Sub-optimal Hankel norm approximation for the Wiener class

Orest V. Iftime and Amol J. Sasane
Faculty of Applied Mathematics,
University of Twente,
P.O. Box 217, 7500 AE Enschede,
The Netherlands.

Abstract

For the Wiener class of matrix-valued functions we obtain a simple frequency domain solution for the sub-optimal Hankel norm approximation problem. The approach is via J-spectral factorization.

1 Introduction

Let G be a transfer function bounded on the imaginary axis and assume that its corresponding Hankel operator is compact. Let σ_k's denote the Hankel singular values of G, and assume that σ satisfies $\sigma_{l+1} < \sigma < \sigma_l$. Roughly speaking, the sub-optimal Hankel norm approximation problem is the following: Find a matrix-valued function K with at most l poles in the closed right half-plane (none of them on the imaginary axis) such that

$$\|G + K\|_\infty \leq \sigma,$$

where $\|\cdot\|_\infty$ denotes the L_∞-norm. The sub-optimal Hankel norm approximation problem has been studied extensively in the literature (see for example, Adamjan et al. [1], Ball and Ran [4], Glover [14], Ran [19], Glover et al. [15], Curtain and Ran [7], Sasane [21], Sasane and Curtain [10]). The new contribution of this paper is to present an elementary derivation of the reduction of the sub-optimal Hankel norm approximation problem to a J-spectral factorization problem. We do this for the Wiener class of transfer functions. Moreover an explicit parameterization of all solutions to the sub-optimal Hankel norm approximation problem is provided.

Although not stated explicitly in their paper, we believe that the paper by Ball and Helton [3] is the first paper which shows the connection between the sub-optimal Hankel norm approximation problem and a J-spectral factorization problem. Various corollaries of this abstract paper have been derived in Ball and Ran [4] and Curtain and Ran [7], but there is a gap between the abstract theory in [3] and the elementary looking corollaries. This motivated the search for an elementary self-contained proof in many papers (see Curtain and Ichikawa [5], Curtain and Oostveen [6], Curtain and Zwart [12], Sasane and Curtain [9], [10], [11] and Iftime and Zwart [17]).

The results presented in this paper refines the preceding lemmas in Curtain and Ichikawa [5], Curtain and Zwart [12], Sasane and Curtain [9], [11] for the Wiener class of transfer functions. All the proofs are based on frequency domain techniques.

2 Notation and preliminaries

In this section we quote some general results and introduce our notation. We begin by defining our class of stable transfer functions (the causal Wiener class) via their impulse responses. We say that $f \in \mathcal{A}$ if f has the representation

$$f(t) = \begin{cases} f_a(t) + f_0\delta(t), & t \geq 0, \\ 0, & t < 0, \end{cases}$$

where $f_0 \in \mathbb{C}$ (the set of complex numbers), $\int_0^\infty |f_a(t)|dt < \infty$, and δ represents the delta distribution at zero. For any $f \in \mathcal{A}$ we define $\hat{f}$, the Laplace transform of f,

$$\hat{f}(s) = \int_0^\infty e^{-st} f_a(t)dt + f_0, \tag{2.1}$$

for $s \in \overline{\mathbb{C}_+}$, where $\overline{\mathbb{C}_+} := \{s \in \mathbb{C} \mid \mathrm{Re}(s) \geq 0\}$. We define the *causal Wiener class* $\hat{\mathcal{A}}$ as

$$\hat{\mathcal{A}} := \left\{ \hat{f} \mid f \in \mathcal{A} \right\}.$$

¿From the definition of $\mathcal{A}$ it is easy to see that for every $f \in \mathcal{A}$, $\hat{f}$ is well-defined on $\overline{\mathbb{C}_+}$, it is bounded and analytic on $\mathbb{C}_+ := \{s \in \mathbb{C} \mid \mathrm{Re}(s) > 0\}$, continuous on $\overline{\mathbb{C}_+}$, and it has a well-defined limit at infinity, that is,

$$\sup_{s \in \in \overline{\mathbb{C}_+},\, |s| \geq \rho} \left| \hat{f}(s) - f_0 \right| \to 0 \quad \text{as } \rho \to \infty.$$

Furthermore, $\hat{\mathcal{A}}$ is a commutative Banach algebra with identity under pointwise addition and multiplication (see [13], Corollary A.7.48). For a complex function f, we use the notation $f^\sim$ to mean the following:

$$f^\sim(s) = \overline{f(-\overline{s})}. \tag{2.2}$$

We consider the algebra

$$\hat{\mathcal{W}} = \left\{ g \in L_\infty(i\mathbb{R}, \mathbb{C}) \mid g(i\cdot) = g_1(i\cdot) + g_2(i\cdot), \text{ with } g_1,\ g_2^\sim \in \hat{\mathcal{A}} \right\},$$

where

$$L_\infty(i\mathbb{R}, \mathbb{C}) = \left\{ f : i\mathbb{R} \to \mathbb{C} \mid \|f\|_{L_\infty} := \operatorname*{ess\,sup}_{s \in i\mathbb{R}} |f(s)| < \infty \right\}.$$

and we call it the *Wiener class* of transfer functions. $\hat{\mathcal{W}}$ is a Banach algebra under pointwise addition, multiplication, and scalar multiplication. The elements of $\hat{\mathcal{W}}$ are bounded and continuous on the imaginary axis, they have a limits at $\pm i\infty$, and these limits are equal.

By $\mathcal{R}_\infty$ we denote the class of proper, rational functions g with complex coefficients such that g has no poles in $\overline{\mathbb{C}_+}$, and has a nonzero limit at infinity. By $\hat{\mathcal{A}}_\infty$ we mean the set of all functions in $\hat{\mathcal{A}}$ that have all their zeros contained in the open right half-plane and a nonzero limit at infinity.

Now we introduce notation for some matrix-valued function spaces which will be used in the sequel.

1. By $\hat{\mathcal{A}}^{p \times m}$ we denote the set of complex $p \times m$ matrix-valued functions with entries in $\hat{\mathcal{A}}$.

