Approximation by analytic matrix functions. The four block problem

V. V. Peller and S. R. Treil*

1. Introduction

The problem of approximating a given scalar function φ on the unit circle $\mathbb{T}$ uniformly by functions analytic in the unit disk $\mathbb{D}$ has been attracting analysts for a long time (see [Kha], [RSh], [Ne], [AAK1-2], [CJ], [PKh]). It was shown in [Kha] that for a continuous function φ such a best approximation is unique while it is not unique in the general case. Later it turned out that this problem is closely related with Hankel operators. Namely, it was proved by Nehari [Ne] that

$$\operatorname{dist}_{L^\infty}(\varphi, H^\infty) = \|H_\varphi\|,$$

where $H_\varphi : \ H^2 \to H^2_- \stackrel{\text{def}}{=} L^2 \ominus H^2$ is the *Hankel operator* with symbol φ defined by

$$H_\varphi f = \mathbb{P}_- \varphi f, \quad f \in H^2,$$

(we denote by $\mathbb{P}_+$ and $\mathbb{P}_-$ the orthogonal projections onto H^2 and H^2_-). Presently the problem of approximating by analytic functions in L^∞ is called *Nehari's problem.*

We shall also need the notion of a Toeplitz operator. Given $\varphi \in L^\infty$ the *Toeplitz operator* $T_\varphi : \ H^2 \to H^2$ is defined by

$$T_\varphi f = \mathbb{P}_+ \varphi f, \quad f \in H^2.$$

Adamyan, Arov and Krein [AAK1-2] found many interesting connections between Hankel operators and Nehari's problem. In particular they found a more

*Both authors are partially supported by NSF grant DMS 9304011

general condition under which a best approximation is unique: if the essential norm $\|H_\varphi\|_{\mathrm{e}}$ is less than $\|H_\varphi\|$, then φ has a unique best approximation. (Recall that for an operator T on Hilbert space

$$\|T\|_{\mathrm{e}} \stackrel{\text{def}}{=} \inf\{\|T-K\| : \ K \text{ is compact}\}).$$

This sufficient uniqueness condition can easily be reformulated in terms of the function φ itself since

$$\|H_\varphi\|_{\mathrm{e}} = \operatorname{dist}_{L^\infty}(\varphi, H^\infty + C)$$

(see [AAK1-2]). They also found a criterion of uniqueness of a best approximation in terms of the corresponding Hankel operator, and in the case of non-uniqueness parametrized all best approximations (optimal solutions of Nehari's problem), see[AAK1-2]. However it is not very easy to verify whether a function φ in L^∞ satisfies the criterion.

Carleson and Jacobs [CJ] studied smoothness properties of the best approximation for smooth functions φ. They proved that if φ belongs to the Hölder–Zygmund class Λ_α, $\alpha > 0$, $\alpha \notin \mathbb{Z}$, then the best approximation also belongs to the same class.

Later in [PKh] more general hereditary properties of the non-linear operator of best approximation were studied. For a large class of function spaces X on $\mathbb{T}$ it was proved that if $\varphi \in X$ and f is the best approximation by analytic functions, then $f \in X$. Note also that in [PKh] Nehari's problem was also applied in prediction theory which led to a new approach to the problem of describing stationary processes satisfying various regularity conditions in terms of their spectral densities.

A new wave of interest in Nehari's problem was caused by the development of H^∞ control theory where Nehari's problem plays a central role (see [Fr]). Moreover for the needs of H^∞ control theory it is important to consider Nehari's problem for matrix-valued functions: given an $n \times m$ matrix function Φ on $\mathbb{T}$ the problem is to approximate Φ by bounded analytic matrix functions Q in the norm

$$\|\Phi - Q\|_\infty \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \|\Phi(\zeta) - F(\zeta)\|,$$

where $\|\cdot\|$ on the right-hand side is the norm of the matrix as an operator from $\mathbb{C}^m$ to $\mathbb{C}^n$. However in contrast with the scalar case we have uniqueness of a best approximation only in exceptional cases. Indeed, let

$$\Phi = \begin{pmatrix} \bar z & 0 \\ 0 & \frac{1}{2}\bar z \end{pmatrix}.$$

Clearly $\operatorname{dist}_{L^\infty}(\bar z, H^\infty) = 1$ and so $\|\Phi - Q\|_\infty \ge 1$ for any $Q \in H^\infty$. However it is easy to see that any function of the form $\begin{pmatrix} 0 & 0 \\ 0 & q \end{pmatrix}$ with $q \in H^\infty$, $\|q\|_\infty \le \frac{1}{2}$, is a best approximation. Intuitively, however, is clear that the "very best" approximation is the zero matrix function $\mathbb{O}$.

In [Y] Young suggested imposing the following additional assumptions on approximating functions. Let Ω_0 be the set of best approximations:

$$\Omega_0 = \{Q \in H^\infty : \ Q \text{ minimizes } \ \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} \|\Phi(\zeta) - Q(\zeta)\|\}.$$

Define inductively the sets Ω_j as follows

$$\Omega_j = \{Q \in \Omega_{j-1} : \ Q \text{ minimizes } \ \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} s_j(\Phi(\zeta) - Q(\zeta))\}$$

(for a matrix (or an operator) A the jth *singular value* $s_j(A)$, $j \ge 0$, is the distance from A to the set of matrices (operators) of rank at most j, $s_0(A) \stackrel{\text{def}}{=} \|A\|$). Elements of $\Omega_{\min\{m,n\}-1}$ are called *superoptimal approximations* of Φ (or superoptimal solutions of Nehari's problem). Put

$$t_j \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} s_j(\Phi(\zeta) - Q(\zeta)), \quad Q \in \Omega_j.$$

The numbers t_j, $0 \le j \le \min\{m,n\}-1$, are called the *superoptimal singular values*. As in the scalar case, $t_0 = \|H_\Phi\|$ (the Hankel operator $H_\Phi : H^2(\mathbb{C}^n) \to H^2(\mathbb{C}^m)$ is defined in the same way as in the scalar case).

Note that Q is a superoptimal solution of Nehari's problem if and only if it lexicographically minimizes the sequence $\{s_j^\infty(\Phi - Q)\}_{j \ge 0}$, where for a matrix function F on $\mathbb{T}$

$$s_j^\infty(F) \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} s_j(F(\zeta)).$$

It was proved in [PY1] that for $\Phi \in H^\infty + C$ there exists a unique superoptimal approximation Q. The method of the proof in [PY1] is based on certain special factorizations of matrix functions (thematic factorizations) and it is constructive. Later in [T] another method was suggested to establish uniqueness in the $H^\infty + C$ case which is based on weighted Nehari's problem.

However in the case when the Hankel operator H_Φ is non-compact, there was no analog of the Adamyan–Arov–Krein sufficient condition for uniqueness in the case of matrix functions.

Nehari's problem is a special case of the so-called *four block* problem which is one of the most important problems in control theory. Let Φ be a block matrix function of the form

$$\Phi = \left(\begin{array}{c|c} \Phi_{11} & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right).$$

Here Φ has size $m \times n$, Φ_{11} has size $m_1 \times n_1$, and Φ_2 has size $m_2 \times n_2$. The four block problem is to minimize

$$\left\| \left(\begin{array}{c|c} \Phi_{11} - Q & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array} \right) \right\|_\infty , \tag{1.1}$$

over bounded analytic functions Q of size $m_1 \times n_1$. A function $Q \in H^\infty(M_{m_1 n_1})$ is called an *optimal solution* of the four block problem if it minimizes the norm (1.1).

The four block problem arises naturally when one considers the following (model-matching) problem in H^∞ control. Let F, G_1 and G_2 be matrix functions of class H^∞. Consider the problem of finding

$$\inf_{Q \in H^\infty} \|F - G_1 Q G_2\|_\infty . \tag{1.2}$$

(the sizes of the matrix functions in (1.2) are such that (1.2) is meaningful). Many problems in H^∞ control reduce to the model-matching problem. In the case when G_1 and G_2 are square matrix functions the problem (1.2) reduces to Nehari's problem while in the general case it reduces to the four block problem (1.1).

By analogy with Hankel operators we define the four block operator $\Gamma_\Phi : H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2}) \to H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$ by

$$\Gamma_\Phi \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \mathbb{P}^- \Phi \begin{pmatrix} f_1 \\ f_2 \end{pmatrix},$$

where $\mathbb{P}^-$ is the orthogonal projection from $L^2(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$ onto $H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$. As in the case of Hankel operators the *infimum* in (1.1) is equal to $\|\Gamma_\Phi\|$ (see [FT]). The matrix function Φ is called a *symbol* of the four block operator Γ_Φ (a four block operator has many different symbols).

As in the case of Nehari's problem we can define the sets Ω_j:

$$\Omega_0 = \left\{ Q \in H^\infty : \ Q \text{ minimizes } \left\| \left(\begin{array}{c|c} \Phi_{11} - F & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array} \right) \right\|_\infty \right\},$$

$$\Omega_j = \left\{ Q \in \Omega_{j-1} : \ Q \text{ minimizes } \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} s_j \left(\left(\begin{array}{c|c} \Phi_{11} - F & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array} \right) (\zeta) \right) \right\}.$$

Q is called a *superoptimal solution* of the four block problem (1.1) if $Q \in \Omega_{\min\{m_1, n_1\} - 1}$. We define the superoptimal singular values of the four block problem (1.1) by

$$t_j = \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} s_j \left(\left(\begin{array}{c|c} \Phi_{11} - Q & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array} \right) (\zeta) \right), \qquad F \in \Omega_j.$$

Clearly, $t_0 = \|\Gamma_\Phi\|$.

We say that a Φ is a *superoptimal symbol* of a four block operator Γ if $\Gamma = \Gamma_\Phi$ and the zero function is a superoptimal solution of the corresponding four block problem.

Using a simple compactness argument one can prove easily that a superoptimal solution always exists. However we cannot expect a sufficient condition for uniqueness of the superoptimal approximation which would be similar to the one found in [PY1] in the case of Nehari's problem. Indeed it can easily be proved that a four block operator cannot be compact unless Φ_{12}, Φ_{21}, and Φ_{22} are identically equal to zero in which case the four block problem is equivalent to Nehari's problem.

The main result of the paper (Theorem 2.1) is a sufficient condition for the four block problem to have a unique superoptimal solution. This result can be considered as an analog of the Adamyan–Arov–Krein theorem mentioned above which deals with Nehari's problem in the scalar case. Note that Theorem 2.1 also gives us a new result for Nehari's problem in the case when the corresponding Hankel operator is non-compact.

The proof is constructive. We give an algorithm to find the unique superoptimal solution. The algorithm is similar to the one given in [PY1], it reduces the problem to the case of matrix functions of lower size. However the proof is considerably more complicated than in the case of Nehari's problem with compact Hankel operator.

In Section 3 we describe briefly the method of factorization and diagonalization. We construct important matrix functions V and W which can be considered as analogs of the thematic functions defined in [PY1].

