresponding graphs, that is

$$\delta(\Sigma_1,\Sigma_2)=\mathrm{g}(\mathscr{G}(\Sigma_1),\mathscr{G}(\Sigma_2)).$$

Let $\mathbf{S}$ denote the set of stable, minimal state space systems Σ with state space dimension n, number of inputs equal to m and number of outputs equal to p. It is known that in the set $\mathbf{S}$, convergence in the gap metric δ is the same as convergence in $H_\infty(\mathbb{C}_+,\mathbb{C}^{p\times m})$ (see for instance Georgiou and Smith [5]):

Theorem 2.2 *If $(\Sigma_k)_{k\geq 1}$ is a sequence in $\mathbf{S}$, $\Sigma\in\mathbf{S}$ and then the following are equivalent:*

1. $\mathrm{g}(\mathscr{G}(\Sigma_k),\mathscr{G}(\Sigma))\to 0$ *as* $k\to\infty$,
2. $\|G_k-G\|_{H_\infty(\mathbb{C}_+,\mathbb{C}^{p\times m})}\to 0$ *as* $k\to\infty$.

In this note we prove that d and δ are equivalent on $\mathbf{S}$.

3 The topology induced by the behavioral distance is finer than that induced by the gap metric on the set S.

We prove that the topology induced by the behavioral distance is finer than that induced by the gap metric on the set **S** by appealing to Theorem 2.2: we show in Theorem 3.3 that if the extended graphs converge in the gap topology of subspaces, then this implies that the transfer functions converge in the H_∞ norm.

We begin by proving two preliminary lemmas which will be used in proving Theorem 2.2: in Lemma 3.1, we express the gap between two extended graphs as the gap between the graphs of multiplication operators on the orthogonal complement of the range of the observability map in the frequency domain, and in Lemma 3.2, we express the orthogonal complement of the range of the observability map as the range of a multiplication map by an inner function.

We first fix some notation: for $K \in H_\infty(\mathbb{C}_+, \mathbb{C}^{k_2\times k_1})$, the linear map $M_K : H_2(\mathbb{C}_+, \mathbb{C}_+^{k_1}) \to H_2(\mathbb{C}_+, \mathbb{C}^{k_2})$ denotes the analytic Toeplitz operator of multiplication by K:

$$M_K : f \mapsto Kf \in H_2(\mathbb{C}_+, \mathbb{C}^{k_2}) \text{ for } f \in H_2(\mathbb{C}_+, \mathbb{C}^{k_1}).$$

The adjoint operator $M_K^* : H_2(\mathbb{C}_+, \mathbb{C}^{k_2}) \to H_2(\mathbb{C}_+, \mathbb{C}^{k_1})$ is then given by

$$M_K^* : f \mapsto P_{H_2(\mathbb{C}_+, \mathbb{C}^{k_1})}(K^* f) \text{ for } f \in H_2(\mathbb{C}_+, \mathbb{C}^{k_2}),$$

where K^* is the matrix-valued function

$$K^*(s) = K(-\bar{s})^*,$$

and $P_{H_2(\mathbb{C}_+, \mathbb{C}^{k_1})} : L_2(i\mathbb{R}, \mathbb{C}^{k_1}) \to H_2(\mathbb{C}_+, \mathbb{C}^{k_1})$ denotes the projection onto the closed subspace $H_2(\mathbb{C}_+, \mathbb{C}^{k_1})$ of $L_2(i\mathbb{R}, \mathbb{C}^{k_1})$.

Let $\mathscr{C} : \mathbb{C}^n \to L_2([0,\infty), \mathbb{C}^p)$ denote the observability map

$$x \mapsto Ce^{\cdot A}x, \;\; x \in \mathbb{C}^n,$$

(where $t \mapsto e^{\cdot A}$ is the (stable) strongly continuous semigroup with infinitesimal generator A) and $\widehat{\mathscr{C}} = \mathscr{F} \circ \mathscr{C}$ with $\mathscr{F}$ equal to the Fourier transform, so

$$\widehat{\mathscr{C}} : x \mapsto C(sI - A)^{-1}x, \;\; x \in \mathbb{C}^n.$$

Lemma 3.1 *If $\Sigma_k, \Sigma \in$ **S**, then*

$$\mathrm{g}(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) = \mathrm{g}\left(\begin{bmatrix} -M_{G_k}^* \\ I \end{bmatrix} (\mathrm{ran}\, \widehat{\mathscr{C}_k})^\perp, \begin{bmatrix} -M_G^* \\ I \end{bmatrix} (\mathrm{ran}\, \widehat{\mathscr{C}})^\perp \right).$$

Proof Using Theorem 2.9 on page 201 of Kato [6], we know that

$$\mathrm{g}(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) = \mathrm{g}((\mathscr{G}_e(\Sigma_k))^\perp, (\mathscr{G}_e(\Sigma))^\perp).$$

So the claim would be proved if we show that for any $\Sigma_0 \in \mathbf{S}$,

$$(\mathscr{G}_e(\Sigma_0))^\perp = \begin{bmatrix} -M_{G_0}^* \\ I \end{bmatrix} (\text{ran } \widehat{\mathscr{C}_0})^\perp.$$

If $\begin{bmatrix} u_0 \\ y_0 \end{bmatrix} \in (\mathscr{G}_e(\Sigma_0))^\perp$, then for all $u \in H_2(\mathbb{C}_+, \mathbb{C}^m)$ and all $x \in \mathbb{C}^n$, there holds that

$$\begin{aligned} 0 &= \langle u_0, u\rangle + \langle y_0, M_{G_0}u + \widehat{\mathscr{C}_0}x\rangle \\ &= \langle u_0, u\rangle + \langle y_0, M_{G_0}u\rangle + \langle y_0, \widehat{\mathscr{C}_0}x\rangle \\ &= \langle u_0, u\rangle + \langle M_{G_0}^* y_0, u\rangle + \langle y_0, \widehat{\mathscr{C}_0}x\rangle \\ &= \langle u_0 + M_{G_0}^* y_0, u\rangle + \langle y_0, \widehat{\mathscr{C}_0}x\rangle. \end{aligned} \tag{4}$$