2. By $\hat{\mathcal{A}}_l^{p\times m}$ we denote the set of complex $p\times m$ matrix-valued functions K of a complex variable with a decomposition
$$K = G + F,$$
where G is a rational matrix-valued transfer function of a system of MacMillan degree *at most equal* to l, with all its poles in the open right half-plane, and $F \in \hat{\mathcal{A}}^{p\times m}$.

3. $\hat{\mathcal{A}}_{[l]}^{p\times m}$ denotes the set of complex $p \times m$ matrix-valued functions K of a complex variable with a decomposition $K = G + F$, where G is a rational matrix-valued transfer function of a system of MacMillan degree *equal* to l, with all its l poles in the open right half-plane, and $F \in \hat{\mathcal{A}}^{p\times m}$.

4. We use the notation $\hat{\mathcal{W}}^{p\times m}$ for the class of $p \times m$ matrix-valued functions with entries in $\hat{\mathcal{W}}$.

We omit the size when there is no danger of confusion. We replace the indices by dots when we leave them unspecified. For complex matrix-valued functions we define
$$G^{\sim}(s) := [G(-\overline{s})]^*,$$
where $\cdot^*$ is used to denote the transpose complex conjugate of a matrix. For scalar functions this corresponds to (2.2). It can be seen that $G^{\sim}(s) = [G(-\overline{s})]^* = G(s)^*$ for all $s \in i\mathbb{R}$.

We will be using the following properties of the above classes of functions. These properties can be proved in a manner analogous to the ones in Section 2.6 of Sasane [20].

P1. If $f \in \hat{\mathcal{A}}$ and $g \in \hat{\mathcal{A}}_\infty$ such that g has at most l zeros (all in the open right half-plane), then $\frac{f}{g} \in \hat{\mathcal{A}}_l$.

P2. If $F \in \hat{\mathcal{A}}^{k\times k}$, $F(i\omega)$ is invertible for every $\omega \in \mathbb{R}$, $\lim_{s\to\pm i\infty} F(s)$ $(= F_\infty)$ is invertible, then $F(\cdot)^{-1} \in \hat{\mathcal{A}}_\bullet^{k\times k}$.

P3. If $K \in \hat{\mathcal{A}}_l^{p\times m}$, then there exists a right coprime factorization of K over $\hat{\mathcal{A}}$, $K = NM^{-1}$, (that is, there exist X and Y in $\hat{\mathcal{A}}^{\bullet\times\bullet}$ such that the following *Bezout identity* holds:
$$XM - YN = I$$
for all $s \in \overline{\mathbb{C}_+}$) where M is rational, $\det(M) \in \mathcal{R}_\infty$ has at most l zeros in $\overline{\mathbb{C}_+}$ and they are all contained in $\mathbb{C}_+$.

P4. If $K \in \hat{\mathcal{A}}_\bullet^{p\times m}$, then given any $\varepsilon > 0$, there exists a $\delta > 0$ such that whenever $0 \leq \zeta \leq \delta$, we have
$$\|K(\zeta + i\cdot)\|_\infty \leq \|K(i\cdot)\|_\infty + \varepsilon,$$
where $\|\cdot\|_\infty$ denotes the L_∞-norm.

P5. If $K \in \hat{\mathcal{A}}_l^{p\times m}$, $K_1 \in \hat{\mathcal{A}}^{p_*\times p}$, and $K_2 \in \hat{\mathcal{A}}^{m\times m_*}$, then $K_1KK_2 \in \hat{\mathcal{A}}_l^{p_*\times m_*}$.

In order to define the Hankel operator, we need the following notations:

$$\begin{aligned}
L_2^n &= \left\{ f : i\mathbb{R} \to \mathbb{C}^n \mid \|f\|_{L_2}^2 = \int_{-\infty}^{+\infty} |f(i\omega)|^2 d\omega < \infty \right\}, \\
H_2^n &= \left\{ f : \mathbb{C}_+ \to \mathbb{C}^n \mid f \quad \text{is analytic in} \quad \mathbb{C}_+ \ \text{ and } \ \|f\|_{H_2}^2 = \sup_{r>0} \int_{-\infty}^{+\infty} \|f(\zeta + j\omega)\|^2 d\omega < \infty \right\}, \\
H_2^{n,\perp} &= \left\{ f : \mathbb{C}_- \to \mathbb{C}^n \mid f \quad \text{is analytic in} \quad \mathbb{C}_0^- \ \text{ and } \ \|f\|_{H_2^\perp}^2 = \sup_{r<0} \int_{-\infty}^{+\infty} \|f(\zeta + j\omega)\|^2 d\omega < \infty \right\},
\end{aligned}$$

where $\mathbb{C}_- := \{s \in \mathbb{C} \mid \mathrm{Re}(s) < 0\}$. It is well known that L_2^n is the direct sum of H_2^n and $H_2^{n,\perp}$ with respect to the usual inner product. The Hankel operator with symbol $G \in L_\infty(i\mathbb{R}, \mathbb{C}^{p\times m})$, is defined as

$$H_G : H_2^m \to H_2^{p,\perp}, \ \ H_G u = \Pi_- G u \quad \text{for all} \ \ u \in H_2^m.$$

where Π_- is the orthogonal projection from L_2^p to H_2^p. Its adjoint is

$$H_G^* : H_2^{p,\perp} \to H_2^m, \ \ H_G^* y = \Pi_+ G^\sim y \quad \text{for all} \quad y \in H_2^{p,\perp},$$

where Π_+ is the orthogonal projection from L_2^m to $H_2^{m,\perp}$. If the Hankel operator with symbol $G \in L_\infty(i\mathbb{R}, \mathbb{C}^{p\times m})$ is compact, then we denote the singular values of H_G (that is, the nonnegative square roots of the eigenvalues of $H_G^* H_G$), by $\sigma_1 \geq \sigma_2 \geq \ldots (\geq 0)$. The σ_k's are then referred to as the *Hankel singular values of* G. If $G(i\cdot)$ is continuous on the imaginary axis with equal limits at $\pm i\infty$, then from Hartman's theorem (see for example Corollary 4.10, page 46, Partington [18]), it follows that the Hankel operator with symbol G is compact.