In Section 4 we use the construction of Section 3 to parametrize all optimal solutions of the four block problem. This allows us to reduce the problem of finding superoptimal solutions to the case of matrix functions of lower size.

Section 5 is devoted to the proof of the fact that the matrix functions V_c and W_c which are submatrices of V and W are left invertible in H^∞. This is one of the principal points in the proof of the main result.

In Section 6 we prove another crucial fact for the proof of the main result. Namely, we show that if Φ satisfies the hypotheses of Theorem 2.1, then the lower order four block problem, obtained as a result of parametrization in Section 4, also satisfies the hypotheses of Theorem 2.1. This makes it possible to continue the process and complete the proof of the main result.

In Section 7 we study superoptimal symbols of four block operators. We obtain certain special factorizations of such symbols (thematic factorizations), and define the indices of such factorizations. In the case of Nehari's problem with compact Hankel operators such factorizations were found in [PY1].

To prove the invariance of indices we introduce in Section 8 the notion of a superoptimal weight for the four block operator. This is an analog of the notion introduced in [T] in the case of compact Hankel operators.

In Section [9] we use superoptimal weights to prove that the sums of the indices in a thematic factorization which correspond to equal superoptimal singular values do not depend on the choice of factorization. In the case of compact Hankel operators this invariance property was proved in [PY2]. Note that as in the case of Nehari's problem for $H^\infty + C$ functions if there are equal superoptimal singular values, the indices can depend on the choice of thematic factorization (see [PY2]).

The last section is devoted to inequalities between the superoptimal singular values and the singular values of the four block operator. We obtain an inequality which is new even in the case of Nehari's problem with compact Hankel operator. It is stronger than the one obtained in [PY2].

Note that in [PY1] hereditary properties of the non-linear operator of superoptimal approximation were studied. It was shown there that for a large class of function spaces X the inclusion $\Phi \in X$ implies that the superoptimal approximant to Φ also belongs to X. It would be interesting to find analogs of such results in the case of the four block problem. In particular we do not know whether the superoptimal solution of the four block problem (1.1) must belong to a Hölder class Λ_α if $\Phi \in \Lambda_\alpha$.

Throughout this paper we shall denote by $M_{m,n}$ the space of $m \times n$ matrices. We shall use the notation $L^\infty(M_{m,n})$ and $H^\infty(M_{m,n})$ for the spaces of bounded and bounded analytic functions which take values in $M_{m,n}$. Sometimes if it does not lead to a confusion, we shall simply write L^∞ and H^∞ instead of $L^\infty(M_{m,n})$ and $H^\infty(M_{m,n})$.

A matrix function $\Theta \in H^\infty$ is called *inner* if $\Theta(\zeta)$ is isometric for almost all $\zeta \in \mathbb{T}$. A matrix function $F \in H^\infty(M_{m,n})$ is called *outer* if $FH^2(\mathbb{C}^n)$ is dense in $H^2(\mathbb{C}^m)$. A function $F \in H^\infty(M_{m,n})$ is called *co-outer* if the transposed function $F^t \in H^\infty(M_{n,m})$ is outer.

2. The main result

In this section we state the main result of the paper as well as important corollaries. Let Φ be a matrix function of the form

$$\Phi = \left(\begin{array}{c|c} \Phi_{11} & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right), \tag{2.1}$$

where Φ has size $m \times n$, Φ_{11} has size $m_1 \times n_1$, and Φ_{22} has size $m_2 \times n_2$. Recall that Γ_Φ is the four block operator defined in Section 1 and $\{t_j\}$ is the sequence of superoptimal singular values.

The following theorem is the main result of the paper.

Theorem 2.1. *Let Φ be a bounded function of the form (2.1). Suppose that $\|\Gamma_\Phi\|_{\mathrm{e}}$ is less than the smallest nonzero superoptimal singular value. Then there exists a unique superoptimal solution Q of the four block problem for Φ. The singular values*

$$s_j\left(\left(\begin{array}{c|c}\Phi_{11}-Q & \Phi_{12}\\ \hline \Phi_{21} & \Phi_{22}\end{array}\right)(\zeta)\right),\qquad 1\le j\le d-1,$$

are constant on $\mathbb{T}$.

The following partial case of Theorem 2.1 improves the result of [PY1] on uniqueness of superoptimal solutions of Nehari's problem.

Theorem 2.2. *Let Φ be bounded matrix function on $\mathbb{T}$. Suppose that $\|H_\Phi\|_{\mathrm{e}}$ is less that the smallest nonzero superoptimal singular value of Nehari's problem. Then Φ has a unique superoptimal approximation Q by bounded analytic matrix functions. The singular values $s_j(\Phi(\zeta)-Q(\zeta))$ are constant on $\mathbb{T}$.*

3. Diagonalization

In this section we start with a maximizing vector of the four block operator and we construct a certain special unitary-valued matrix. This allows us to achieve a diagonalization. Later using this diagonalization we shall reduce the problem to the case of a matrix function of a lower size.

Lemma 3.1. *Let $\mathbf{v}$ be an $n\times 1$ inner matrix function. Then there exists a co-outer function $V_c\in H^\infty(M_{n,n-1})$ such that the matrix function*

$$\mathbf{V}\stackrel{\mathrm{def}}{=}\begin{pmatrix}\mathbf{v} & \overline{V}_c\end{pmatrix}$$

is unitary-valued on $\mathbb{T}$.

In [PY1] a stronger result was obtained. It was shown that all minors of $\mathbf{V}$ on the first column are in H^∞. This property of analyticity of minors was essential for the proof of the uniqeness of a superoptimal solution of Nehari's problem which was given in [PY1]. Earlier the existence of a co-outer V_c satisfying the requirement of Lemma 3.1 was proved in [Va], however the property of analyticity of minors was not noticed in [Va]. It also can be shown that if we $V_c^{(1)}$ and $V_c^{(2)}$ are $n\times(n-1)$ co-outer functions satisfying the requirements of Lemma 3.1, then there exists a constant unitary matrix U such that $V_c^{(1)} = V_c^{(2)}U$ (see [Va], [PY1]).

To start the procedure we need a maximizing vector of the four block operator Γ_Φ, i.e., a nonzero vector $f \in H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2})$ such that $\|\Gamma_\Phi f\| = \|\Gamma_\Phi\| \cdot \|f\|$. If Φ satisfies the hypotheses of Theorem 2.1 and $\Gamma_\Phi \neq \mathbb{O}$, then $\|\Gamma_\Phi\| = t_0 > \|\Gamma_\Phi\|_{\rm e}$ and so a maximizing vector for Γ_Φ exists.

The following fact is well-known in the case of Hankel operators (see [AAK3]). The proof of it in the case of four block operators is similar.

Lemma 3.2. *Let Φ be a matrix function on $\mathbb{T}$ of the form* (2.1) *and such that $\|\Phi\|_\infty = \|\Gamma_\Phi\|$. Suppose that f is a maximizing vector for Γ_Φ and put $g = t_0^{-1}\Gamma_\Phi f$. Then $\Gamma_\Phi f = \Phi f$ and $\|g(\zeta)\|_{\mathbb{C}^m} = \|f(\zeta)\|_{\mathbb{C}^n}$ a.e. on $\mathbb{T}$. Furthermore $\|\Phi(\zeta)\| = \|\Phi\|_\infty$ a.e. on $\mathbb{T}$.*

Proof. We have

$$\|\Gamma_\Phi f\|_2 = \|\mathbb{P}^-\Phi f\|_2 \le \|\Phi f\|_2 \le \|\Phi\|_\infty \|f\|_2 = \|\Gamma_\Phi\| \cdot \|f\|_2 = \|\Gamma_\Phi f\|_2.$$

It follows that all inequalities in this chain are, in fact, equalities. The fact that $\|\mathbb{P}^-\Phi\|_2 = \|\Phi\|_2$ certainly means that $\Phi f \in H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$ and so $\Gamma_\Phi f = \Phi f$. The equality $\|\Phi f\|_2 = \|\Phi\|_\infty \|f\|_2$ implies that $\|g(\zeta)\|_{\mathbb{C}^m} = \|f(\zeta)\|_{\mathbb{C}^n}$ a.e. on $\mathbb{T}$, which in turn implies that $\|\Phi(\zeta)\| = \|\Phi\|_\infty$ for almost all $\zeta \in \mathbb{T}$. ∎

Lemma 3.3. *Let Φ be a matrix function of the form* (2.1) *such that $\|\Gamma_\Phi\|_{\rm e} < \|\Gamma_\Phi\|$. Suppose that $f = f_1 \oplus f_2$ is a maximizing vector for Γ_Φ and $t_0 g = g_1 \oplus g_2 = \Gamma_\Phi f$, where $f_1 \in H^2(\mathbb{C}^{n_1})$, $f_2 \in L^2(\mathbb{C}^{n_2})$, $g_1 \in H^2_-(\mathbb{C}^{m_1})$, and $g_2 \in L^2(\mathbb{C}^{m_2})$. Then*

$$\|f_1(\zeta)\|^2_{\mathbb{C}^{n_1}} \ge \frac{\|\Gamma_\Phi\|^2 - \|\Gamma_\Phi\|^2_{\rm e}}{\|\Gamma_\Phi\|^2}\|f(\zeta)\|^2_{\mathbb{C}^n}, \quad \|g_1(\zeta)\|^2_{\mathbb{C}^{m_1}} \ge \frac{\|\Gamma_\Phi\|^2 - \|\Gamma_\Phi\|^2_{\rm e}}{\|\Gamma_\Phi\|^2}\|g(\zeta)\|^2_{\mathbb{C}^m}$$

almost everywhere on $\mathbb{T}$.