In particular, with $u = 0$, we obtain that $\langle y_0, \widehat{\mathscr{C}_0}x\rangle = 0$ for all $x \in \mathbb{C}^n$ and so $y_0 \in (\text{ran } \widehat{\mathscr{C}_0})^\perp$. From (4), it now follows that since $\langle u_0 + M_{G_0}^* y_0, u\rangle = 0$ for all $u \in H_2(\mathbb{C}_+, \mathbb{C}^m)$, there holds that $u_0 + M_{G_0}^* y_0 = 0$, that is, $u_0 = -M_{G_0}^* y_0$. Consequently,

$$\begin{bmatrix} u_0 \\ y_0 \end{bmatrix} \in \begin{bmatrix} -M_{G_0}^* \\ I \end{bmatrix} (\text{ran } \widehat{\mathscr{C}_0})^\perp.$$

Conversely, if $\begin{bmatrix} u_0 \\ y_0 \end{bmatrix} \in \begin{bmatrix} -M_{G_0}^* \\ I \end{bmatrix} (\text{ran } \widehat{\mathscr{C}_0})^\perp$, then $y_0 \in (\text{ran } \widehat{\mathscr{C}_0})^\perp$ and $u_0 = -M_{G_0}^* y_0$. Let $\begin{bmatrix} u \\ y \end{bmatrix} \in \mathscr{G}_e(\Sigma_0)$, that is $y = M_{G_0}u + \widehat{\mathscr{C}_0}x$ for some $x \in \mathbb{C}^n$. Then

$$\begin{aligned} \left\langle \begin{bmatrix} u_0 \\ y_0 \end{bmatrix}, \begin{bmatrix} u \\ y \end{bmatrix} \right\rangle &= \langle u_0, u\rangle + \langle y_0, y\rangle \\ &= \langle -M_{G_0}^* y_0, u\rangle + \langle y_0, M_{G_0}u + \widehat{\mathscr{C}_0}x\rangle \\ &= -\langle y_0, M_{G_0}u\rangle + \langle y_0, M_{G_0}u\rangle + \langle y_0, \widehat{\mathscr{C}_0}x\rangle \\ &= 0, \end{aligned}$$

and so $\begin{bmatrix} u_0 \\ y_0 \end{bmatrix} \in (\mathscr{G}_e(\Sigma_0))^\perp$. This completes the proof. ∎

Lemma 3.2 *If $\Sigma \in \mathbf{S}$, then there exists an inner $\Theta \in H_\infty(\mathbb{C}_+, \mathbb{C}^{p\times p})$ and a $F \in H_\infty(\mathbb{C}_+, \mathbb{C}^{m\times p})$ such that $G = \Theta F^*$ and $(\text{ran } \widehat{\mathscr{C}})^\perp = M_\Theta H_2(\mathbb{C}_+, \mathbb{C}^p)$.*

Proof Consider first the scalar case $m = p = 1$. We assume in addition that the poles of G are all simple. Then $G(s)$ has a partial fraction representation $G(s) = D + \sum_{j=1}^n \frac{r_j}{s - p_j}$ with distinct poles $p_1, \dots, p_n$ in the right half plane. Then $G(s) = D + C(sI - A)^{-1}B$ is a minimal realization for G with

$$C = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}, \qquad A = \begin{bmatrix} p_1 & & \\ & \ddots & \\ & & p_n \end{bmatrix}, \qquad B = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}.$$

Let $\Theta(s) = \prod_{j=1}^{n} \frac{s+\overline{p_j}}{s-p_j}$ be the inner function with poles at $p_1, \ldots, p_n$. Set $F(s) = \Theta(s)\overline{G(-\bar{s})}$. Then we see that the zeros of $\Theta(s)$ cancel out the poles of $\overline{G(-\bar{s})}$ in $\mathbb{C}_+$ and hence $F \in H_\infty(\mathbb{C}_+)$. Moreover we have the representation $G(s) = \Theta(s)\overline{F(-\bar{s})}$ for G.

Note next that

$$\operatorname{ran} \widehat{\mathscr{C}} = \left\{ \frac{c_1}{s-p_1} + \cdots + \frac{c_n}{s-p_n} \;\middle|\; c_j \in \mathbb{C} \text{ for } 1 \le j \le n \right\}.$$

Thus $f \in (\operatorname{ran} \widehat{\mathscr{C}})^\perp$ means that $f \perp \frac{1}{\cdot - p_j}$, or $\int_{i\mathbb{R}} \frac{1}{s+\overline{p_j}} f(s)\, ds = 0$, for $j = 1, \ldots, n$. Viewing the integral as a contour integral and using the Residue Theorem, we see that this is equivalent to $f(-\overline{p_j}) = 0$ for $j = 1, \ldots, n$. This in turn amounts to f having a factorization $f = \Theta g$ with g analytic on $\mathbb{C}_+$. We conclude that $(\operatorname{ran} \widehat{\mathscr{C}})^\perp = \Theta H^2(\mathbb{C}_+)$, and the lemma is proved for the scalar simple-pole case

For the general case, the ideas are the same but it is convenient to use the formalism from [2] to handle the additional matrix zero-pole structure. Let $G(s) = D + C(sI-A)^{-1}B$ be a minimal realization for G. As we are assuming that G is stable, $\sigma(A) \subset \{s \in \mathbb{C} \mid \operatorname{Re}(s) < 0\}$. Let $\Theta(s)$ be the $p \times p$-matrix inner function having right pole pair (C, A), i.e., $\Theta(s) = I - C(sI-A)^{-1}H^{-1}C^*$ where H is the unique solution of the Lyapunov equation

$$HA + A^*H + C^*C = 0 \tag{5}$$

(see [2, Theorem 6.1.4]). As Θ and G have the same right pole pair (C, A) over $\mathbb{C}_-$ and $\Theta(s)^{-1} = \Theta(-\bar{s})^*$ is analytic on $\mathbb{C}_-$, it follows that G has a factorization $G = \Theta F'$ with F' analytic on $\mathbb{C}_-$ (see [2, Proposition 12.1.1] for a precise statement). If we then set $F(s) = F'(-\bar{s})^*$, we have $G(s) = \Theta(s)F(-\bar{s})^*$ with $F \in H_\infty(\mathbb{C}_+, \mathbb{C}^{m\times p})$.