Let $G^\sim \in \hat{\mathcal{A}}^{m\times p}$ be a given matrix-valued function and let σ be a real number such that $\sigma_{l+1} < \sigma$. Then, the sub-optimal Hankel norm approximation problem that we consider is the following: Find $K \in \hat{\mathcal{A}}_l^{p\times m}$ such that $\|G(i\cdot) + K(i\cdot)\|_\infty \leq \sigma$.

The following theorem is a consequence of a slightly more general result proved by Sasane and Curtain in [9]. They give sufficient conditions for the sub-optimal Hankel norm approximation problem to have a solution.

Theorem 2.1. *Suppose that the following assumptions hold:*

S1. The matrix-valued function $G \in \hat{\mathcal{W}}^{p\times m}$ *(let* σ_k*'s denote the Hankel singular values of* G*).*

S2. $\sigma_{l+1} < \sigma < \sigma_l$.

S3. There exists a $\Lambda \in \hat{\mathcal{A}}^{(p+m)\times(p+m)}$ *such that*

$$\begin{bmatrix} I_p & 0 \\ G(s)^* & I_m \end{bmatrix} \begin{bmatrix} I_p & 0 \\ 0 & -\sigma^2 I_m \end{bmatrix} \begin{bmatrix} I_p & G(s) \\ 0 & I_m \end{bmatrix} = \Lambda(s)^* \begin{bmatrix} I_p & 0 \\ 0 & -I_m \end{bmatrix} \Lambda(s)$$

for all $s \in i\mathbb{R}$.

S4 The matrix-valued function Λ *is invertible as an element of* $\hat{\mathcal{A}}^{(p+m)\times(p+m)}$*, that is, there exists a* $V \in \hat{\mathcal{A}}^{(p+m)\times(p+m)}$ *such that* $\Lambda(s)V(s) = I_{p+m}$ *for all* $s \in \overline{\mathbb{C}_+}$.

S5. $\lim_{\omega\to\pm\infty} \Lambda(i\omega) = \begin{bmatrix} I_p & 0 \\ 0 & \sigma I_m \end{bmatrix}$, *S6.* $\Lambda_{11}(\cdot)^{-1} \in \hat{\mathcal{A}}_l^{p\times p}$,

then $K \in \hat{\mathcal{A}}_l^{p\times m}$ *and* $\|G(i\cdot)+K(i\cdot)\|_\infty \le \sigma$ *iff* $K(\cdot) = R_1(\cdot)R_2(\cdot)^{-1}$, *where*

$$\begin{bmatrix} R_1(\cdot) \\ R_2(\cdot) \end{bmatrix} = \Lambda(\cdot)^{-1} \begin{bmatrix} Q(\cdot) \\ I_m \end{bmatrix}$$

for some $Q \in \hat{\mathcal{A}}^{p\times m}$ *satisfying* $\|Q(i\cdot)\|_\infty \le 1$.

Remark 2.1. *The conditions S3-S4 say that the matrix-valued function*

$$W(s) := \begin{bmatrix} I_p & 0 \\ G(s)^* & I_m \end{bmatrix} \begin{bmatrix} I_p & 0 \\ 0 & -\sigma^2 I_m \end{bmatrix} \begin{bmatrix} I_p & G(s) \\ 0 & I_m \end{bmatrix} \tag{2.3}$$

admits a J*-spectral factorization (see the exact definition in the following section).*

In this paper, our main result is the following:

Theorem 2.2. *Let* G *be such that* $G^\sim \in \hat{\mathcal{A}}^{m\times p}$ *and let* σ *be a strictly positive real number such that* $\sigma \neq \sigma_k$ *for all* $k \in \mathbb{N}$. *Then there exists a* $\Lambda \in \hat{\mathcal{A}}^{(p+m)\times(p+m)}$ *such that S3, S4 and S5 hold. Moreover, the following are equivalent:*

1. $\sigma_{l+1} < \sigma < \sigma_l$.
2. *There exists a* $K \in \hat{\mathcal{A}}_{[l]}^{p\times m}$ *such that* $\|G(i\cdot)+K(i\cdot)\|_\infty \le \sigma$.
3. *The matrix-valued function* $\Lambda \in \hat{\mathcal{A}}^{(p+m)\times(p+m)}$ *which satisfies S3-S5, satisfies also* $\Lambda_{11}(\cdot)^{-1} \in \hat{\mathcal{A}}_{[l]}^{p\times p}$.

Furthermore, all solutions to the sub-optimal Hankel norm approximation problem are given by

$$K(\cdot) = R_1(\cdot)R_2(\cdot)^{-1},$$

where

$$\begin{bmatrix} R_1(\cdot) \\ R_2(\cdot) \end{bmatrix} = \Lambda(\cdot)^{-1} \begin{bmatrix} Q(\cdot) \\ I_m \end{bmatrix}$$

for some $Q \in \hat{\mathcal{A}}^{p\times m}$ *satisfying* $\|Q(i\cdot)\|_\infty \le 1$.

Remark 2.2. *The above theorem generalizes the result obtained, for* $\sigma > \sigma_1 = \|H_G\|$, *by Iftime and Zwart in [17]. In this case, the sub-optimal Hankel norm approximation problem becomes the so called sub-optimal Nehari problem. The sub-optimal Nehari problem can also be seen as an application of the results obtained by Ball et al. in [2], using the band method approach.*

3 Existence of a J-spectral factorization

We consider the signature matrix

$$J_{\sigma,p,m} = \begin{bmatrix} I_p & 0 \\ 0 & -\sigma^2 I_m \end{bmatrix},$$

where p and m are in $\mathbb{N}$ and σ is a strictly positive real number. Sometimes we simply use J_σ for the above, and if σ is 1, we use $J_{p,m}$ or simple, J.

Definition 3.1. *Let $W = W^{\sim} \in \hat{\mathcal{W}}^{k\times k}$. We say that the matrix-valued function W has a* J-spectral factorization *if there exists an invertible $\Lambda \in \hat{\mathcal{A}}^{k\times k}$ such that $\Lambda(\cdot)^{-1} \in \hat{\mathcal{A}}^{k\times k}$, and the equality*

$$W(s) = \Lambda^{\sim}(s)J\Lambda(s)$$

is satisfied for all $s \in i\mathbb{R}$. Such a matrix-valued function Λ will be called a J-spectral factor *of W.*

We now introduce the concept of equalizing vectors.