Proof. By subtracting an optimal solution, we can assume that $\|\Phi\|_\infty = \|\Gamma_\Phi\|$. By Lemma 3.2, $t_0 g(\zeta) = \Phi(\zeta) f(\zeta)$ and $\|g(\zeta)\|_{\mathbb{C}^m} = \|f(\zeta)\|_{\mathbb{C}^n}$ a.e. on $\mathbb{T}$. Therefore

$f(\zeta)$ is a maximizing vector for $\Phi(\zeta)$ and $g(\zeta)$ is a maximizing vector for $\Phi^*(\zeta)$. We have $g_2 = \begin{pmatrix} \Phi_{21} & \Phi_{22} \end{pmatrix} f$. It is well known that

$$\left\| \Gamma_{\left(\begin{array}{c|c} \mathbb{O} & \mathbb{O} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right)} \right\|_{\mathrm{e}} = \left\| \Gamma_{\left(\begin{array}{c|c} \mathbb{O} & \mathbb{O} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right)} \right\| = \left\| \begin{pmatrix} \Phi_{21} & \Phi_{22} \end{pmatrix} \right\|_{\infty}$$

and so

$$\left\| \begin{pmatrix} \Phi_{21} & \Phi_{22} \end{pmatrix} \right\|_{\infty} \le \|\Gamma_\Phi\|_{\mathrm{e}}.$$

Hence

$$\|\Gamma_\Phi\| \cdot \|g_2(\zeta)\|_{\mathbb{C}^{m_2}} \le \left\| \begin{pmatrix} \Phi_{21} & \Phi_{22} \end{pmatrix} \right\|_{\infty} \|f(\zeta)\|_{\mathbb{C}^n} \le \|\Gamma_\Phi\|_{\mathrm{e}} \|f(\zeta)\|_{\mathbb{C}^n}.$$

On the other hand

$$\|g(\zeta)\|^2_{\mathbb{C}^m} = \|g_1(\zeta)\|^2_{\mathbb{C}^{m_1}} + \|g_2(\zeta)\|^2_{\mathbb{C}^{m_2}} = \|f(\zeta)\|^2_{\mathbb{C}^n}.$$

Therefore

$$\|g_1\|^2_{\mathbb{C}^{m_1}} \ge \frac{\|\Gamma_\Phi\|^2 - \|\Gamma_\Phi\|^2_{\mathrm{e}}}{\|\Gamma_\Phi\|^2} \|g(\zeta)\|^2_{\mathbb{C}^m}.$$

To prove the inequality for $f_1(\zeta)$ we can use the same argument since $t_0 f(\zeta) = \Phi^*(\zeta) g(\zeta)$ and

$$\left\| \begin{pmatrix} \Phi^*_{12} & \Phi^*_{22} \end{pmatrix} \right\| = \left\| \begin{pmatrix} \Phi_{12} \\ \Phi_{22} \end{pmatrix} \right\|_{\infty} = \left\| \Gamma_{\left(\begin{array}{c|c} \mathbb{O} & \Phi_{12} \\ \hline \mathbb{O} & \Phi_{22} \end{array}\right)} \right\|_{\mathrm{e}} = \left\| \Gamma_{\left(\begin{array}{c|c} \mathbb{O} & \Phi_{12} \\ \hline \mathbb{O} & \Phi_{22} \end{array}\right)} \right\|. \blacksquare$$

Corollary 3.4. *Let Φ be a matrix function of the form* (2.1) *such that $\|\Gamma_\Phi\|_{\mathrm{e}} < \|\Gamma_\Phi\|$. Suppose that* $\min\{m_1, n_1\} = 1$. *Then there exists a unique optimal solution of the four block problem.*

Proof. It is sufficient to consider the case $n_1 = 1$. To obtain the result in the case $m_1 = 1$ we can pass to the transpose of Φ. Let $f = f_1 \oplus f_2$ be a maximizing vector of Γ_Φ and let $Q \in H^\infty(M_{m_1 1})$ be an optimal solution of the four block problem, i.e.,

$$\left\| \begin{pmatrix} \Phi_{11} - Q & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \right\| = \|\Gamma_\Phi\|.$$

It follows from Lemma 3.3 that f_1 is a nonzero scalar function in H^2. By Lemma 3.2

$$\begin{pmatrix} \Phi_{11} - Q & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \Gamma_\Phi f.$$

Therefore Qf_1 is uniquely determined by Γ_Φ and f (Q is a column matrix) and since $f_1 \neq \mathbb{O}$, it follows that Q is uniquely determined by Φ. ■

Now, following ideas of [PY1] we construct diagonalizing matrix functions V and W in the following way.

Take a maximizing vector $f = f_1 \oplus f_2$, $f_1 \in H^2(\mathbb{C}^{n_1})$, $f_2 \in L^2(\mathbb{C}^{n_2})$. Let h be a scalar outer function such that $|h(\zeta)|^2 = \|f(\zeta)\|^2_{\mathbb{C}^n}$, $\zeta \in \mathbb{T}$. Such a function h always exists since $f_1 \in H^2$ and $f_1 \neq \mathbb{O}$. Denote by ϑ a greatest common inner divisor of the entries of f_1 (it may happen that $\vartheta = \mathbf{1}$). Define the vector $v = v_1 \oplus v_2 \in H^\infty(\mathbb{C}^{n_1}) \oplus L^\infty(\mathbb{C}^{n_2})$ by $v = \bar{\vartheta} f/h$. Clearly, $\|v(z)\| = 1$ a.e. on $\mathbb{T}$. The vector v will be the first column of the matrix V.

We can represent the column function v_1 as $v_1 = v_{(o)} v_{(i)}$ where $v_{(o)}$ is a scalar outer function such that $|v_{(o)}(\zeta)| = \|v_1(\zeta)\|_{\mathbb{C}^{n_1}}$ a.e. on $\mathbb{T}$ and $v_{(i)}$ is an inner column function (the inner part of v_1), i.e., $\|v_{(i)}(\zeta)\|_{\mathbb{C}^{n_1}} = 1$ a.e. on $\mathbb{T}$. Applying Lemma 3.1 to $\mathbf{v} = v_{(i)}$, we obtain an inner and co-outer matrix V_c such that the matrix $\begin{pmatrix} v_{(i)} & \overline{V}_c \end{pmatrix}$ is unitary-valued. Note that the vector-function $v_{(i)}$ is pointwise orthogonal to any column of $\overline{V}_c$ a.e. on $\mathbb{T}$, and so the same is true for v_1. So the matrix function

$$\begin{pmatrix} v_1 & \overline{V}_c \\ v_2 & \mathbb{O} \end{pmatrix} \tag{3.1}$$

is isometric almost everywhere on $\mathbb{T}$.

It is easy to see that we can complete this matrix function by adding m_2 measurable column functions to obtain a unitary-valued function. Indeed it is sufficient to complete the matrix function to a square matrix function whose columns are pointwise linear independent and then apply the Gram-Schmidt orthogonalization process to the columns. To this end we can approximate our matrix function uniformly by step functions which take isometric values. Clearly, we can find a unitary completion for each step function. It is easy to see that if the distance from a step function to our initial function is sufficiently small, then the columns of our initial function and the columns we added to the step function are linearly independent.

Let V be a unitary-valued completion of the matrix (3.1). Then V has the form

$$V = \left(\begin{array}{cc|c} v_1 & \overline{V}_c & \star \\ \hline v_2 & \mathbb{O} & \star \end{array}\right)$$

Let us now construct a unitary-valued matrix W in a similar way. Let $t_0 g = \Gamma_\Phi f$. Then by Lemma 3.2, $\|g(\zeta)\| = \|f(\zeta)\| = |h(\zeta)|$, $\zeta \in \mathbb{T}$. Let τ be a greatest common inner divisor of all entries of $\bar{z}\overline{g}_1$ (recall that $g = g_1 \oplus g_2$, $g_1 \in H^2_-(\mathbb{C}^{m_1})$, $g_2 \in L^2(\mathbb{C}^{m_2})$). Define the column function $w = w_1 \oplus w_2 \in H^\infty(\mathbb{C}^{m_1}) \oplus L^\infty(\mathbb{C}^{m_2})$ by $w \stackrel{\text{def}}{=} \bar{z}\bar{\tau}\bar{g}/h$. By analogy with (3.1) we can find a co-outer matrix function W_c such that the matrix function $\begin{pmatrix} w_1 & \overline{W}_c \\ w_2 & \mathbb{O} \end{pmatrix}$ takes isometric values on $\mathbb{T}$. We can complete this matrix function to a unitary-valued matrix function and define W to be its transpose:

$$W^t = \left(\begin{array}{cc|c} w_1 & \overline{W}_c & \star \\ \hline w_2 & \mathbb{O} & \star \end{array}\right).$$

To prove Theorem 2.1 we shall proceed as follows. Let Q_0 be an optimal solution of the four block problem. We shall prove in the next section that the matrix function $\begin{pmatrix} \Phi_{11} - Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}$ admits a representation

$$\left(\begin{array}{c|c} \Phi_{11} - Q_0 & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right) = W^* \left(\begin{array}{cc|c} t_0 u_0 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \Phi_{11}^{(1)} & \Phi_{12}^{(1)} \\ \hline \mathbb{O} & \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{array}\right) V^*,$$

where $u_0 = \bar{z}\bar{\vartheta}\bar{\tau}\bar{h}/h$ and $\Phi_{11}^{(1)}$ is a matrix function of size $(m_1 - 1) \times (n_1 - 1)$. We shall also prove in Section 4 that if Q is another optimal solution, then

$$\left(\begin{array}{c|c} \Phi_{11} - Q & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right) = W^* \left(\begin{array}{cc|c} t_0 u_0 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \Phi_{11}^{(1)} - Q_1 & \Phi_{12}^{(1)} \\ \hline \mathbb{O} & \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{array}\right) V^*,$$

where $Q_1 \in H^\infty(M_{m_1-1,n_1-1})$ and

$$\left\| \begin{pmatrix} \Phi_{11}^{(1)} - Q_1 & \Phi_{12}^{(1)} \\ \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{pmatrix} \right\|_\infty \le t_0.$$

Since V and W are unitary-valued, it is easy to see that Q is a superoptimal solution to the four block problem for the matrix function Φ if and only if Q_1 is a superoptimal solution to the four block problem for the matrix function

$$\Phi^{(1)} = \left(\begin{array}{c|c} \Phi_{11}^{(1)} & \Phi_{12}^{(1)} \\ \hline \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{array}\right).$$

Moreover, if $t_0, t_1, \cdots, t_{d-1}$ is sequence of superoptimal singular values of the four block problem for Φ, then $t_1, \cdots, t_{d-1}$ is the sequence of superoptimal singular values of the four block problem for $\Phi^{(1)}$. This reduction allows us to diminish the size of the matrix function Φ_{11}.

If $\Gamma_{\Phi^{(1)}} = \mathbb{O}$, we clearly have uniqueness. To continue this process we have to be able to find a maximizing vector for the four block operator $\Gamma_{\Phi^{(1)}}$. We can certainly do that if its essential norm is still less than the smallest nonzero superoptimal singular value. In Section 6 we shall prove that $\|\Gamma_{\Phi^{(1)}}\|_{\mathrm{e}} \le \|\Gamma_{\Phi}\|_{\mathrm{e}}$ which will allow us to continue the process and reduce Theorem 2.1 to Corollary 3.4.

4. Parametrization of optimal solutions

In this section we describe the optimal solutions of the four block problem in case when $\|\Gamma_\Phi\|_{\mathrm{e}} < \|\Gamma_\Phi\|$.

Lemma 4.1. *Let Φ be a block matrix function of the form* (2.1) *such that $\|\Gamma_\Phi\|_{\mathrm{e}} < \|\Gamma_\Phi\|$, and let V and W be the matrix functions constructed in Section 3. Then there exists a unimodular function u_0 such that any optimal solution Q_0 of the four block problem satisfies*

$$\left(\begin{array}{c|c} \Phi_{11} - Q_0 & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right) = W^* \left(\begin{array}{cc|c} t_0 u_0 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \Phi_{11}^{(1)} & \Phi_{12}^{(1)} \\ \hline \mathbb{O} & \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{array}\right) V^*, \tag{4.1}$$

where $\Phi_{11}^{(1)}$ is a matrix function of size $m_1 - 1 \times n_1 - 1$. The unimodular function u_0 admits a representation $u_0 = \bar{z}\bar{b}\bar{h}/h$, where h is an outer function in H^2 and b is a finite Blaschke product. Moreover, the Toeplitz operator T_{u_0} is Fredholm and $\operatorname{Range} T_{u_0} = H^2$.