By definition $\operatorname{ran} \widehat{\mathscr{C}} = \{C(\cdot I - A)^{-1}x \mid x \in \mathbb{C}^n\}$. Thus $f \in (\operatorname{ran} \widehat{\mathscr{C}})^\perp$ means that

$$\int_{i\mathbb{R}} (sI + A^*)^{-1} C^* f(s)\, ds = 0.$$

From Theorem 12.3.1 in [2] (using that a $\mathbb{C}_+$-null-pole-triple for Θ is $(0, 0; -A^*, C^*; 0)$), we see that this condition is equivalent to f having a factorization as $f = \Theta g$ with g analytic on $\mathbb{C}_+$. We conclude that $(\operatorname{ran} \widehat{\mathscr{C}})^\perp = \Theta H_2(\mathbb{C}_+, \mathbb{C}^p)$ as asserted. ■

Using the results from Lemmas 3.1 and 3.2, we are now ready to prove the following result.

Theorem 3.3 *Let $(\Sigma_k)_{k\ge1}$ be a sequence of systems in $\mathbf{S}$ and let $\Sigma \in \mathbf{S}$. For each $k \ge 1$, let G_k denote the transfer function of Σ_k and let let G denote the transfer function of Σ.*

If

$$\mathrm{g}(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) \to 0 \text{ as } k \to \infty, \tag{6}$$

then

$$\|G_k - G\|_{H_\infty(\mathbb{C}_+, \mathbb{C}^{p\times m})} \to 0 \text{ as } k \to \infty. \tag{7}$$

Proof The proof is long and so we have divided it into a sequence of steps.

STEP 1. From Lemma 3.1, it follows that

$$g(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) = g\left(\begin{bmatrix} -M_{G_k}^* \\ I \end{bmatrix} (\text{ran } \widehat{\mathscr{C}_k})^\perp, \begin{bmatrix} -M_G^* \\ I \end{bmatrix} (\text{ran } \widehat{\mathscr{C}})^\perp \right).$$

Using Lemma 3.2, we have

$$(\text{ran } \widehat{\mathscr{C}_k})^\perp = M_{\Theta_k} H_2(\mathbb{C}_+, \mathbb{C}^p),$$

and so

$$\begin{bmatrix} -M_{G_k}^* \\ I \end{bmatrix} (\text{ran } \widehat{\mathscr{C}_k})^\perp = \begin{bmatrix} -M_{F_k} \\ M_{\Theta_k} \end{bmatrix} H_2(\mathbb{C}_+, \mathbb{C}^p),$$

where Θ_k is inner, $G_k = \Theta_k F_k^*$, and $F_k \in H_\infty(\mathbb{C}_+, \mathbb{C}^{m \times p})$. Similarly

$$\begin{bmatrix} -M_G^* \\ I \end{bmatrix} (\text{ran } \widehat{\mathscr{C}})^\perp = \begin{bmatrix} -M_F \\ M_\Theta \end{bmatrix} H_2(\mathbb{C}_+, \mathbb{C}^p),$$

where Θ is inner, $G = \Theta F^*$, and $F \in H_\infty(\mathbb{C}_+, \mathbb{C}^{m \times p})$. Hence

$$g(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) = g\left(\begin{bmatrix} -M_{F_k} \\ M_{\Theta_k} \end{bmatrix} H_2(\mathbb{C}_+, \mathbb{C}^p), \begin{bmatrix} -M_F \\ M_\Theta \end{bmatrix} H_2(\mathbb{C}_+, \mathbb{C}^p) \right). \tag{8}$$

STEP 2. The projection Π_k onto $\begin{bmatrix} -M_{F_k} \\ M_{\Theta_k} \end{bmatrix} H_2(\mathbb{C}_+, \mathbb{C}^p)$ is given by

$$\begin{aligned} \Pi_k &= \begin{bmatrix} -M_{F_k} \\ M_{\Theta_k} \end{bmatrix} (M_{F_k}^* M_{F_k} + M_{\Theta_k}^* M_{\Theta_k})^{-1} \begin{bmatrix} -M_{F_k}^* & M_{\Theta_k}^* \end{bmatrix} \\ &= \begin{bmatrix} -M_{F_k} \\ M_{\Theta_k} \end{bmatrix} (I + M_{F_k}^* M_{F_k})^{-1} \begin{bmatrix} -M_{F_k}^* & M_{\Theta_k}^* \end{bmatrix}, \end{aligned} \tag{9}$$

since Θ_k is inner. Similarly the projection Π onto $\begin{bmatrix} -M_F \\ M_\Theta \end{bmatrix} H_2(\mathbb{C}_+, \mathbb{C}^p)$ is given by

$$\Pi = \begin{bmatrix} -M_F \\ M_\Theta \end{bmatrix} (I + M_F^* M_F)^{-1} \begin{bmatrix} -M_F^* & M_\Theta^* \end{bmatrix}. \tag{10}$$

In view of (8), the assumption that $\mathrm{g}(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) \to 0$ as $k \to \infty$ means simply that $\Pi_k \to \Pi$ in operator norm, from which we get, using (9) and (10), that

$$M_{F_k}(I + M_{F_k}^* M_{F_k})^{-1} M_{F_k}^* \longrightarrow M_F (I + M_F^* M_F)^{-1} M_F^* \tag{11}$$

$$M_{\Theta_k}(I + M_{F_k}^* M_{F_k})^{-1} M_{F_k}^* \longrightarrow M_\Theta (I + M_F^* M_F)^{-1} M_F^* \tag{12}$$

$$M_{\Theta_k}(I + M_{F_k}^* M_{F_k})^{-1} M_{\Theta_k}^* \longrightarrow M_\Theta (I + M_F^* M_F)^{-1} M_\Theta^* \tag{13}$$

in operator norm as $k \to \infty$.