Definition 3.2. *A vector u is an* equalizing vector *for the matrix-valued function $W \in \hat{\mathcal{W}}^{k_1\times k_2}$ if u is a nonzero element of $H_2^{k_2}$ and Wu is in $H_2^{k_1,\perp}$.*

The following theorem gives equivalent conditions for the existence of a J-spectral factorization for a matrix-function $W = W^{\sim} \in \hat{\mathcal{W}}^{k\times k}$. A proof can be found in [16].

Theorem 3.1. *Let $W = W^{\sim} \in \hat{\mathcal{W}}^{k\times k}$ be such that $\det W(s) \neq 0$, for all $s \in i\mathbb{R} \cup \{\pm i\infty\}$. Then the following statements are equivalent*

1. *The matrix-valued function W admits a J-spectral factorization;*
2. *The matrix-valued function W has no equalizing vectors.*

Theorem 3.2. *Let G be a matrix-valued function of a complex variable such that $G^{\sim} \in \hat{\mathcal{A}}^{m\times p}$ and σ a positive real number such that $\sigma \neq \sigma_k$ for all $k \in \mathbb{N}$. Then there exists a $(p+m)\times(p+m)$ matrix-valued function of a complex variable $\Lambda \in \hat{\mathcal{A}}$ such that W, defined by*

$$W(s) = \begin{bmatrix} I_p & G(s) \\ 0 & I_m \end{bmatrix}^{\sim} J_{\sigma,p,m} \begin{bmatrix} I_p & G(s) \\ 0 & I_m \end{bmatrix}, \tag{3.4}$$

has a $J_{p,m}$-spectral factorization

$$W(s) = \Lambda(s)^{\sim} J_{p,m}\Lambda(s). \tag{3.5}$$

Moreover, if G is strictly proper, then Λ can be chosen such that

$$\lim_{\omega\to\pm\infty} \Lambda(i\omega) = \begin{bmatrix} I_p & 0 \\ 0 & \sigma I_m \end{bmatrix}. \tag{3.6}$$

Proof. It is easy to see that $W(s) = W^{\sim}(s)$ and $\det(W(s)) \neq 0$ for all $s \in \mathbb{R} \cup \{\pm i\infty\}$. In order to prove that the matrix-valued function $W(s)$ has a J-spectral factorization, it is enough to show that $W(s)$ has no equalizing vectors (see Theorem 3.1).

Let u be an equalizing vector for the matrix-valued function W, that is,

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in H_2, \ , u \neq 0, \ Wu = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in H_2^{\perp}. \tag{3.7}$$

So we have that

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = Wu = \begin{bmatrix} I_p & 0 \\ G^{\sim} & I_m \end{bmatrix} \begin{bmatrix} I_p & 0 \\ 0 & -\sigma^2 I_m \end{bmatrix} \begin{bmatrix} I_p & G \\ 0 & I_m \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} I_p & G \\ G^{\sim} & G^{\sim}G - \sigma^2 I_m \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

which is equivalent to

$$v_1 = u_1 + Gu_2 = v_1 \quad \text{and} \quad v_2 = G^{\sim}u_1 + G^{\sim}Gu_2 - \sigma^2 u_2.$$

In the first equality we split Gu_2 using the projections Π_- and Π_+. We obtain that

$$u_1 + \Pi_+ Gu_2 = v_1 - \Pi_- Gu_2 \quad \text{and} \quad G^{\sim}(u_1 + Gu_2) - \sigma^2 u_2 = v_2.$$

From (3.7) and the definition of the projection operators we have that the left-hand side of the first equality $u_1 + \Pi_+ Gu_2 \in H_2$ and the right-hand side $v_1 - \Pi_- Gu_2 \in H_2^{\perp}$. This implies that

$$u_1 + \Pi_+ Gu_2 = 0 \quad \text{and} \quad v_1 - \Pi_- Gu_2 = 0. \tag{3.8}$$

Now we replace u_1 in the second equality and split the term $G^{\sim}\Pi_- Gu_2$ using the projections. We have that

$$\begin{aligned} & G^{\sim}\Pi_- Gu_2 - \sigma^2 u_2 = v_2 \\ \Leftrightarrow\ & \Pi_- G^{\sim}\Pi_- Gu_2 + \Pi_+ G^{\sim}\Pi_- Gu_2 - \sigma^2 u_2 = v_2 \\ \Leftrightarrow\ & \Pi_+ G^{\sim}\Pi_- Gu_2 - \sigma^2 u_2 = v_2 - \Pi_- G^{\sim}\Pi_- Gu_2. \end{aligned}$$

Using similar arguments as before we have that

$$\Pi_+ G^{\sim}\Pi_- Gu_2 - \sigma^2 u_2 = 0,$$

which is equivalent to $\left(H_G^* H_G - \sigma^2 I_m\right) u_2 = 0$. Since σ is not a singular value of the Hankel operator, we obtain that u_2 must be zero. From (3.8) we see that also u_1 must be zero, so $u = 0$. We conclude that the matrix-valued function W has no equalizing vectors, which implies that W has a J-spectral factorization (3.5).

If G is a strictly proper matrix-valued function we see that the limit of W at $\pm i\infty$ is the identity matrix. Consequently, it is easy to check that if there exists a J-spectral factor Λ_0 which has the limit say Λ_∞ at $\pm\infty$, then Λ defined by

$$\Lambda(s) = \begin{bmatrix} I_p & 0 \\ 0 & \sigma I_m \end{bmatrix} \Lambda_\infty^{-1}\, \Lambda_0(s)$$

is clearly a J-spectral factor with the limit $\begin{bmatrix} I_p & 0 \\ 0 & \sigma I_m \end{bmatrix}$ at $\pm i\infty$. □

4 Proof of Theorem 2.2

In this section we prove the main result of this paper. We consider a matrix-valued function of a complex variable G such that $G^{\sim} \in \hat{\mathcal{A}}^{m\times p}$. Let σ_k denote the Hankel singular values of G, and let σ be a positive real number such that $\sigma \neq \sigma_k$ for all $k \in \mathbb{N}$. Theorem 3.2 shows that conditions S3, S4 and S5 are satisfied. The equivalence between the first and the second items of Theorem 2.2.