Proof. Let $f = f_1 \oplus f_2$ be a maximizing vector for Γ_Φ and let $\Gamma_\Phi f = t_0 g = t_0(g_1 \oplus g_2)$. Put $u_0 = \bar{z}\bar{\vartheta}\bar{\tau}\bar{h}/h$ (see the construction of the matrix functions V and W in Section 3). By Lemma 3.2, $f(\zeta)$ is a maximizing vector for $\begin{pmatrix} \Phi_{11}(\zeta) - Q_0(\zeta) & \Phi_{12}(\zeta) \\ \Phi_{21}(\zeta) & \Phi_{22}(\zeta) \end{pmatrix}$ almost everywhere on $\mathbb{T}$ and

$$\begin{pmatrix} \Phi_{11} - Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} f = t_0 g.$$

Therefore $g(\zeta)$ is a maximizing vector for $\begin{pmatrix} \Phi_{11}(\zeta)-Q_0(\zeta) & \Phi_{12}(\zeta) \\ \Phi_{21}(\zeta) & \Phi_{22}(\zeta) \end{pmatrix}^*$ for almost all $\zeta\in\mathbb{T}$ and so

$$\begin{pmatrix} \Phi_{11}-Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}^* g = t_0 f. \tag{4.2}$$

Since $f=\vartheta h\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\bar z\bar g=\tau h\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, we have

$$\vartheta h\begin{pmatrix} \Phi_{11}-Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=t_0\bar z\bar\tau\bar h\begin{pmatrix} \overline{w}_1 \\ \overline{w}_2 \end{pmatrix}.$$

It follows from the definition of the matrix functions V and W (see Section 3) that

$$\vartheta h\begin{pmatrix} \Phi_{11}-Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}V\begin{pmatrix} \mathbf{1} \\ \mathbb{O} \\ \vdots \\ \mathbb{O} \end{pmatrix}=t_0\bar z\bar\tau\bar h W^*\begin{pmatrix} \mathbf{1} \\ \mathbb{O} \\ \vdots \\ \mathbb{O} \end{pmatrix}.$$

It is easy to see that the first column of $W\begin{pmatrix} \Phi_{11}-Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}V$ has the form $\begin{pmatrix} t_0u_0 & \mathbb{O} & \cdots & \mathbb{O} \end{pmatrix}^t$, where $u_0 \stackrel{\text{def}}{=} \bar z\bar\vartheta\bar\tau\bar h/h$. Similarly, using (4.2) we find that the first row of $W\begin{pmatrix} \Phi_{11}-Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}V$ has the form $\begin{pmatrix} t_0u_0 & \mathbb{O} & \cdots & \mathbb{O} \end{pmatrix}$, which proves that $\begin{pmatrix} \Phi_{11}-Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}$ has the form (4.1).

Let us show that the Toeplitz operator T_{u_0} is Fredholm and is onto. Clearly, $\|H_{u_0}\|=1$, since $\|H_{u_0}h\|_2=\|\mathbb{P}_-\bar z\bar\vartheta\bar\tau\|_2=\|\bar z\bar\vartheta\bar\tau\|_2=\|h\|_2$. We claim that $\|H_{u_0}\|_{\mathrm{e}}<1$. Indeed, let f be a scalar function in H^2. We have

$$\begin{aligned}
\Gamma_\Phi vf &= \mathbb{P}_-\begin{pmatrix} \Phi_{11}-Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}\begin{pmatrix} v_1 f \\ v_2 f \end{pmatrix} \\
&= \mathbb{P}_-W^*\begin{pmatrix} t_0u_0 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \Phi_{11}^{(1)} & \Phi_{12}^{(1)} \\ \mathbb{O} & \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{pmatrix}V^*\begin{pmatrix} v_1 f \\ v_2 f \end{pmatrix} \\
&= \mathbb{P}_-W^*\begin{pmatrix} t_0u_0 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \Phi_{11}^{(1)} & \Phi_{12}^{(1)} \\ \mathbb{O} & \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{pmatrix}\begin{pmatrix} f \\ \mathbb{O} \\ \vdots \\ \mathbb{O} \end{pmatrix}=t_0\mathbb{P}_-\begin{pmatrix} \overline{w}_1u_0f \\ \overline{w}_2u_0f \end{pmatrix}
\end{aligned}$$

Therefore

$$\mathbb{P}_- w^t \Gamma_\Phi v f = t_0 \mathbb{P}_- w^t \mathbb{P}_- \begin{pmatrix} \overline{w}_1 u_0 f \\ \overline{w}_2 u_0 f \end{pmatrix} = t_0 \mathbb{P}_- \begin{pmatrix} w_1 & w_2 \end{pmatrix} \begin{pmatrix} \overline{w}_1 u_0 f \\ \overline{w}_2 u_0 f \end{pmatrix} = t_0 H_{u_0} f,$$

whence

$$t_0 \|H_{u_0}\|_{\mathrm{e}} \le \|v\|_2 \|w\|_2 \|\Gamma_\Phi\|_{\mathrm{e}} = \|\Gamma_\Phi\|_{\mathrm{e}} < t_0,$$

which implies that $\|H_{u_0}\|_{\mathrm{e}} < 1$.

Since $\|H_{u_0}\|_{\mathrm{e}} = \operatorname{dist}_{L^\infty}(u_0, H^\infty + C)$ (see e.g., [S], [Ni]), it follows that $\|H_{u_0}\|_{\mathrm{e}} = \lim_{j\to\infty} \operatorname{dist}_{L^\infty}(z^j u_0, H^\infty)$. We have $\|H_{u_0} h\| = \|h\|$. So $\operatorname{dist}(u_0, H^\infty) = \|H_{u_0}\| = 1$. Therefore there exists a $j \in \mathbb{Z}_+$ such that $\operatorname{dist}_{L^\infty}(z^j u_0, H^\infty) = 1$ and $\operatorname{dist}_{L^\infty}(z^{j+1} u_0, H^\infty) < 1$. This means that $T_{z^{j+1}u_0}$ is left invertible and $T_{z^j u_0}$ is not left invertible which implies that $T_{z^{j+1}u_0}$ is invertible (see [Ni]). Clearly, $T_{z^{j+1}u_0} = T_{u_0} T_{z^{j+1}}$, $T_{z^{j+1}}$ is Fredholm and so is T_{u_0}.

Since u_0 has the form $u_0 = \bar z \bar\vartheta \bar\tau \bar h / h$, where ϑ and τ are inner and h is an outer function in H^2, the Toeplitz operator has dense range (see [PKh]) which together with the Fredholmness of T_{u_0} implies that T_{u_0} is onto.

It remains to show that both ϑ and τ are finite Blaschke products. Indeed, if κ is an inner divisor of $\vartheta\tau$, it is easy to see that $\kappa h \in \operatorname{Ker} T_{u_0}$ and since T_{u_0} is Fredholm, $\operatorname{Ker} T_{u_0}$ is finite dimensional, which implies that both ϑ and τ are finite Blaschke products. ■

Theorem 4.2. *Let Φ be a block matrix function of the form (2.1) such that $\|\Gamma_\Phi\|_{\mathrm{e}} < \|\Gamma_\Phi\|$ and let Q_0 be an optimal solution of the four block problem. Suppose that V, W, u_0, $\Phi_{11}^{(1)}$, $\Phi_{12}^{(1)}$, $\Phi_{21}^{(1)}$, $\Phi_{22}^{(1)}$ satisfy (4.1) holds. Let Q be a matrix function of size $m_1 \times n_1$. Then Q is an optimal solution of the four block problem if and only if there exists $Q_1 \in H^\infty(M_{m_1-1,n_1-1})$ that satisfies the following conditions:*

$$\left(\begin{array}{c|c} \Phi_{11} - Q & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right) = W^* \left(\begin{array}{cc|c} t_0 u_0 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \Phi_{11}^{(1)} - Q_1 & \Phi_{12}^{(1)} \\ \hline \mathbb{O} & \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{array}\right) V^*, \tag{4.3}$$

$$\left\| \begin{pmatrix} \Phi_{11}^{(1)} - Q_1 & \Phi_{12}^{(1)} \\ \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{pmatrix} \right\|_\infty \le t_0. \tag{4.4}$$

To prove Theorem 4.2 we need the following result from [PY1]:

Lemma 4.3. *Let $\mathbf{V}, \mathbf{W}$ be L^∞ matrix functions on $\mathbb{T}$, of types $n \times n$, $m \times m$ respectively, which are unitary-valued a.e. and are of the form*

$$\mathbf{V} = \begin{pmatrix} \mathbf{v} & \overline{V}_c \end{pmatrix}, \quad \mathbf{W}^t = \begin{pmatrix} \mathbf{w} & \overline{W}_c \end{pmatrix},$$

where $\mathbf{v}, V_c, \mathbf{w}, W_c$ *are* H^∞ *matrix functions,* $\mathbf{v}$ *and* $\mathbf{w}$ *are column functions, and* V_c, W_c *are co-outer. Then*

$$\mathbf{W}H^\infty(M_{m,n})\mathbf{V}\bigcap\begin{pmatrix}\mathbb{O} & \mathbb{O}\\ \mathbb{O} & L^\infty(M_{m-1,n-1})\end{pmatrix}=\begin{pmatrix}\mathbb{O} & \mathbb{O}\\ \mathbb{O} & H^\infty(M_{m-1,n-1})\end{pmatrix}.$$

Proof of Theorem 4.2. Let Q be an optimal solution. By Lemma 4.1 $W\left(\begin{array}{c|c} Q_0-Q & \mathbb{O}\\ \hline \mathbb{O} & \mathbb{O}\end{array}\right)V$ has the form

$$\left(\begin{array}{cc|c}\mathbb{O} & \mathbb{O} & \mathbb{O}\\ \mathbb{O} & \star & \star\\ \hline \mathbb{O} & \star & \star\end{array}\right)$$

(the upper left block is scalar). On the other hand it is easy to see from the definition of V and W (see Section 3) that

$$W\left(\begin{array}{c|c} Q_0-Q & \mathbb{O}\\ \hline \mathbb{O} & \mathbb{O}\end{array}\right)V=\left(\begin{array}{c|c}\begin{pmatrix} w_1 & \overline{W}_c\end{pmatrix}^t(Q_0-Q)\begin{pmatrix} v_1 & \overline{V}_c\end{pmatrix} & \mathbb{O}\\ \hline \mathbb{O} & \mathbb{O}\end{array}\right).$$