STEP 3. If T is a bounded linear operator on a Hilbert space $\mathscr{H}$ such that $I+T^*T$ is invertible, then $I+TT^*$ is also invertible and

$$(I+TT^*)^{-1} = I - T(I+T^*T)^{-1}T^*. \tag{14}$$

This can be verified directly by checking that

$$(I-T(I+T^*T)^{-1}T^*)(I+TT^*) = I = (I+TT^*)(I-T(I+T^*T)^{-1}T^*).$$

From (14),

$$T(I+T^*T)^{-1}T^* = I-(I+TT^*)^{-1}. \tag{15}$$

Since $(I+T^*T)T^* = T^*(I+TT^*)$, by operating from the left and right by $(I+T^*T)^{-1}$ and $(I+TT^*)^{-1}$, respectively, we also obtain the identity

$$T^*(I+TT^*)^{-1} = (I+T^*T)^{-1}T^*. \tag{16}$$

Applying the identity (15) to (11), we obtain

$$I-(I+M_{F_k}M_{F_k}^*)^{-1} \longrightarrow I-(I+M_F M_F^*)^{-1},$$

and so

$$(I+M_{F_k}M_{F_k}^*)^{-1} \longrightarrow (I+M_F M_F^*)^{-1}$$

in operator norm as $k \to \infty$. Since the inverse map $\cdot^{-1}$ is continuous on the Banach space of continuous linear operators on a Hilbert space, it follows that

$$I+M_{F_k}M_{F_k}^* \longrightarrow I+M_F M_F^* \tag{17}$$

in operator norm as $k \to \infty$.

Applying the identity (16) to (12), we obtain

$$M_{\Theta_k}M_{F_k}^*(I+M_{F_k}M_{F_k}^*)^{-1} \longrightarrow M_\Theta M_F^*(I+M_F M_F^*)^{-1} \tag{18}$$

in operator norm as $k \to \infty$.

Finally, multiplying the sequence (18) by $(I+M_{F_k}M_{F_k}^*)$ and using that (17) and together with the fact that operator multiplication is continuous in the uniform topology, it follows that

$$M_{G_k} = M_{\Theta_k}M_{F_k}^* \longrightarrow M_G = M_\Theta M_F^*$$

in operator norm as $k \to \infty$. Thus we obtain (7). ■

Corollary 3.4 *The topology induced by the behavioral distance d is finer than that induced by the gap metric δ on the set* **S**.

4 The topology induced by the behavioral distance is coarser than that induced by the gap metric on the set S.

In this section we show that the topology induced by the behavioral distance is also coarser than that induced by the gap metric on the set **S**, by showing that if the transfer functions converge in the H_∞ norm, then the extended graphs converge in the gap topology of subspaces.

We show that under some conditions on the chosen realizations,

$$G_k \overset{H_\infty(\mathbb{C}_+,\mathbb{C}^{p\times m})}{\longrightarrow} G \text{ implies that } C_k \longrightarrow C \tag{19}$$

in Theorem 4.4 below, which (together with some other properties of the chosen realizations) will enable us to prove the convergence of the Θ_k and F_k constructed in Lemma 3.2. This then yields convergence of the extended graphs in Theorem 4.5.

The simple example with

$$\begin{aligned} G_k(s) &= \frac{1}{s+1}, \qquad A_k = -1,\ B_k = \frac{1}{k},\ C_k = k,\ D_k = 0, \text{ and} \\ G(s) &= \frac{1}{s+1}, \qquad A = -1,\ B = 1,\ C = 1,\ D = 0 \end{aligned}$$

demonstrates that (19) does not hold with every realization of the transfer function. So one looks for an appropriate realization for which the implication in (19) holds. We do this by appealing to Theorem 1.3 (page 303, Staffans [11]), where it is shown that every factorization of the Hankel operator induces a realization of the transfer function. For our purposes, we will use the following extreme factorization: $\Gamma = \Gamma I_{L_2}$. Recall that if $h \in L_1((0,\infty),\mathbb{C}^{p\times m})$ denotes the inverse Laplace transform of a transfer function in S, then the associated *Hankel operator* $\Gamma \in \mathscr{L}(L_2((0,\infty),\mathbb{C}^m), L_2((0,\infty),\mathbb{C}^p))$ is defined by

$$(\Gamma u)(t) = \int_0^\infty h(t+\tau)u(\tau)d\tau, \ \ t \geq 0, \text{ for } u \in L_2((0,\infty),\mathbb{C}^m).$$

Let $\mathscr{X} = \operatorname{ran}(\Gamma^*\Gamma)^{\frac{1}{2}} \subset L_2((0,\infty),\mathbb{C}^m)$, and let $P_{\mathscr{X}} : L_2((0,\infty),\mathbb{C}^m) \to L_2((0,\infty),\mathbb{C}^m)$ denote the projection operator onto the closed subspace $\mathscr{X}$. (Note that $\mathscr{X}$ is finite dimensional.)