1. $\sigma_{l+1} < \sigma < \sigma_l$.

2. There exists a $K \in \hat{\mathcal{A}}^{p\times m}_{[l]}$ such that $\|G(i\cdot) + K(i\cdot)\|_\infty \le \sigma$.

is a consequence of the fact that

$$\inf_{K\in\hat{\mathcal{A}}_l} \|G(i\cdot) + K(i\cdot)\|_\infty = \sigma_{l+1}. \tag{4.9}$$

This can be proved as in Sasane [22].

In the following two lemmas we prove the equivalence between the last two items of Theorem 2.2. We start with the implication "3. ⇒ 2.".

Lemma 4.1. *Let $\Lambda \in \hat{\mathcal{A}}^{(p+m)\times(p+m)}$ be a matrix-valued function which satisfies S3-S5, and $\Lambda_{11}(\cdot)^{-1} \in \hat{\mathcal{A}}^{p\times p}_{[l]}$. Then there exists $K_0 \in \hat{\mathcal{A}}^{p\times m}_{[l]}$ such that $\|G(i\cdot) + K_0(i\cdot)\|_\infty \le \sigma$.*

Proof. Define

$$K_0(s) := V_{12}(s)V_{22}(s)^{-1},$$

where V is the inverse of Λ. The rest of the proof follows as in Sasane [20] (Chapter 4, Theorem 4.2.5 and Corollary 4.2.6). □

The following lemma proves the implication "2. ⇒ 3.". The proof is the same as in Curtain and Sasane [8] and in Sasane [22], but here we consider a different transfer function algebra.

Lemma 4.2. *Suppose that there exists a $K_* \in \hat{\mathcal{A}}^{p\times m}_{[l]}$ such that $\|G(i\cdot) + K_*(i\cdot)\|_\infty \le \sigma$. Let $\Lambda \in \hat{\mathcal{A}}^{(p+m)\times(p+m)}$ be a matrix-valued function which satisfies S3-S5. Then $\Lambda_{11}(\cdot)^{-1} \in \hat{\mathcal{A}}^{p\times p}_{[l]}$.*

Proof. We will split the proof in 6 steps. In the first two steps we prove some properties of V, the inverse of Λ. In the third step we prove that $V_{22}(\cdot)^{-1} \in \hat{\mathcal{A}}^{m\times m}_{[l_*]}$ for some $l_* \in \mathbb{N}$. In the fourth step we define

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} := \Lambda \begin{bmatrix} N \\ M \end{bmatrix}$$

where $K = NM^{-1}$ is a right-coprime factorization of K over $\hat{\mathcal{A}}$, and prove that U_2 is invertible over the imaginary axis and $\|U_1U_2^{-1}\|_\infty \le 0$. Using the Nyquist index, in Step 5 we show that $l_* \le l$. Finally, we obtain in the last step that $\Lambda_{11}(\cdot)^{-1} \in \hat{\mathcal{A}}^{p\times p}_{[l]}$.

Step 1. From S5,

$$V(s) - \begin{bmatrix} I_p & 0 \\ 0 & \frac{1}{\sigma}I_m \end{bmatrix} = V(s)\left(\begin{bmatrix} I_p & 0 \\ 0 & \sigma I_m \end{bmatrix} - \Lambda(s)\right)\begin{bmatrix} I_p & 0 \\ 0 & \frac{1}{\sigma}I_m \end{bmatrix},$$

and the fact that $V(\cdot) \in \hat{\mathcal{A}}$, it follows that

$$\lim_{\substack{|s|\to\infty \\ s\in\overline{\mathbb{C}_+}}} V(s) = \begin{bmatrix} I_p & 0 \\ 0 & \frac{1}{\sigma}I_m \end{bmatrix}. \tag{4.11}$$

Step 2. The matrix-valued function Λ satisfies S3, and so, taking inverses, we obtain

$$V(i\omega)\begin{bmatrix} I_p & 0 \\ 0 & -I_m \end{bmatrix}V(i\omega)^* = \begin{bmatrix} I_p & G(i\omega) \\ 0 & -I_m \end{bmatrix}^{-1}\begin{bmatrix} I_p & 0 \\ 0 & -\frac{1}{\sigma^2}I_m \end{bmatrix}\begin{bmatrix} I_p & 0 \\ G(i\omega)^* & -I_m \end{bmatrix}^{-1}. \tag{4.12}$$

for all $\omega \in \mathbb{R}$. Considering the $(2,2)$-block of the above yields

$$V_{21}(i\omega)V_{21}(i\omega)^* - V_{22}(i\omega)V_{22}(i\omega)^* = -\frac{1}{\sigma^2}I_m, \text{ where } \omega \in \mathbb{R}. \tag{4.13}$$

Thus for $u \in \mathbb{C}^{\mathtt{m}}$ we have

$$\|V_{22}(i\omega)^*u\|^2 = \|V_{21}(i\omega)^*u\|^2 + \frac{1}{\sigma^2}\|u\|^2.$$

So, if $V_{22}(i\omega)^*u = 0$ for all $\omega \in \mathbb{R}$, then $u = 0$. Hence it follows that $V_{22}(i\omega)^*$ is invertible for all $\omega \in \mathbb{R}$, or equivalently, $V_{22}(i\omega)$ is invertible for all $\omega \in \mathbb{R}$.

From (4.13), we have $\left\|V_{22}(i\omega)^{-1}V_{21}(i\omega)u\right\|^2 - \|u\|^2 = -\frac{1}{\sigma^2}\left\|V_{22}(i\omega)^{-1}u\right\|^2$. Let $M > 0$ be such that $\|V_{22}(i\omega)\| \le M$ for all $\omega \in \mathbb{R}$. We obtain $\|u\|^2 \le \|V_{22}(i\omega)\|^2 \left\|V_{22}(i\omega)^{-1}u\right\|^2 \le M^2\left\|V_{22}(i\omega)^{-1}u\right\|^2$. Thus

$$\left\|V_{22}(i\omega)^{-1}V_{21}(i\omega)\right\|^2 \le 1 - \frac{1}{\sigma^2M^2} < 1 \text{ for all } \omega \in \mathbb{R},$$

and so we have $\left\|V_{22}(i\cdot)^{-1}V_{21}(i\cdot)\right\|_\infty < 1$.