Therefore

$$\begin{pmatrix} w_1 & \overline{W}_c\end{pmatrix}^t(Q_0-Q)\begin{pmatrix} v_1 & \overline{V}_c\end{pmatrix}=\begin{pmatrix}\mathbb{O} & \mathbb{O}\\ \mathbb{O} & F\end{pmatrix}$$

for some $F\in L^\infty(M_{m_1-1,n_1-1})$ (the upper left corner of the matrix function on the right hand side is scalar). Let $v_1=v_{(o)}v_{(i)}$, $w_1=w_{(o)}w_{(i)}$, where $v_{(o)}$ and $w_{(o)}$ are scalar outer functions, and $v_{(i)}$ and $w_{(i)}$ are inner column functions. We have

$$\begin{aligned}&\begin{pmatrix} w_{(i)} & \overline{W}_c\end{pmatrix}^t(Q_0-Q)\begin{pmatrix} v_{(i)} & \overline{V}_c\end{pmatrix}\\ &=\begin{pmatrix} w_{(o)} & \mathbb{O}\\ \mathbb{O} & I\end{pmatrix}\begin{pmatrix} w_1 & \overline{W}_c\end{pmatrix}^t(Q_0-Q)\begin{pmatrix} v_1 & \overline{V}_c\end{pmatrix}\begin{pmatrix} v_{(o)} & \mathbb{O}\\ \mathbb{O} & I\end{pmatrix}\\ &=\begin{pmatrix} w_{(o)} & \mathbb{O}\\ \mathbb{O} & I\end{pmatrix}\begin{pmatrix}\mathbb{O} & \mathbb{O}\\ \mathbb{O} & F\end{pmatrix}\begin{pmatrix} v_{(o)} & \mathbb{O}\\ \mathbb{O} & I\end{pmatrix}=\begin{pmatrix}\mathbb{O} & \mathbb{O}\\ \mathbb{O} & F\end{pmatrix}.\end{aligned}$$

Put $\mathbf{v}=v_{(i)}$, $\mathbf{w}=w_{(i)}$. Clearly, the matrix functions $\mathbf{V}=\begin{pmatrix}\mathbf{v} & \overline{V}_c\end{pmatrix}$ and $\mathbf{W}=\begin{pmatrix}\mathbf{w} & \overline{W}_c\end{pmatrix}$ satisfy the hypotheses of Lemma 4.3. Therefore $F\in H^\infty(M_{m_1-1,n_1-1})$, which proves that Q satisfies (4.3) with $Q_1=-F$. Since Q is an optimal solution, (4.4) obviously holds.

Conversely, suppose that Q_1 is a function in $H^\infty(M_{m_1-1,n_1-1})$ satisfying (4.3). Then it follows from Lemma 4.3 that there exists a function $G \in H^\infty(M_{m_1,n_1})$ such that

$$\begin{pmatrix} w_1 & \overline{W}_c \end{pmatrix}^t G \begin{pmatrix} v_1 & \overline{V}_c \end{pmatrix} = \begin{pmatrix} \mathbb{O} & \mathbb{O} \\ \mathbb{O} & -Q_1 \end{pmatrix},$$

which implies that

$$\left(\begin{array}{c|c} \Phi_{11} - (G + Q_0) & \Phi_{12} \\ \hline \Phi_{21} & \Phi_{22} \end{array}\right) = W^* \left(\begin{array}{cc|c} t_0 u_0 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \Phi_{11}^{(1)} - Q_1 & \Phi_{12}^{(1)} \\ \hline \mathbb{O} & \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{array}\right) V^*$$

and so $Q = G + Q_0 \in H^\infty(M_{m_1,n_1})$. Clearly, (4.4) implies now that Q is an optimal solution. ■

It is easy to see that Theorem 4.2 reduces the problem of finding a superoptimal solution for Φ to the same problem for the matrix function $\left(\begin{array}{c|c} \Phi_{11}^{(1)} & \Phi_{12}^{(1)} \\ \hline \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{array}\right)$ which has a lower size.

5. The matrix functions V_c and W_c are left invertible in H^∞

In the last section we reduced the problem of finding a superoptimal solution for Φ to the same problem for $\Phi^{(1)} \stackrel{\text{def}}{=} \left(\begin{array}{c|c} \Phi_{11}^{(1)} & \Phi_{12}^{(1)} \\ \hline \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{array}\right)$. If we could continue this process, we would eventually reduce the problem to the case $\min\{m_1, n_1\} = 1$ and it would follow from Corollary 3.4 that there is a unique superoptimal solution to the four block problem for Φ. The main problem now is to prove that the four block operator Γ_{Φ_1} has a maximizing vector. This is certainly the case if $\|\Gamma_{\Phi_1}\|_{\mathrm{e}} \le \|\Gamma_\Phi\|_{\mathrm{e}}$. To prove this inequality we use an idea of [PY2] based on the solution of the so-called matricial corona problem for the matrix functions V_c and W_c. However in our case the solvability of this corona problem is much harder than in [PY2] where $V_c,\ W_c \in QC$.

In this section we shall prove that the matrix functions V_c and W_c are left invertible in H^∞ (in other words the corona problem is solvable for them) which we shall use in the next section to prove that $\|\Gamma_{\Phi_1}\|_{\mathrm{e}} \le \|\Gamma_\Phi\|_{\mathrm{e}}$.

Theorem 5.1. *If $\|\Gamma_\Phi\|_{\mathrm{e}} < \|\Gamma_\Phi\|$, then the matrix functions V_c and W_c defined in Section 3 are left invertible in H^∞.*

Clearly, it is sufficient to prove that W_c is left invertible in H^∞, which means that there exists a matrix function Ω in $H^\infty(M_{n_1,n_1-1})$ such that $\Omega(\zeta)W_c(\zeta) = I$ for every $\zeta \in \mathbb{D}$. To show the left invertibility of V_c, it is sufficient to apply Theorem 5.1 to the transposed function Φ^t and use the equalities $\|\Gamma_\Phi\| = \|\Gamma_{\Phi^t}\|$ and $\|\Gamma_\Phi\|_{\mathrm{e}} = \|\Gamma_{\Phi^t}\|_{\mathrm{e}}$, which follow immediately from the obvious identity

$$\Gamma_{\Phi^t} = J\Gamma^*_\Phi J,$$

where $J\rho \stackrel{\mathrm{def}}{=} \bar{z}\bar{\rho}$ for a vector function ρ in L^2.

Recall that $\overline{w} = \overline{w}_1 \oplus \overline{w}_2$ is the first column of W^*. Denote by w_{1_r}, $1 \le r \le n_1$, the components of w_1. We have $w_1 = w_{(o)}w_{(i)}$, where $w_{(o)}$ is a scalar outer function in H^2 and $w_{(i)}$ is an inner column function.

Lemma 5.2. *The vectorial Toeplitz operator $T_{\overline{w}_1} : H^2 \to H^2(\mathbb{C}^{n_1})$ is left invertible.*

Proof. First of all, $\operatorname{Ker} T_{\overline{w}_1} = \{\mathbb{O}\}$. Indeed, assume that $\psi \in \operatorname{Ker} T_{\overline{w}_1}$. Then $\overline{w}_{1_r}\psi \in H^2_-$ for $1 \le r \le n_1$. Since $\mathbf{1}$ is a greatest inner divisor of the components of w_1, it follows from Beurling's theorem that the functions

$$\{\sum_{r=1}^{n_1} \kappa_r w_{1_r} : \ \kappa_r \in H^2\}$$

form a dense subset in H^2. Therefore we can approximate ψ in the L^1-norm by functions of the form $\sum_{r=1}^{n_1} \bar{\kappa}_r \bar{w}_{1_r}\psi$, each of which belongs to $H^1_- \stackrel{\mathrm{def}}{=} \{\varphi \in L^1 : \ \hat{\varphi}(k) = 0 \text{ for } k \ge 0\}$. Hence $\psi \in H^1_-$ and since $\psi \in H^2$, it follows that $\psi = \mathbb{O}$.

If $T_{\overline{w}_1}$ is not left invertible, there exists a sequence of scalar functions $\{\varphi_j\}_{j\ge 0}$ in H^2 such that $\|\varphi_j\| = 1$ and $\varphi_n \to \mathbb{O}$ in the weak topology and $\|T_{\overline{w}_1}\varphi_n\| \to 0$. By Lemma 4.1 the operator T_{u_0} is onto and so there exists a sequence $\{\omega_n\}_{n\ge 0}$ of scalar functions in $(\operatorname{Ker} T_{u_0})^\perp$ such that $T_{u_0}\omega_n = \varphi_n$. Since T_{u_0} is Fredholm, $\omega_n \to \mathbb{O}$ weakly. Put $\rho_n \stackrel{\mathrm{def}}{=} \omega_n v \in H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2})$, where v is the first column of V. Let Q_0 be an optimal solution of the four block problem for Φ. By (4.1) we have

$$\begin{pmatrix} \Phi_{11} - Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \rho_j = W^* \begin{pmatrix} t_0 u_0 \omega_j \\ \mathbb{O} \\ \vdots \\ \mathbb{O} \end{pmatrix} = t_0 u_0 \omega_j \overline{w} = t_0(\varphi_j + \varphi_j^-)\overline{w},$$

for some functions $\varphi_j^- \in H^2_-$. It follows that

$$\begin{aligned}\|\Gamma_\Phi \rho_j\|^2 &= \left\|\mathbb{P}^- \begin{pmatrix} \Phi_{11} - Q_0 & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \rho_j \right\|^2 = \|t_0 \mathbb{P}^-(\varphi_j + \varphi_j^-)\overline{w}\|^2 \\ &= \|t_0(\varphi_j + \varphi_j^-)\overline{w}\|^2 - \|\mathbb{P}_+ t_0(\varphi_j + \varphi_j^-)\overline{w}_1\|^2 \\ &= \|t_0 u_0 \omega_j \overline{w}\|^2 - \|t_0 \mathbb{P}_+ \varphi_j \overline{w}_1\|^2 \\ &= \|t_0 u_0 \omega_j \overline{v}\|^2 - \|t_0 \mathbb{P}_+ \varphi_j \overline{w}_1\|^2 = t_0^2(\|\rho_j\|^2 - \|T_{\overline{w_1}} \varphi_j\|^2),\end{aligned}$$

since $\|v(\zeta)\|_{\mathbb{C}^{n_1}} = \|w(\zeta)\|_{\mathbb{C}^{m_1}}$.

Taking into account that $\|T_{\overline{w_1}} \varphi_j\| \to 0$ and $\rho_j \to \mathbb{O}$ weakly, we obtain $\|\Gamma_\Phi\|_e = t_0 = \|\Gamma_\Phi\|$ which contradicts the hypotheses of the lemma. ■

The next step is to prove that the Toeplitz operator $T_{\overline{w}_{(i)}}$ is left invertible, where $w_{(i)}$ is the inner part of w_1. We need the following well known facts. Let $\chi = \{\chi_j\}_{1 \le j \le k}$ be a column function in $H^\infty(\mathbb{C}^k)$. Then it is left invertible in H^∞ (i.e. there exist functions κ_j, $1 \le j \le k$, such that $\sum_{j=1}^k \kappa_j(\zeta)\chi_j(\zeta) = 1$ for all $\zeta \in \mathbb{D}$) if and only if the Toeplitz operator $T_{\bar{\chi}}$ is left invertible (see [Ar]). Note that by the Carleson corona theorem (see e.g., [Ni]) χ is left invertible if and only if $\inf_{\zeta \in \mathbb{D}} \|\chi(\zeta)\|_{\mathbb{C}^k} > 0$. This result was generalized in [SNF2] for the case of matrix (and even operator) functions: let Ξ be a matrix function in H^∞, then Ξ is left invertible in H^∞ if and only if the Toeplitz operator $T_{\overline{\Xi}}$ is left invertible.