Lemma 4.1 *Let $(\Sigma_k)_{k\geq 1}$ be a sequence of systems in* **S** *and let $\Sigma \in$* **S***. For each $k \geq 1$, let G_k denote the transfer function of Σ_k and let let G denote the transfer function of Σ. Furthermore, let Γ_k and Γ be the Hankel operators associated with the inverse Laplace transforms h_k and h of G_k and G respectively, with $\mathscr{X}_k = \operatorname{ran}(\Gamma_k^*\Gamma_k)^{\frac{1}{2}}$ and $\mathscr{X} = \operatorname{ran}(\Gamma^*\Gamma)^{\frac{1}{2}}$.*

If $G_k \to G$ in $H_\infty(\mathbb{C}_+,\mathbb{C}^{p\times m})$ as $k \to \infty$, then

1. *$\Gamma_k \to \Gamma$ in the operator norm as $k \to \infty$.*
2. *$P_{\mathscr{X}_k} \to P_{\mathscr{X}}$ in the operator norm as $k \to \infty$.*

Proof The first part follows for instance from Lemma 8.2.3.c (page 397, Curtain and Zwart [3]) combined with Lemma 8.1.2.a (page 388 of [3]):

$$\|\Gamma_k - \Gamma\| \leq \|G_k - G\|_{H_\infty(\mathbb{C}_+,\mathbb{C}^{p\times m})} \to 0 \text{ as } k \to \infty.$$

That $P_{\mathscr{X}_k} \to P_{\mathscr{X}}$ can be seen as follows. Let $\sigma_1^{(k)} \geq \ldots \geq \sigma_n^{(k)} > 0$ denote the n Hankel singular values of Γ_k, and $\sigma_1 \geq \ldots \geq \sigma_n > 0$ those of Γ. From the convergence of Γ_k to Γ in the operator norm, and the upper semicontinuity of the spectrum in the operator norm (see Theorem 3.1 on page 208 of Kato [6]), there exists an open interval (a,b) with $0 < a,b < +\infty$, such that for n sufficiently large, $\sigma_1^{(k)},\ldots,\sigma_n^{(k)} \in (a,b)$ and $\sigma_1,\ldots,\sigma_n \in (a,b)$. Let C be a simple, closed, rectifiable curve that encloses an open set containing (a,b) in its interior. Then we have

$$P_{\mathscr{X}_k} = \frac{1}{2\pi i}\int_C (\lambda I - (\Gamma_k^*\Gamma_k)^{\frac{1}{2}})^{-1} d\lambda.$$

From the continuity of the resolvent and the square root it follows easily that $P_{\mathscr{X}_k} \to P$. ■

The closed subspace $\mathscr{X}$ induces a decomposition of $L_2((0,\infty),\mathbb{C}^m)$:

$$L_2((0,\infty),\mathbb{C}^m) = \mathscr{X}^\perp \oplus \mathscr{X}.$$

Lemma 4.2 *Let $(\Sigma_k)_{k\geq 1}$ be a sequence of systems in* **S** *and let $\Sigma \in$* **S**. *For each $k \geq 1$, let G_k denote the transfer function of Σ_k and let let G denote the transfer function of Σ. Furthermore, let Γ_k and Γ be the Hankel operators associated with the inverse Laplace transforms h_k and h of G_k and G respectively, with $\mathscr{X}_k = \mathrm{ran}(\Gamma_k^*\Gamma_k)^{\frac{1}{2}}$ and $\mathscr{X} = \mathrm{ran}(\Gamma^*\Gamma)^{\frac{1}{2}}$.*

If $G_k \to G$ in $H_\infty(\mathbb{C}_+,\mathbb{C}^{p\times m})$ as $k \to \infty$, then there exists a k_0 (large enough) such that for $k \geq k_0$,

$$\mathscr{X}_k = \{W_k x + Ux \mid x \in \mathbb{C}^n\} \tag{20}$$

for a unique bounded linear operator $W_k : \mathbb{C}^n \to \mathscr{X}^\perp (\subset L_2((0,\infty),\mathbb{C}^m))$, and a fixed unitary identification map $U : \mathbb{C}^n \to \mathscr{X} (\subset L_2((0,\infty),\mathbb{C}^m))$.

Proof We show that for large k, $\mathscr{X}_k$ satisfies

$$\mathscr{X}_k \dot{+} \mathscr{X}^\perp = L_2((0,\infty),\mathbb{C}^m). \tag{21}$$

We begin by showing that

$$\mathscr{X}_k \cap \mathscr{X}^\perp = \{0\}. \tag{22}$$

Suppose that $x \in L_2((0,\infty),\mathbb{C}^m)$ is in $\mathscr{X}_k \cap \mathscr{X}^\perp$ with $\|x\| = 1$. Then we have

$$\|P_{\mathscr{X}_k} - P_{\mathscr{X}}\| = \sup_{\{v \mid \|v\|=1\}} \|(P_{\mathscr{X}_k} - P_{\mathscr{X}})v\| \geq \sup_{\{x \mid \|x\|=1\ x\in\mathscr{X}^\perp\}} \|(P_{\mathscr{X}_k} - P_{\mathscr{X}})x\| = \|x\| \geq 1.$$

Since $P_{\mathscr{X}_k} \to P_{\mathscr{X}}$, it follows that there exists a k_0 such that $k \geq k_0$ implies that $\|P_{\mathscr{X}_k} - P_{\mathscr{X}}\| < 1$. So we have proved (22).

Next we show that

$$\mathscr{X}_k + \mathscr{X}^\perp = L_2((0,\infty),\mathbb{C}^m). \tag{23}$$

Suppose that there exists $x \in L_2((0,\infty),\mathbb{C}^m)$ with $\|x\| = 1$ and $x \perp (\mathscr{X}_k + \mathscr{X}^\perp)$. In particular, $x \perp \mathscr{X}^\perp$ so $x \in \mathscr{X}$. Then also $x \perp \mathscr{X}_k$. Thus

$$\|P_{\mathscr{X}_k} - P_{\mathscr{X}}\| \geq \|(P_{\mathscr{X}_k} - P_{\mathscr{X}})x\| = \|P_{\mathscr{X}}x\| = \|x\| = 1.$$

Hence for k large enough, no such x can exist, and we conclude that $\mathscr{X}_k + \mathscr{X}^\perp$ is dense in $L_2((0,\infty),\mathbb{C}^m)$. Since $\mathscr{X}^\perp$ has finite codimension, every superspace of $\mathscr{X}^\perp$ is closed. So $\mathscr{X}_k + \mathscr{X}^\perp$ is closed and consequently (23) holds. From (22) and (23), we obtain that (21) holds.