Step 3. From (4.11), we know that

$$\lim_{\substack{|s|\to\infty \\ s\in\overline{\mathbb{C}_+}}} V_{22}(s) = \frac{1}{\sigma}I_m.$$

Thus applying property P2 to $V_{22}(\cdot)$, we obtain that $V_{22}(\cdot)^{-1} \in \hat{\mathcal{A}}_{[l_*]}^{m\times m})$ for some $l_* \in \mathbb{N}$.

Step 4. Let $K_*(\cdot) \in \hat{\mathcal{A}}_{[l]}^{p\times m}$ satisfy $\|G(i\cdot) + K_*(i\cdot)\|_\infty \le \sigma$ and suppose it has the coprime factorization $K_* = NM^{-1}$ over $\hat{\mathcal{A}}$, where N and M are in $\hat{\mathcal{A}}$, M is rational, and $\det(M) \in \mathcal{R}_\infty$ has l zeros in $\overline{\mathbb{C}_+}$ and none on the imaginary axis. Define

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} := \begin{bmatrix} \Lambda_{11}N + \Lambda_{12}M \\ \Lambda_{21}N + \Lambda_{22}M \end{bmatrix} = \Lambda \begin{bmatrix} N \\ M \end{bmatrix} = \Lambda \begin{bmatrix} K_* \\ I_m \end{bmatrix} M. \tag{4.14}$$

We prove that U_2 is invertible over the imaginary axis and $\|U_1U_2^{-1}\|_\infty < 1$. First we prove that $\ker(U_2(i\omega)) = 0$ for all $\omega \in \mathbb{R}$. From (4.14) we have that

$$\begin{bmatrix} U_1(i\omega) \\ U_2(i\omega) \end{bmatrix} = \Lambda(i\omega)\begin{bmatrix} I_{\mathtt{p}} & G(i\omega) \\ 0 & I_m \end{bmatrix}^{-1}\begin{bmatrix} G(i\omega) + K_*(i\omega) \\ I_m \end{bmatrix} M(i\omega),$$

for all $\omega \in \mathbb{R}$. Note that the following equality holds

$$U_1(i\omega)^*U_1(i\omega) - U_2(i\omega)^*U_2(i\omega) = \begin{bmatrix} U_1(i\omega) \\ U_2(i\omega) \end{bmatrix}^* \begin{bmatrix} I_p & 0 \\ 0 & -I_m \end{bmatrix}\begin{bmatrix} U_1(i\omega) \\ U_2(i\omega) \end{bmatrix},$$

for all $\omega \in \mathbb{R}$. Multiplying the equality (3.5) to the left and to the right with appropriate matrices, we have that

$$\begin{bmatrix} I_p & 0 \\ G(i\omega)^* & I_m \end{bmatrix}^{-1} \Lambda(i\omega)^* \begin{bmatrix} I_p & 0 \\ 0 & -I_m \end{bmatrix} \Lambda(i\omega) \begin{bmatrix} I_p & G(i\omega) \\ 0 & I_m \end{bmatrix}^{-1} = \begin{bmatrix} I_p & 0 \\ 0 & -\sigma^2 I_m \end{bmatrix}.$$

Thus

$$U_1^* U_1 - U_2^* U_2 = M^* \begin{bmatrix} G + K_* \\ I_m \end{bmatrix}^* \begin{bmatrix} I_p & 0 \\ 0 & -\sigma^2 I_m \end{bmatrix} \begin{bmatrix} G + K_* \\ I_m \end{bmatrix} M \tag{4.15}$$

on the imaginary axis. Hence for all $u \in \mathbb{C}^m$ and all $\omega \in \mathbb{R}$, we have from equation (4.15) that

$$\|U_1(i\omega)u\|^2 - \|U_2(i\omega)u\|^2 = \|(G(i\omega) + K_*(i\omega))M(i\omega)u\|^2 - \sigma^2\|M(i\omega)u\|^2 \leq 0. \tag{4.16}$$

Since $\|G(i\cdot) + K_*(i\cdot)\|_\infty \leq \sigma$, and $M(i\omega)$ is invertible on the imaginary axis, we conclude that U_1 and U_2 satisfy the following inequality:

$$\|U_1(i\omega)u\| \leq \|U_2(i\omega)u\|. \tag{4.17}$$

Multiplying to the left the equality (4.14) with V, the inverse of Λ, we obtain that

$$V \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} K_* \\ I_m \end{bmatrix} M, \tag{4.18}$$

and so

$$V_{21}U_1 + V_{22}U_2 = M. \tag{4.19}$$

We claim that $\ker(U_2(i\omega)) = \{0\}$ for all $\omega \in \mathbb{R}$. Suppose on the contrary that there exists $0 \neq u_0 \in \mathbb{C}^m$ and a $\omega_0 \in \mathbb{R}$ such that $U_2(i\omega_0)u_0 = 0$. Then from (4.17) and (4.19), we obtain $M(i\omega_0)u_0 = 0$, which implies that $u_0 = 0$, a contradiction.
From (4.16), we deduce that

$$\left\|U_1(i\omega)U_2(i\omega)^{-1}y\right\|^2 \leq \|y\|^2 \text{ for all } \omega \in \mathbb{R},$$

and so $U_1(i\cdot)U_2(i\cdot)^{-1} \in L_\infty(\mathbb{R}, \mathbb{C}^{p\times m})$ satisfies $\left\|U_1(i\cdot)U_2(i\cdot)^{-1}\right\|_\infty \leq 1$.

Step 5. We prove, using the Nyquist index, that $l_* < l$, where $l_* \in \mathbb{N}$ is the one from Step 3. Consider U_1 and U_2 as defined in (4.14). We know that Λ_{21} is strictly proper and both Λ_{22} and M are proper with invertible limits at infinity in $\overline{\mathbb{C}_+}$. Thus from (4.14) we see that

$$\lim_{\substack{|s|\to\infty \\ s\in\overline{\mathbb{C}_+}}} U_2(s) \text{ exists and is invertible.} \tag{4.20}$$

Thus it follows that $s \mapsto \det(U_2(s))$ has only finitely many zeros in $\overline{\mathbb{C}_+}$, and they are all contained in $\mathbb{C}_+$.