Lemma 5.3. *Under the hypotheses of Theorem 5.1 the Toeplitz operator $T_{\overline{w}_{(i)}}$ is left invertible.*

Proof. By Lemma 5.2, $T_{\overline{w}_1}$ is left invertible. By Arveson's theorem mentioned above w_1 is left invertible in H^∞. We have $w_1 = w_{(o)} w_{(i)}$, where $w_{(o)}$ is a scalar outer function in H^∞ and $w_{(i)}$ is an inner column function. Obviously, it follows that $w_{(i)}$ is left invertible in H^∞. Again by Arveson's theorem this implies that $T_{\overline{w}_{(i)}}$ is left invertible. ■

We need the following result proved in [P].

Theorem 5.4. *Let $\mathbf{W}$ be a unitary-valued matrix function of the form $\mathbf{W}^t = \begin{pmatrix} \mathbf{w} & \overline{W}_c \end{pmatrix}$, where $\mathbf{w}$ is a co-outer inner column, and W_c is a co-outer inner function. Then the Toeplitz operator $T_{\mathbf{W}^t}$ has trivial kernel and dense range, and the operators $H^*_{\mathbf{W}^t} H_{\mathbf{W}^t}$ and $H^*_{(\mathbf{W}^t)^*} H_{(\mathbf{W}^t)^*}$ are unitarily equivalent.*

The following result can easily be deduced from Theorem 5.4

Theorem 5.5. *Let $\mathbf{W}$ be a matrix function satisfying the hypotheses of Theorem 5.4. Suppose that $\|H_{\overline{\mathbf{w}}}\| < 1$. Then the Toeplitz operator $T_{\mathbf{W}^t}$ is invertible.*

Proof. Clearly, $\|H_{(\mathbf{W}^t)^*}\| = \|H_{\mathbf{w}^*}\| = \|H_{\overline{\mathbf{w}}^t}\|$. It is easy to see that $\|H_{\overline{\mathbf{w}}^t}\| = \|H_{\overline{\mathbf{w}}}\| < 1$. By Theorem 5.4 the operators $H^*_{\mathbf{W}^t}H_{\mathbf{W}^t}$ and $H^*_{(\mathbf{W}^t)^*}H_{(\mathbf{W}^t)^*}$ are unitarily equivalent. Therefore $\|H_{\mathbf{W}^t}\| = \|H_{(\mathbf{W}^t)^*}\| < 1$. Since $\mathbf{W}^t$ takes isometric values on $\mathbb{T}$, it is easy to see that

$$\|T_{\mathbf{W}^t}F\|_2^2 + \|H_{\mathbf{W}^t}F\|_2^2 = \|\mathbf{W}^tF\|_2^2 = \|F\|_2^2$$

for every vector function F. Consequently, $T_{\mathbf{W}^t}$ is left invertible if and only if $\|H_{\mathbf{W}^t}\| < 1$. It follows that both $T_{\mathbf{W}^t}$ and $T_{(\mathbf{W}^t)^*}$ are left invertible which means that $T_{\mathbf{W}^t}$ is invertible. ■

Proof of Theorem 5.1. Put $\mathbf{w} \stackrel{\text{def}}{=} w_{(i)}$ and let $\mathbf{W}^t = \begin{pmatrix} w_{(i)} & \overline{W}_c \end{pmatrix}$. By Lemma 5.3, $T_{\overline{w}_{(i)}}$ is left invertible. Since $w_{(i)}$ takes isometric values on $\mathbb{T}$, we have as in the proof of Theorem 5.5

$$\|T_{\overline{w}_{(i)}}\omega\|^2 + \|H_{\overline{w}_{(i)}}\omega\|^2 = \|\omega\|^2$$

for every $\omega \in H^2$. Hence $\|H_{\overline{w}_{(i)}}\| < 1$ and so by Theorem 5.5 the operator $T_{\mathbf{W}^t}$ is invertible. Clearly, it follows that $T_{\overline{W}_c}$ is left invertible, since $T_{\overline{W}_c}$ can be interpreted as a restriction of $T_{\mathbf{W}^t}$. Therefore by the Sz.-Nagy–Foias theorem mentioned above W_c is left invertible in H^∞. ■

6. The essential norm of $\Gamma_{\Phi^{(1)}}$

In Section 4 we reduced the proof of Theorem 2.1 to the fact that $\|\Gamma_{\Phi_{(1)}}\|_e \le \|\Gamma_\Phi\|_e$, where the matrix function $\Phi^{(1)} = \left(\begin{array}{c|c} \Phi^{(1)}_{11} & \Phi^{(1)}_{12} \\ \hline \Phi^{(1)}_{21} & \Phi^{(1)}_{22} \end{array}\right)$ is defined in (4.1). In this section we are going to use the facts that V_c and W_c are left invertible (see Section 5) to prove this inequality which will complete the proof of Theorem 2.1.

The idea behind the proof is the following. We use the fact that

$$\|\Gamma_{\Phi^{(1)}}\|_e = \inf\{\limsup_j \|\Gamma_{\Phi^{(1)}}\xi_j\|_2\},$$

where the *infimum* is taken over all sequences $\{\xi_j\}$ in $H^2(\mathbb{C}^{n_1-1}) \oplus L^2(\mathbb{C}^{n_2})$ such that $\|\xi_j\|_2 = 1$ and $\xi_j \to \mathbb{O}$ weakly. Given such a sequence $\{\xi_j\}$ we construct another sequence $\{\rho_j\}$ in $H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2})$ such that $\|\rho_j\|_2 = 1$, $\rho_j \to \mathbb{O}$ weakly and

$$\limsup_j \|\Gamma_\Phi \rho_j\|_2 \geq \limsup_j \|\Gamma_{\Phi^{(1)}} \xi_j\|_2.$$

To this end we are going to use a construction which is similar to the one used in [PY2].

Let W be the unitary-valued matrix function constructed in Section 3. Consider the matrix W^* which has the form

$$W^* = \left(\begin{array}{cc|c} \overline{w}_1 & W_c & F \\ \hline \overline{w}_2 & \mathbb{O} & G \end{array} \right).$$

Let

$$\beta = \left(\begin{array}{c|c} W_c & F \\ \hline \mathbb{O} & G \end{array} \right). \tag{6.1}$$

To use a construction similar to the one given in [PY2], we have to find a left inverse of β of a special form. Recall that we have proved in Section 5 that W_c is left invertible in H^∞. Let W_c^{li} be an H^∞ left inverse of W_c.

Lemma 6.1. *Let β be the matrix function defined by* (6.1). *Then the matrix function G is invertible in L^∞ and there exists a bounded left inverse of β of the form*

$$B = \left(\begin{array}{c|c} W_c^{\mathrm{li}} & X \\ \hline \mathbb{O} & G^{-1} \end{array} \right). \tag{6.2}$$

Proof. Suppose that G is not invertible in L^∞. Then there exists a sequence $\{\xi_j\}_{j\geq 0}$ in $L^2(\mathbb{C}^{m_2})$ such that $\|\xi_j\|_2 = 1$ and $\|G\xi_j\|_2 \to 0$.

It is easy to see from Lemma 3.3 that

$$\|w_1(\zeta)\|^2_{\mathbb{C}^{m_1}} \geq \frac{\|\Gamma_\Phi\|^2 - \|\Gamma_\Phi\|^2_{\mathrm{e}}}{\|\Gamma_\Phi\|^2} \stackrel{\mathrm{def}}{=} \delta < 1, \quad \zeta \in \mathbb{T}. \tag{6.3}$$

Since the column $\overline{w}_1(\zeta)$ is orthogonal to the columns of $W_c(\zeta)$ a.e. on $\mathbb{T}$, it follows that the matrix function $\begin{pmatrix} \overline{w}_1 & W_c \end{pmatrix}$ is invertible in L^∞. Therefore there exists a bounded sequence $\{\eta_j\}$ in $L^2(\mathbb{C}^{m_1})$ such that

$$\begin{pmatrix} \overline{w}_1 & W_c \end{pmatrix} \eta_j + F\xi_j = \mathbb{O}.$$

Then

$$W^* \begin{pmatrix} \eta_j \\ \xi_j \end{pmatrix} = \begin{pmatrix} \mathbb{O} \\ \begin{pmatrix} \overline{w}_2 & \mathbb{O} \end{pmatrix} \eta_j + G\xi_j \end{pmatrix}.$$

It follows from (6.3) that $\|w_2(\zeta)\|^2_{\mathbb{C}^{m_2}} \le 1-\delta$, $\zeta \in \mathbb{T}$. Since $\|G\xi_j\|_2 \to 0$, we have for large values of j

$$\left\| W^* \begin{pmatrix} \eta_j \\ \xi_j \end{pmatrix} \right\|_2 < \|\eta_j\|_2 \le \left\| \begin{pmatrix} \eta_j \\ \xi_j \end{pmatrix} \right\|_2,$$

which contradicts the fact that W is unitary-valued.

Let now B be a matrix in the form (6.2). Clearly $B\beta = I$ if and only if

$$W_c^{\rm li} F + XG = \mathbb{O}.$$

Since G is invertible in L^∞, we can always find a matrix function X in L^∞ which satisfies this equality. ■

Remark. In the same way we can consider the submatrix α of the matrix $\overline{V}$ constructed in Section 3,

$$\alpha = \left(\begin{array}{c|c} V_c & \star \\ \hline \mathbb{O} & \star \end{array} \right)$$

and prove that α has a left inverse in the form

$$A = \left(\begin{array}{c|c} V_c^{\rm li} & \star \\ \hline \mathbb{O} & \star \end{array} \right),$$

where $V_c^{\rm li}$ is an H^∞ left inverse of V_c.

To prove the main result of this section we need the following lemma which in the case of Nehari's problem was proved in [PY2] (see Lemma 2.1 there).