Now we show that any subspace $\mathscr{X}_k$ satisfying (21) is a graph space, that is, there exists a unique bounded linear operator $W_k \in \mathscr{L}(\mathbb{C}^n, L_2((0,\infty),\mathbb{C}^m))$ such that (20) holds. Note that from (21), in particular, given $x \in \mathscr{X}$, there exists $x_\perp \in \mathscr{X}^\perp$ such that

$$x + x_\perp \in \mathscr{X}_k.$$

This $x_\perp$ is unique. Indeed if $x'_\perp \in \mathscr{X}^\perp$ is also such that $x'_\perp + x \in \mathscr{X}_k$, then

$$(x_\perp + x) - (x'_\perp + x) = x_\perp - x'_\perp \in \mathscr{X}_k \cap \mathscr{X}^\perp = \{0\}$$

which implies that $x_\perp = x'_\perp$. Define $M_k : \mathscr{X} \to \mathscr{X}^\perp$ by $M_k x = x_\perp$. Then it can be checked that M_k is linear and that

$$\mathscr{X}_k = \{M_k x + x \mid x \in \mathscr{X}\}.$$

As a consequence of the closed graph theorem, M_k is bounded. Since $\mathscr{X}$ is a finite dimensional space with dimension n, it follows that there is an isomorphism $U : \mathbb{C}^n \to \mathscr{X}$. Then W_k defined by $M_k U$ satisfies (20). ■

Define

$$U_k = (W_k + U)(I + W_k^* W_k)^{-\frac{1}{2}} \in \mathscr{L}(\mathbb{C}^n, L_2((0,\infty),\mathbb{C}^m)). \tag{24}$$

As $W_k^* U = U^* W_k = 0$ (since U and W_k have orthogonal ranges) and $U^* U = I_{\mathbb{C}^n}$, we see that $U_k^* U_k = I_{\mathbb{C}^n}$. It is also easily checked that $P_{\mathscr{X}_k} = U_k U_k^*$.

Lemma 4.3 *Let $(\Sigma_k)_{k\geq 1}$ be a sequence of systems in* $\mathbf{S}$ *and let $\Sigma \in \mathbf{S}$. For each $k \geq 1$, let G_k denote the transfer function of Σ_k and let let G denote the transfer function of Σ. Let the associated operators W_k, U_k, U be given as in (20).*

If $G_k \to G$ in $H_\infty(\mathbb{C}_+, \mathbb{C}^{p\times m})$ as $k \to \infty$, then

1. *$W_k \to 0$ in the operator norm as $k \to \infty$.*
2. *$U_k \to U$ in the operator norm as $k \to \infty$.*

Proof Since

$$P_{\mathscr{X}_k} = U_k U_k^* = (W_k + U)(I + W_k^* W_k)^{-1}(W_k^* + U^*) \longrightarrow P_{\mathscr{X}} = UU^*$$

it follows that

$$P_{\mathscr{X}^\perp} P_{\mathscr{X}_k} P_{\mathscr{X}^\perp} \longrightarrow P_{\mathscr{X}^\perp} U U^* P_{\mathscr{X}^\perp} = 0$$

as $k \to \infty$, where

$$P_{\mathscr{X}^\perp} P_{\mathscr{X}_k} P_{\mathscr{X}^\perp} = W_k (I + W_k^* W_k)^{-1} W_k^* = -I + (I + W_k W_k^*)^{-1}.$$

We conclude that $(I + W_k W_k^*) \to I$ and hence $W_k \to 0$. As $U_k = (W_k + U)(I + W_k^* W_k)^{-1/2}$, we see next that $U_k \to U$ in operator norm as $k \to \infty$. ■

Following Staffans [11] (Theorem 1.3 on page 302), by using the extremal factorization $\Gamma_k = \Gamma_k I_{L_2}$ for the Hankel associated with the inverse Laplace transform of G_k, it can be checked that G_k has a realization (A_k, B_k, C_k, D_k) with state space $\mathbb{C}^n$, input space $\mathbb{C}^m$ and output space $\mathbb{C}^p$ such that:

1. $A_k \in \mathbb{C}^{n \times n}$ is the infinitesimal generator of the semigroup $e^{tA_k} = U_k^* S(t) U_k$, for $t \geq 0$, where $S(\tau) : L_2((0,\infty), \mathbb{C}^m) \to L_2((0,\infty), \mathbb{C}^m)$ denotes the shift operator:

$$(S(\tau) f)(t) = f(t + \tau), \;\; t \in (0,\infty), \text{ for } f \in L_2((0,\infty), \mathbb{C}^m).$$

2. The input map $\mathscr{B}_k : L_2((0,\infty), \mathbb{C}^m) \to \mathbb{C}^n$ is given by $\mathscr{B}_k = U_k^*$.

3. The output map $\mathscr{C}_k : \mathbb{C}^n \to L_2((0,\infty), \mathbb{C}^p)$ is given by $\mathscr{C}_k = \Gamma_k U_k$.

4. The input-output map $\mathscr{D}_k : L_2((0,\infty), \mathbb{C}^m) \to L_2((0,\infty), \mathbb{C}^p)$ is given by

$$\mathscr{D}_k u = \mathscr{F}_p^{-1}(G_k(\mathscr{F}_m u)), \;\; u \in L_2((0,\infty), \mathbb{C}^m),$$

   where $\mathscr{F}_m : L_2((0,\infty), \mathbb{C}^m) \to H_2(\mathbb{C}_+, \mathbb{C}^m)$ and $\mathscr{F}_p^{-1} : H_2(\mathbb{C}_+, \mathbb{C}^m) \to L_2((0,\infty), \mathbb{C}^p)$ denote the Fourier transformation and the inverse Fourier transformation, respectively.