The zeros of $\det(V_{22})$, $\det(M)$ and $\det(U_2)$ are contained in some half-plane $\mathbb{C}_{\varepsilon,+}$, where $\varepsilon > 0$. Since $\left\|V_{22}(i\cdot)^{-1}V_{21}(i\cdot)\right\|_\infty < 1$, there exists a $0 < r < 1$ such that $\left\|V_{22}(i\cdot)^{-1}V_{21}(i\cdot)\right\|_\infty = 1 - r$. It follows from P4 that there exists a $\delta_1 > 0$ such that $\delta_1 < \varepsilon$ and for any ζ satisfying $0 < \zeta < \delta_1$, $\left\|V_{22}(\zeta + i\cdot)^{-1}V_{21}(\zeta + i\cdot)\right\|_\infty \leq 1 - \frac{r}{2}$. Similarly it follows from Lemma P4 that there exists a $\delta_2 > 0$ such that $\delta_2 < \varepsilon$ and for any ζ satisfying $0 < \zeta < \delta_2$,

$$\|U_1(\zeta + i\cdot)U_2(\zeta + i\cdot)\|_\infty \leq 1 + \frac{\frac{r}{4}}{1 - \frac{r}{4}} = \frac{1}{1 - \frac{r}{4}}.$$

Let $\delta := \min\{\delta_1, \delta_2\}$, and fix a ζ satisfying $0 < \zeta < \delta$. Define

$$\phi(\alpha, s) = \det\left(\alpha V_{21}(\zeta + s)U_1(\zeta + s) + V_{22}(\zeta + s)U_2(\zeta + s)\right),$$

where $\alpha \in [0, 1]$.

a. We know that

$$\begin{aligned} \phi(0, \cdot) &= \det\left(V_{22}(\zeta + \cdot)U_2(\zeta + \cdot)\right) \quad \text{and} \\ \phi(1, \cdot) &= \det\left(V_{21}(\zeta + \cdot)U_1(\zeta + \cdot) + V_{22}(\zeta + \cdot)U_2(\zeta + \cdot)\right) \end{aligned}$$

are meromorphic (in fact analytic!) in $\mathbb{C}_{-\zeta/2,+}$.

b. $\phi(0, \cdot)$ has a nonzero limit at infinity in $\overline{\mathbb{C}_+}$: $\det(V_{22})$ has a nonzero limit at infinity in $\overline{\mathbb{C}_+}$ and $\det(U_2)$ has a nonzero limit at infinity in $\overline{\mathbb{C}_+}$ (see (4.20)).

$\phi(1, \cdot)$ has a nonzero limit at infinity in $\overline{\mathbb{C}_+}$, since V_{21} is strictly proper, U_1 is proper in $\overline{\mathbb{C}_+}$, and the above.

c. $(\alpha, s) \mapsto \phi(\alpha, s) : [0, 1] \times i\mathbb{R} \to \mathbb{C}$ is a continuous function, and

$$\begin{aligned} \phi(0, i\omega) &= \det\left(V_{22}(\zeta + i\omega)U_2(\zeta + i\omega)\right) \\ &= \det(V_{22}(\zeta + i\omega)) \det\left(U_2(\zeta + i\omega)\right), \quad \text{and} \\ \phi(1, i\omega) &= \det\left(V_{21}(\zeta + i\omega)U_1(\zeta + i\omega) + V_{22}(\zeta + i\omega)U_2(\zeta + i\omega)\right). \end{aligned}$$

d. We have

$$\begin{aligned} \phi(\alpha, i\omega) &= \det(V_{22}(\zeta + i\omega)) \det\left(U_2(\zeta + i\omega)\right) \\ &\quad \det\left(I + \alpha V_{22}(\zeta + i\omega)^{-1}V_{21}(\zeta + i\omega)U_1(\zeta + i\omega)U_2(\zeta + i\omega)^{-1}\right) \\ &\neq 0, \end{aligned}$$

since

$$\begin{aligned} &\left\|\alpha V_{22}(\zeta + i\cdot)^{-1}V_{21}(\zeta + i\cdot)U_1(\zeta + i\cdot)U_2(\zeta + i\cdot)^{-1}\right\|_\infty \\ \leq\; & 1 \left\|V_{22}(\zeta + i\cdot)^{-1}V_{21}(\zeta + i\cdot)\right\|_\infty \left\|U_1(\zeta + i\cdot)U_2(\zeta + i\cdot)^{-1}\right\|_\infty \\ \leq\; & \left[1 - \frac{r}{2}\right] \frac{1}{1 - \frac{r}{4}} < 1, \end{aligned}$$

$\det(V_{22}(\zeta + i\omega)) \neq 0$ and $\det\left(U_2(\zeta + i\omega)\right) \neq 0$.

e. $\phi(\alpha, \infty) \neq 0$, since V_{21} is strictly proper, U_1 is proper in $\overline{\mathbb{C}_+}$, and $\det(V_{22})\det(U_2)$ has a nonzero limit at infinity in $\overline{\mathbb{C}_+}$.

Thus the assumptions in Lemma A.1.18 (Curtain and Zwart [13], page 570) are satisfied by ϕ, and hence it follows that the Nyquist indices of $\phi(0, \cdot)$ and $\phi(1, \cdot)$ are the same. Consequently, the number of zeros are the same (the number of poles is zero, as $\phi(0, \cdot)$, $\phi(1, \cdot)$ are analytic in $\mathbb{C}_{-\frac{\delta}{2},+}$) and so the sum of the number of zeros of $s \mapsto \det(V_{22}(\zeta + s))$ in $\overline{\mathbb{C}_0^+}$ plus the number of zeros of $s \mapsto \det\left(U_2(\zeta + s)\right)$ in $\overline{\mathbb{C}_+}$ equals the number of zeros of $s \mapsto \det\left(V_{21}(\zeta + s)U_1(\zeta + s) + V_{22}(\zeta + s)U_2(\zeta + s)\right)$ ($= \det(M(\zeta + s)$, using (4.19)) in $\overline{\mathbb{C}_+}$.