Lemma 6.2. *Let η be a vector function in $H^2_-(\mathbb{C}^{m_1-1}) \oplus L^2(\mathbb{C}^{m_2})$ and let χ be the scalar function in H^2 defined by*

$$\chi = -\mathbb{P}_+ w^t B^* \eta.$$

Then

$$W^* \begin{pmatrix} \chi \\ \eta \end{pmatrix} \in H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2}).$$

Proof. Since W is unitary-valued, we have

$$I = W^*W = \overline{w}w^t + \beta\beta^*$$

and hence

$$\beta = \beta(B\beta)^* = \beta\beta^* B^* = (I - \overline{w}w^t)B^*.$$

Therefore

$$W^*\begin{pmatrix}\chi\\ \eta\end{pmatrix}=\begin{pmatrix}\overline{w} & \beta\end{pmatrix}\begin{pmatrix}\chi\\ \eta\end{pmatrix}=\overline{w}\chi+(I-\overline{w}w^t)B^*\eta=B^*\eta+\overline{w}(\chi+w^tB^*\eta).$$

It is easy to see from (6.2) that $B^*\eta\in H^2_-(\mathbb{C}^{m_1})\oplus L^2(\mathbb{C}^{m_2})$. Since $w\in H^\infty(\mathbb{C}^{m_1})\oplus L^\infty(\mathbb{C}^{m_2})$, it follows that $\overline{w}(\chi+w^tB^*\eta)=\overline{w}\mathbb{P}_-w^tB^*\eta\in H^2_-$, which proves the result. ■

Remark. It is easy to see that if $\{\eta_j\}$ is a sequence of functions in $H^2_-(\mathbb{C}^{m_1-1})\oplus L^2(\mathbb{C}^{m_2})$ which converges weakly to $\mathbb{O}$, the above construction produces a sequence of scalar functions $\{\chi_j\}$ in H^2, $\chi_j=-\mathbb{P}_+(w^tB^*\eta_j)$, which also converges weakly to $\mathbb{O}$.

Now we are in a position to prove that $\|\Gamma_{\Phi_1}\|_{\rm e}\le\|\Gamma_\Phi\|_{\rm e}$, where the matrix function $\Phi^{(1)}=\left(\begin{array}{c|c}\Phi^{(1)}_{11} & \Phi^{(1)}_{12}\\ \hline \Phi^{(1)}_{21} & \Phi^{(1)}_{22}\end{array}\right)$ is defined in (4.1).

Theorem 6.3. *Let Φ be a matrix function of the form* (2.1) *such that $\|\Gamma_\Phi\|_{\rm e}<\|\Gamma_\Phi\|$. Then $\|\Gamma_{\Phi^{(1)}}\|_{\rm e}\le\|\Gamma_\Phi\|_{\rm e}$.*

Proof. Let $\{\xi_j\}$ be a sequence of functions in $H^2_-(\mathbb{C}^{n_1-1})\oplus L^2(\mathbb{C}^{n_2})$ such that $\|\xi_j\|_2=1$ and $\xi_j\to\mathbb{O}$ weakly. Put $\eta_j=\Gamma_{\Phi^{(1)}}\xi_j$. We are going to construct a sequence of functions $\{\xi_j^\#\}$ in $H^2_-(\mathbb{C}^{n_1})\oplus L^2(\mathbb{C}^{n_2})$ such that $\frac{\xi_j^\#}{\|\xi_j^\#\|_2}\to\mathbb{O}$ weakly, and $\frac{\|\eta_j^\#\|_2}{\|\xi_j^\#\|_2}\ge\|\eta_j\|_2$, where $\eta_j^\#\stackrel{\rm def}{=}\Gamma_\Phi\xi_j^\#$. As we have explained in the beginning of the section, this would imply the desired inequality (put $\rho_j=\frac{\xi_j^\#}{\|\xi_j^\#\|_2}$).

To this end we apply Lemma 6.2 to the sequence $\{\eta_j\}$. We obtain a sequence of scalar H^2 functions $\{\chi_j\}$ such that

$$W^*\begin{pmatrix}\chi_j\\ \eta_j\end{pmatrix}\in H^2_-(\mathbb{C}^{m_1})\oplus L^2(\mathbb{C}^{m_2}).$$

Put

$$\xi_j^\#=A^t\xi_j+q_jv,$$

where v is the first column of V and A is the left inverse of α described in the Remark after Lemma 6.1. The scalar functions q_j will be chosen later.

We have

$$\begin{pmatrix}t_0u_0 & \mathbb{O}\\ \mathbb{O} & \Phi^{(1)}\end{pmatrix}V^*\xi_j^\#=\begin{pmatrix}t_0u_0q_j+t_0u_0v^*A^t\xi_j\\ \Phi^{(1)}\xi_j\end{pmatrix}. \tag{6.4}$$

Since the Toeplitz operator T_{u_0} is onto, we can pick q_j as a solution of the equation

$$\mathbb{P}_+(t_0u_0q_j + t_ou_ov^*A^t\xi_j) = \chi_j.$$

Clearly, we may choose the q_j so that $q_j \to \mathbb{O}$ weakly. Indeed, we may put

$$q_j = t_0^{-1}(T_{u_0}|(\operatorname{Ker} T_{u_0})^\perp)^{-1}(t_0\chi_j - \mathbb{P}_+u_0v^*A^t\xi_j).$$

It follows that $\xi_j^\# \to \mathbb{O}$ weakly.

Let us show that the sequence $\{\xi_j^\#\}$ has the required properties. Since $\xi_j^\# \to \mathbb{O}$ weakly, to prove that $\frac{\xi_j^\#}{\|\xi_j^\#\|_2} \to \mathbb{O}$ weakly, we have to estimate $\|\xi_j^\#\|_2$ from below. We have

$$\|\xi_j^\#\|_2^2 = \|V^*\xi_j^\#\|_2^2 = \|q_j + v^*A^t\xi_j\|_2^2 + \|\xi_j\|_2^2 \ge 1. \tag{6.5}$$

To complete the proof it remains to show that $\|\Gamma_\Phi\xi_j^\#\|_2 \ge \|\eta_j\|_2$.

Recall that $\eta_j = \Gamma_{\Phi^{(1)}}\eta_j = \mathbb{P}^-\Phi^{(1)}\xi_j$ and so $\Phi^{(1)}\xi_j - \eta_j \in H^2(\mathbb{C}^{m_1-1}) \oplus \{\mathbb{O}\}$. It is easy to see from the definition of W (see Section 3) that

$$W^*\begin{pmatrix} \mathbb{O} \\ \Phi^{(1)}\xi_j - \eta_j \end{pmatrix} \in H^2(\mathbb{C}^{m_1}) \oplus \{\mathbb{O}\}.$$

It follows now from (6.4) that

$$\eta_j^\# = \mathbb{P}^-\Phi\xi_j^\# = \mathbb{P}^-W^*\begin{pmatrix} t_0u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{pmatrix}V^*\xi_j^\# = \mathbb{P}^-W^*\begin{pmatrix} \chi_j + \omega_j \\ \eta_j \end{pmatrix},$$

where

$$\omega_j \stackrel{\text{def}}{=} \mathbb{P}_-(t_0u_0q_j + t_0u_0v^*A^t\xi_j).$$

Since the first column of W^* is $\overline{w}_1 \oplus \overline{w}_2$ and $w_1 \in H^\infty$, it follows that

$$W^*\begin{pmatrix} \omega_j \\ \mathbb{O} \end{pmatrix} \in \begin{pmatrix} H^2_-(\mathbb{C}^{m_1}) \\ L^2(\mathbb{C}^{m_2}) \end{pmatrix}.$$

We have chosen $\{\chi_j\}$ so that

$$W^*\begin{pmatrix} \chi_j \\ \eta_j \end{pmatrix} \in \begin{pmatrix} H^2_-(\mathbb{C}^{m_1}) \\ L^2(\mathbb{C}^{m_2}) \end{pmatrix}.$$

Therefore

$$\mathbb{P}^-W^*\begin{pmatrix} \chi_j + \omega_j \\ \eta_j \end{pmatrix} = W^*\begin{pmatrix} \chi_j + \omega_j \\ \eta_j \end{pmatrix}.$$

Hence
$$\|\eta_j^{\#}\|_2^2 = \|\chi_j + \omega_j\|_2^2 + \|\eta_j\|_2^2 = t_0^2\|q_j + v^* A^t \xi_j\|_2^2 + \|\Gamma_{\Phi^{(1)}}\xi_j\|^2.$$
Since $\|\Gamma_{\Phi^{(1)}}\| \le t_0$, this together with (6.5) yields
$$\frac{\|\eta_j^{\#}\|_2}{\|\xi_j^{\#}\|_2} \ge \frac{\|\eta_j\|_2}{\|\xi_j\|_2} = \|\eta_j\|_2,$$
which completes the proof. ■

As we have already explained, Theorem 6.3 allows us to complete the proof of Theorem 2.1.

Proof of Theorem 2.1. By Theorem 4.2 the four block for Φ has a unique solution if so does the four block problem for $\Phi^{(1)}$ and the superoptimal singular values of the four block problem for $\Phi^{(1)}$ are $t_1, t_2, \cdots, t_{d-1}$. By Theorem 6.3, $\|\Gamma_{\Phi^{(1)}}\|_{\mathrm{e}} \le \|\Gamma_\Phi\|_{\mathrm{e}}$. If $\Gamma_{\Phi^{(1)}} = \mathbb{O}$, we certainly have uniqeness. Otherwise we can continue this process.

Doing in this way we may stop the process if we get on a certain stage the zero four block operator or, otherwise, we eventually reduce the problem to the case $d = 1$. Uniqueness follows now from Corollary 3.4.

The fact that the singular values are constant on $\mathbb{T}$ follows immediately from the facts that V and W are unitary-valued and u_0 is unimodular, and from Lemma 3.2. ■

7. Thematic factorizations and indices of superoptimal singular values

In this section we analyze the algorithm described in Section 3 and obtain certain special factorizations of superoptimal symbols of four block operators satisfying the hypotheses of Theorem 2.1. Following [PY1] we shall call such factorizations *thematic.*

In Section 3 we have constructed matrix functions V and W associated with the four block problem. By analogy with [PY1] we shall call matrix functions of the form V or W *thematic functions.*

To state the result we may assume without loss of generality that $n_1 \le m_1$ (otherwise we can take the transpose).

Theorem 7.1. *Let Φ be a superoptimal symbol of the four block operator Γ which satisfies the hypotheses of Theorem 2.1 and suppose that $n_1 \le m_1$. Then Φ admits the following factorization*

$$\Phi = W_0^* W_1^* W_2^* \cdots W_{d-1}^* D V_{d-1}^* \cdots V_2^* V_1^* V_0^* \tag{7.1}$$

where

$$D = \left(\begin{array}{cccc|c} t_0 u_0 & \mathbb{O} & \cdots & \mathbb{O} & \\ \mathbb{O} & t_1 u_1 & \cdots & \mathbb{O} & \\ \vdots & \vdots & \ddots & \vdots & \\ \mathbb{O} & \mathbb{O} & \cdots & t_{d-1} u_{d-1} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \cdots & \mathbb{O} & \\ \vdots & \vdots & \ddots & \vdots & \\ \mathbb{O} & \mathbb{O} & \cdots & \mathbb{O} & \\ \hline & & \mathbb{O} & & \star \end{array} \right),$$

the u_j are unimodular functions such that the Toeplitz operator T_{u_j} is Fredholm and $\operatorname{ind} T_{u_j} > 0$*, and the matrix functions W_j and V_j have the form*

$$V_j = \begin{pmatrix} I_j & 0 \\ 0 & \breve{V}_j \end{pmatrix} \qquad W_j = \begin{pmatrix} I_j & 0 \\ 0 & \breve{W}_j \end{pmatrix},$$

where $\breve{V}_j$, $\breve{W}_j$ are thematic matrix functions and I_j is the identity $j \times j$ matrix.