In light of these remarks, we have the following result:

Theorem 4.4 *Let $(\Sigma_k)_{k \geq 1}$ be a sequence of systems in $\mathbf{S}$ and let $\Sigma \in \mathbf{S}$. For each $k \geq 1$, let G_k denote the transfer function of Σ_k and let A_k, B_k, C_k, D_k be defined as in items 1, 2, 3 and 4 above. Furthermore, let G denote the transfer function of Σ and let A, B, C, D be defined as in items 1, 2, 3 and 4 above. If $G_k \to G$ in $H_\infty(\mathbb{C}_+, \mathbb{C}^{p \times m})$ as $k \to \infty$, then*

1. *$\Theta_k \to \Theta$ in $L_\infty(i\mathbb{R}, \mathbb{C}^{p \times p})$ as $k \to \infty$,*

2. *$F_k^* \to F^*$ in $L_\infty(\mathbb{C}_+, \mathbb{C}^{p \times m})$ as $k \to \infty$,*

where Θ_k, F_k and Θ, F are constructed as in the proof of Lemma 3.2.

Proof

STEP 1. In this step we use the Lebesgue dominated convergence theorem to prove that $A_k \to A$ in operator norm as $k \to \infty$.

By the previous Lemma, we know that $U_k \to U$ in operator norm as $k \to \infty$. Since $e^{tA_k} = U_k^* S(t) U_k$ and $e^{tA} = U^* S(t) U$, we conclude that for each fixed t, $e^{tA_k} \to e^{tA}$, and so we have pointwise convergence on $(0,\infty)$.

As U_k converges to U in operator norm, it is uniformly bounded: there exists a $M > 0$ such that $\|U_k\| = \|U_k^*\| \le M$ for all k. But $(S(t))_{t\ge 0}$ is a contraction semigroup and so $\|e^{tA_k}\| \le M^2$. Thus the semigroups are uniformly bounded with a uniform bound M^2, and so we have a dominating function $M^2 e^{-\mathrm{Re}(\omega)t}$ for each $\omega \in \mathbb{C}_+$: $\|e^{-\omega t} e^{tA_k}\| \le M^2 e^{-\mathrm{Re}(\omega)t} \in L_1(0,\infty)$.

Using the fact that the resolvent of the infinitesimal generator of a strongly continuous semigroup is the Laplace transform of the semigroup (see for instance Theorem 3.2.9.(i) on page 103 of Staffans [12]), and the Lebesgue dominated convergence theorem, we obtain

$$(\omega I - A_k)^{-1} = \int_0^\infty e^{-\omega t} U_k^* S(t) U_k dt \overset{k\to\infty}{\longrightarrow} \int_0^\infty e^{-\omega t} U^* S(t) U dt = (\omega I - A)^{-1}.$$

By the continuity of the inverse, we conclude that $A_k \to A$ in operator norm as $k \to \infty$.

STEP 2. We have that $\Gamma_k \to \Gamma$ and $U_k \to U$ as $k \to \infty$ in the respective operator norms. Since $\mathscr{C}_k = \Gamma_k U_k$, it is evident that $\mathscr{C}_k \to \mathscr{C}$ in $\mathscr{L}(\mathbb{C}^n, L_2((0,\infty),\mathbb{C}^p))$ as $k \to \infty$. We claim that in fact

$$\mathscr{C}_k \to \mathscr{C} \text{ in } \mathscr{L}(\mathbb{C}^n, W^{1,2}((0,\infty),\mathbb{C}^p)) \text{ as } k \to \infty, \tag{25}$$

where $W^{1,2}((0,\infty),\mathbb{C}^p)$ denotes the Sobolev space:

$$W^{1,2}((0,\infty),\mathbb{C}^p) := \left\{ f \in L_2((0,\infty),\mathbb{C}^p) \;\middle|\; \frac{df}{dt} \in L_2((0,\infty),\mathbb{C}^p) \right\},$$

equipped with the norm

$$\|f\|_{W^{1,2}} = \left(\|f\|_{L_2}^2 + \left\| \frac{df}{dt} \right\|_{L_2}^2 \right)^{\frac{1}{2}}.$$

Indeed (25) amounts to showing that for each $x \in \mathbb{C}^n$,

$$\frac{d}{dt}\mathscr{C}_k x \to \frac{d}{dt}\mathscr{C} x \text{ in } L_2((0,\infty),\mathbb{C}^p) \text{ as } k \to \infty,$$

which is the same as $\mathscr{C}_k A_k x \to \mathscr{C} A x$ in $L_2((0,\infty),\mathbb{C}^p)$ as $k \to \infty$. As we know that $\mathscr{C}_k \to \mathscr{C}$ and $A_k \to A$ in the appropriate spaces, the claim (25) follows. Since point evaluation is continuous in the Sobolev norm, for each $x \in \mathbb{C}^n$ we have

$$C_k x = (\mathscr{C}_k x)(0) \overset{k\to\infty}{\longrightarrow} (\mathscr{C} x)(0) = Cx \text{ in } \mathbb{C}^p.$$

Thus $C_k \to C$ in matrix norm as $k \to \infty$.

STEP 3. The solution to the Lyapunov equation (5) is given by $H_k = \mathscr{C}_k^* \mathscr{C}_k$ and so we see that $H_k \to H$ in $\mathbb{C}^{n\times n}$ as $k \to \infty$. From the continuity of the inverse, it also follows that $H_k^{-1} \to H^{-1}$ in $\mathbb{C}^{n\times n}$ as $k \to \infty$.