In particular, we obtain that the number of zeros of $s \mapsto \det\left(V_{22}(\zeta + s)\right)$ in $\overline{\mathbb{C}_+}$ is less than or equal to l. But since the choice of ζ can be made arbitrarily small, it follows that $s \mapsto \det(V_{22})$ has at most l zeros in $\overline{\mathbb{C}_+}$. Thus $V_{22}(\cdot) \in \hat{\mathcal{A}}_{[l_*]}^{m\times m}$ where $l_* \leq l$.

Step 6. Finally it can be checked easily that $\Lambda_{11}^{-1} = V_{11} - V_{12}V_{22}^{-1}V_{21}$ and $V_{22}^{-1} = \Lambda_{22} - \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12}$. It follows from P5 that $\Lambda_{11}(\cdot)^{-1} \in \hat{\mathcal{A}}_{l_*}^{p\times p}$. If $\Lambda_{11}(\cdot)^{-1} \in \hat{\mathcal{A}}_{[k]}$ with $k < l_*$, using once more P5 we obtain that $V_{22}(\cdot)^{-1} \in \hat{\mathcal{A}}_k$, which is a contradiction. Using (4.9) and Lemma 4.1, we obtain that $\Lambda_{11}(\cdot)^{-1} \in \hat{\mathcal{A}}_{[l]}^{p\times p}$. □

Remark 4.1. *Finally we remark that under the assumptions of Theorem 2.2, if $\sigma_l < \sigma < \sigma_{l+1}$, then all solutions to the sub-optimal Hankel norm approximation problem are given by*

$$K(\cdot) = R_1(\cdot)R_2(\cdot)^{-1},$$

where

$$\begin{bmatrix} R_1(\cdot) \\ R_2(\cdot) \end{bmatrix} = \Lambda(\cdot)^{-1} \begin{bmatrix} Q(\cdot) \\ I_m \end{bmatrix}$$

for some $Q \in \hat{\mathcal{A}}^{p\times m}$ satisfying $\|Q(i\cdot)\|_\infty \leq 1$. This follows as in the proof of Theorem 2.1.

References

[1] V.M. Adamjan, D.Z. Arov, and M.G. Krein. Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur-Takagi problem. *Math. Sb. (N.S.)*, 86:34–75, 1971.

[2] J.A. Ball, I. Gohberg, and M.A. Kaashoek. *Recent development in operator theory and its applications*. Number 87 in Operator theory, Advances and Applications. Birkhäuser, Basel, 1996.

[3] J.A. Ball and J.W. Helton. A Beurling-Lax theorem for the Lie group $U(m,n)$ which contains most classical interpolation theory. *Journal of Operator Theory*, 9:107–142, 1983.

[4] J.A. Ball and A.C.M. Ran. Optimal Hankel norm model reductions and Wiener-Hopf factorization I: The canonical case. *SIAM Journal on Control and Optimization*, 25(2):362–382, 1987.

[5] R.F. Curtain and A. Ichikawa. The Nehari problem for infinite-dimensional systems of parabolic type. *Integral Equations and Operator Theory*, 26:29–45, 1996.

[6] R.F. Curtain and J.C. Oostveen. The Nehari problem for nonexponentially stable systems. *Integral Equations and Operator Theory*, 31:307–320, 1998.

[7] R.F. Curtain and A. Ran. Explicit formulas for Hankel norm approximations of infinite-dimensional systems. *Integral Equations and Operator Theory*, 13:455–469, 1989.

[8] R.F. Curtain and A.J. Sasane. A frequency domain solution to the sub-optimal Hankel norm approximation problem. Submitted for publication in Integral Equations and Operator Theory.

[9] R.F. Curtain and A.J. Sasane. Sub-optimal Hankel norm approximation for the analytic class of infinite-dimensional systems. To appear in Integral Equations and Operator Theory.

[10] R.F. Curtain and A.J. Sasane. Sub-optimal Hankel norm approximation for the analytic class of infinite-dimensional systems (extended version). Technical report, IWI 2000-4-01, University of Groningen, 2000. Available at `http://www/math.rug.nl/~amols/index.html`.

[11] R.F. Curtain and A.J. Sasane. Sub-optimal Hankel norm approximation for the Pritchard-Salamon class of infinite-dimensional systems. *Integral Equations and Operator Theory*, 2001.

[12] R.F. Curtain and H.J. Zwart. The Nehari problem for the Pritchard-Salamon class of infinite-dimensional linear systems: a direct approach. *Integral Equations and Operator Theory*, 18:130–153, 1994.

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

[14] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L_∞ error bounds. *International Journal of Control*, 39:1115–1193, 1984.

[15] K. Glover, R.F. Curtain, and J.R. Partington. Realization and approximation of linear infinite-dimensional systems with error bounds. *SIAM Journal on Control and Optimization*, 26:863–898, 1988.

[16] O.V. Iftime and H.J. Zwart. J-spectral factorization and equalizing vectors. *Systems and Control Letters*, 43:321–327, 2001.

[17] O.V. Iftime and H.J. Zwart. Nehari problems and equalizing vectors for infinite-dimensional sytems. *System and Control Letters*, 45:217–225, 2002.

[18] J.R. Partington. *An Introduction to Hankel Operators.* Number London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 1988.

[19] A. Ran. Hankel norm approximation for infinite-dimensional systems and Wiener-Hopf factorization. In R.F. Curtain, editor, *Modelling Robustness and Sensitivity Reduction in Control Systems*, NATO ASI Series, pages 57–70. Springer-Verlag, 1986.

[20] A.J. Sasane. *Hankel norm approximation for infinite-dimensional systems.* PhD thesis, University of Groningen, September 2001.

[21] A.J. Sasane. Hankel norm approximation for nonexponentially stable infinite-dimensional systems. *Submitted to Integral Equations and Operator Theory*, 2001.

[22] A.J. Sasane. *Hankel norm approximation for infinite-dimensional systems.* Number 277 in Lecture notes in Control and Information Sciences. Springer-Verlag, 2002.