STEP 4. We know that $A_k \to A$ in $\mathbb{C}^{n\times n}$ as $k \to \infty$, and so using the continuity of the spectral set (see for instance Theorem 10.20 on page 257 of Rudin [9]), we see that given $\varepsilon > 0$, there exists a large enough K such that $k \geq K$ implies that $\sigma(A_k) \subset \sigma(A) + B(0,\varepsilon)$. Here $B(0,\varepsilon)$ denotes the ball with center 0 and radius ε in $\mathbb{C}$, and for a square matrix M, $\sigma(M)$ is used to denote its set of eigenvalues. Since $\sigma(A) \subset \{s \in \mathbb{C} \mid \mathrm{Re}(s) < 0\}$, it follows that there exists a positive ε and a $K \in \mathbb{N}$ such that for all $k \geq K$,

$$\sigma(A) \cup \sigma(A_k) \subset \{s \in \mathbb{C} \mid \mathrm{Re}(s) < -\varepsilon\}.$$

Consequently, there exist positive constants M_1, M_2 such that

$$\|e^{tA_k} - e^{tA}\| \leq M_1 e^{-\varepsilon t} + M_2 e^{-\varepsilon t} = (M_1 + M_2) e^{-\varepsilon t}.$$

Hence from the Lebesgue dominated convergence theorem, we have

$$\|e^{\cdot A_k} - e^{\cdot A}\|_{L_1((0,\infty),\mathbb{C}^{n\times n})} \to 0$$

as $k \to \infty$. Using continuity of the Laplace transform (see for instance Property A.6.2.a on page 636 of Curtain and Zwart [3]), it follows that

$$\|(\cdot I - A_k)^{-1} - (\cdot I - A)^{-1}\|_{H_\infty(\mathbb{C}_+,\mathbb{C}^{n\times n})} \to 0$$

as $k \to \infty$. Hence

$$\Theta_k(\cdot) = I - C_k(\cdot I - A_k)^{-1} H_k^{-1} C_k^* \longrightarrow \Theta(\cdot) = I - C(\cdot I - A)^{-1} H^{-1} C^*$$

in $L_\infty(i\mathbb{R}, \mathbb{C}^{p\times p})$ as $k \to \infty$. Since $L_\infty(i\mathbb{R}, \mathbb{C}^{p\times p})$ is a Banach algebra, from the continuity of the inverse, we have $\Theta_k^{-1} \to \Theta^{-1}$ in $L_\infty(i\mathbb{R}, \mathbb{C}^{p\times p})$ as $k \to \infty$. Finally

$$F_k^* = \Theta_k^{-1} G_k \longrightarrow \Theta^{-1} G = F^*$$

in $L_\infty(i\mathbb{R}, \mathbb{C}^{p\times m})$ as $k \to \infty$. ∎

Using the above result, we now obtain the following:

Theorem 4.5 *If $(\Sigma_k)_{k\geq 1}$ is a sequence in* **S**, *$\Sigma \in$* **S** *and there holds that*

$$\|G_k - G\|_{H_\infty(\mathbb{C}_+,\mathbb{C}^{p\times m})} \to 0 \textit{ as } k \to \infty,$$

then $\mathrm{g}(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) \to 0$ *as* $k \to \infty$.

Proof We use the formula established in STEP 1 of the proof of Theorem 3.3. Indeed from Theorem 4.4 above, we know that $M_{\Theta_k} \to M_\Theta$ and $M_{F_k} \to M_F$ in operator norm as $k \to \infty$. Consequently from (9) and (10), we see that $\Pi_k \to \Pi$ in operator norm as $k \to \infty$, and so $\mathrm{g}(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) \to 0$ as $k \to \infty$. ∎

Corollary 4.6 *The topology induced by the behavioral distance d is coarser than that induced by the gap metric δ on the set* **S**.

5 The topologies induced by the behavioral distance and the gap metric coincide on the set S.

We summarize the results from §3 and §4 below:

Theorem 5.1 *Let* $(\Sigma_k)_{k\geq 1}$ *be a sequence in* $\mathbf{S}$, *and let* $\Sigma \in \mathbf{S}$. *The following are equivalent:*

1. $\mathrm{g}(\mathscr{G}_e(\Sigma_k), \mathscr{G}_e(\Sigma)) \to 0$ *as* $k \to \infty$
2. $\mathrm{g}(\mathscr{G}(\Sigma_k), \mathscr{G}(\Sigma)) \to 0$ *as* $k \to \infty$
3. $\|G_k - G\|_{H_\infty(\mathbb{C}_+, \mathbb{C}^{p\times m})} \to 0$ *as* $k \to \infty$.

Proof This follows from Theorems 2.2, 3.3 and 4.5. ∎

Corollary 5.2 *The topologies induced by the behavioral distance and the gap metric coincide on the set* $\mathbf{S}$.

Acknowledgements: A useful discussion with Professor Tryphon Georgiou (Department of Electrical and Computer Engineering, University of Minnesota) is gratefully acknowledged.

References

[1] T.Ya. Azizov and I.S. Iokhvidov. *Linear Operators in Spaces with an Indefinite Metric*. John-Wiley, 1989.

[2] J.A. Ball, I. Gohberg and L. Rodman, *Interpolation of Rational Matrix Functions*, OT45 Birkhäuser-Verlag, 1990.

[3] R.F. Curtain and H.J. Zwart. *An Introduction to Infinite-Dimensional Systems Theory*. Springer-Verlag, 1995.

[4] T.T. Georgiou and M.C. Smith. Optimal robustness in the gap metric. *IEEE Transactions on Automatic Control*, AC-35(6):673–686, 1990.

[5] T.T. Georgiou and M.C. Smith. Topological approaches to robustness. In *Proceedings of the 10th International Conference on Analysis and Optimization of Systems*, June 9-12, École des Mines, Sophia-Antipolis, 1992.

[6] T. Kato. *Perturbation Theory for Linear Operators*. Springer, 1995.

[7] J.W. Polderman and J.C. Willems. *Introduction to Mathematical Systems Theory*. Springer, 1998.

[8] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 3rd edition, 1987.

[9] W. Rudin. *Functional Analysis*. McGraw-Hill, 2nd edition, 1991.

[10] A.J. Sasane. Distance between behaviours. *International Journal of Control*, 76:1214–1223, 2003.

[11] O.J. Staffans. Admissible factorizations of Hankel operators induce well-posed linear systems. *Systems and Control Letters*, 37:301–307, 1999.

[12] O.J. Staffans. *Well-Posed Linear Systems*. To be published by Cambridge University Press in 2004. Electronic copy available at `http://www.abo.fi/~staffans/publ.htm`.

[13] G. Zames and A.K. El-Sakkary. Unstable systems and feedback: the gap metric. In *Proceedings of the Allerton Conference*, 380–385, 1980.