It is easy to see that the successive application of the algorithm described in Section 3 gives us a desired factorization.

Remark. As in the case of Nehari's problem (see [PY1]) it is easy to see that if a matrix function admits a factorization of the form (7.1), then it is the superoptimal symbol of the corresponding four block operator.

We can associate with the factorization (7.1) the *factorization indices k_j* which are defined in the case $t_j \ne 0$. We put $k_j = \operatorname{ind} T_{u_j} = \dim \operatorname{Ker} T_{u_j}$.

It was shown in [PY1] that even for Nehari's problem the indices depend on the choice of a thematic factorization rather than on the function Φ itself. However, it was shown in [PY2] that for Nehari's problem with compact Hankel operator the sum of the indices corresponding to equal superoptimal singular values is an invariant (i.e. does not depend on the choice of a factorization).

The same turns out to be true for the four block problem too, and we shall prove this later in Section 9. Moreover, the sum of the indices corresponding to equal superoptimal singular values admits a quite natural and simple geometric interpretation. To give this interpretation we have to introduce a new object — the so–called *superoptimal weight.*

8. Superoptimal weight

Let $\mathcal{W} \in L^\infty(M_{n,n})$ be a *matrix weight*, i.e. a bounded matrix-valued function on $\mathbb{T}$, whose values are nonnegative selfadjoint $n \times n$ matrices.

Given a four block operator $\Gamma : H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2}) \to H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$, $n_1 + n_2 = n$, we call a weight $\mathcal{W}$ *admissible* if

$$\|\Gamma f\|^2 \le (\mathcal{W} f, f) \stackrel{\text{def}}{=} \int_{\mathbb{T}} (\mathcal{W}(\zeta) f(\zeta), f(\zeta)) d\boldsymbol{m}(\zeta), \quad f \in H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2}).$$

We need the following result which we call *Generalized Nehari's Theorem.*

Theorem 8.1. *Given a four block operator Γ_Φ and an admissible weight $\mathcal{W}$ there exists a symbol Φ of Γ (i.e. an operator-valued function Φ such that $\Gamma = \Gamma_\Phi$) satisfying $\Phi^*\Phi \le \mathcal{W}$.*

If Φ is a symbol of Γ satisfying $\Phi^*\Phi \le \mathcal{W}$, we say that Φ is *dominated by* the admissible weight $\mathcal{W}$.

In the case $\mathcal{W} \equiv cI$, $v \in \mathbb{R}_+$, this result was established in [FT], and this is an analog of Nehari's theorem for four block operators. In the general case the result follows from Theorem 1.1 of [TV], since the four block operator Γ acting from the space $H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2})$ endowed with the weighted norm $\|\cdot\|_{\mathcal{W}}$ to the space $H^2(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$ satisfies the hypothesis of the theorem.

For the sake of completeness we deduce Theorem 8.1 from the analog of Nehari's theorem mentioned above.

Proof of theorem 8.1. Define $\mathcal{W}_\varepsilon = \mathcal{W} + \varepsilon I$. Since $\mathcal{W}_\varepsilon \ge \varepsilon I$, it admits a factorization $\mathcal{W}_\varepsilon = G_\varepsilon^* G_\varepsilon$, where $G_\varepsilon \in H^\infty(M_{n,n})$ is a matrix function which is invertible in H^∞ (see [R]). The weight $\mathcal{W}_\varepsilon$ is clearly admissible, so

$$\|\Gamma f\| \le \|G_\varepsilon f\|, \qquad f \in H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2}),$$

which is equivalent to the fact that

$$\|\Gamma G_\varepsilon^{-1} f\| \le \|f\|, \qquad f \in H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2}),$$

Since $G_\varepsilon^{-1} \in H^\infty(M_{n,n})$, we can consider the operator $\Gamma G_\varepsilon^{-1}$ as a four block operator. By the analog of Nehari's theorem it has a symbol Ψ_ε such that $\|\Psi_\varepsilon\|_\infty \le 1$. Then the function $\Phi_\varepsilon = \Psi G_\varepsilon$ is a symbol of Γ and

$$\Phi_\varepsilon^* \Phi_\varepsilon = G_\varepsilon^* \Psi_\varepsilon^* \Psi_\varepsilon G_\varepsilon \le G_\varepsilon^* G_\varepsilon = \mathcal{W}_\varepsilon = \mathcal{W} + \varepsilon I.$$

It remains to chose a sequence $\{\varepsilon_j\}$ converging to 0 and such that the sequence $\{\Phi_{\varepsilon_j}\}$ converges to a matrix function, say $\Phi \in L^\infty$, in the $*$-weak topology. Clearly, Φ is a symbol of Γ dominated by $\mathcal{W}$. ■

Definition. Let $\mathcal{W}$ be an admissible weight for the four block operator G. Consider the numbers

$$s_j^\infty(\mathcal{W}) \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} s_j(\mathcal{W}(\zeta)), \quad 0 \le j \le d-1, \quad d = \min\{m_1, n_1\}.$$

The admissible weight $\mathcal{W}$ is called *superoptimal* if it lexicographically minimizes the numbers $s_0^\infty(\mathcal{W}),\ s_1^\infty(\mathcal{W}), \cdots, s_{d-1}^\infty(\mathcal{W})$ among all admissible weights, i.e.,

$$s_0^\infty(\mathcal{W}) = \min\{s_0^\infty(\mathcal{V}) :\ \mathcal{V} \text{ is admissible}\},$$

$$s_1^\infty(\mathcal{W}) = \min\{s_1^\infty(\mathcal{V}) :\ \mathcal{V} \text{ is admissible, } s_0^\infty(\mathcal{V}) \text{ is minimal possible}\}, \quad \text{etc.}$$

The following lemma shows that under the the hypotheses of Theorem 2.1 a superoptimal weight exists. However, a superoptimal weight is not unique in general. The lemma also shows that a superoptimal weight is nevertheless "essentially" unique for our purposes.

Let λ_a, $a \in \mathbb{R}$, be the function on $\mathbb{R}$ defined by

$$\lambda_a(t) \stackrel{\text{def}}{=} \begin{cases} t, & t \ge a \\ 0, & t < a \end{cases}.$$

Lemma 8.2. *Let Γ be a four block operator satisfying the hypothesis of Theorem 3.3. Let Φ be the superoptimal symbol of Γ. Then*

1. *$\Phi^*\Phi$ is a superoptimal weight for Γ;*
2. *If $\mathcal{W}$ and $\mathcal{W}'$ are two superoptimal weights, then $\lambda_a(\mathcal{W}) = \lambda_a(\mathcal{W}')$ for any $a \ge t_{d-1}$.*

Proof. Note that Φ is a symbol of Γ dominated by the weight $\Phi^*\Phi$. Suppose that $\Phi^*\Phi$ is not a superoptimal weight, i.e. that there exists an admissible weight $\mathcal{W}$ such that for some j_0, $0 \le j_0 \le d-1$,

$$s_{j_0}^\infty(\mathcal{W}) < s_{j_0}^\infty(\Phi^*\Phi), \quad s_j^\infty(\mathcal{W}) = s_j^\infty(\Phi^*\Phi), \quad 0 \le j \le j_0.$$

Let Ψ be a symbol of Γ dominated by the weight $\mathcal{W}$. Then

$$s_{j_0}^\infty(\Psi) < s_{j_0}^\infty(\Phi), \quad s_j^\infty(\Psi) = s_j^\infty(\Phi), \quad 0 \le j \le j_0,$$

which contradicts the fact that Φ is the superoptimal symbol of Γ. Therefore $\Phi^*\Phi$ is a superoptimal weight.

Let now $\mathcal{W}$ be a superoptimal weight, and let Ψ be a symbol of Γ dominated by $\mathcal{W}$. Then Ψ lexicographically minimizes $(s_0^\infty(\Psi),\ s_1^\infty(\Psi), \ldots s_{d-1}^\infty(\Psi))$ and so Ψ

coincides with the superoptimal symbol Φ. So, for any superoptimal weight $\mathcal{W}$ the superoptimal symbol Φ is the unique symbol of Γ dominated by $\mathcal{W}$. This means that $\Phi^*\Phi \leq \mathcal{W}$ for any superoptimal weight $\mathcal{W}$. Together with the equalities $s_j^\infty(\mathcal{W}) = s_j^\infty(\Phi^*\Phi) = t_j^2$ this implies the second part of the lemma. ■

Denote by Λ_a the function defined by

$$\Lambda_a(t) \stackrel{\text{def}}{=} \begin{cases} t, & t \geq a \\ a, & t < a \end{cases} \tag{8.1}$$

The following fact is an easy consequence of Lemma 8.2.

Corollary 8.3. *Let $\mathcal{W}$ and $\mathcal{W}'$ be two superoptimal weights. Then $\Lambda_a(\mathcal{W}) = \Lambda_a(\mathcal{W}')$ for any $a \geq t_{d-1}$.*

It is easy to see that if $a = t_{d-1}$ and $\mathcal{W}$ is a superoptimal weight, the weight $\Lambda_a(\mathcal{W})$ is the (unique) maximal superoptimal weight.

9. Invariance of indices

The main result of this section shows that the sum of the indices of a thematic factorization of a superoptimal symbol does not depend on the choice of a factorization. To prove this fact we shall use the same construction which was used in Section 6 to prove Theorem 6.3.

Let

$$a_0 > a_1 > \cdots > a_l$$

be all the distinct nonzero superoptimal singular values of a four block operator Γ which satisfies the hypotheses of Theorem 2.1. Let Φ be the superoptimal symbol of Γ and let k_j be the indices of a thematic factorization of Φ of the form (7.1). Consider the sum of the indices that correspond to equal superoptimal singular values:

$$\nu_r = \sum_{\{j: t_j = a_r\}} k_j, \quad 0 \leq r \leq l.$$

The following theorem is the main result of the section.

Theorem 9.1. *The numbers ν_r do not depend on the choice of thematic factorization of Φ.*

We are going to deduce Theorem 9.1 from Theorem 9.3 below, which describes the numbers ν_r in terms of a superoptimal weight $\mathcal{W}$.

We say that a nonzero function $\xi \in H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2})$ is a *maximizing vector for an admissible weight* $\mathcal{W}$ if

$$\|\Gamma\xi\|^2 = (\mathcal{W}\xi, \xi).$$

Lemma 9.2. *Let Γ be a four block operator, $\mathcal{W}$ an admissible weight for Γ, and Φ a symbol of Γ dominated by $\mathcal{W}$. Let ξ be a maximizing vector for $\mathcal{W}$. Then $\Phi\xi \in H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$, i.e., $\Gamma_\Phi\xi = \Phi\xi$.*

Note that for $\mathcal{W} \equiv cI$, $c \in \mathbb{R}_+$, this was proved in Lemma 3.2.

Proof. We have

$$(\mathcal{W}\xi, \xi) = \|\Gamma_\Phi\xi\|_2^2 = \|\mathbb{P}^-\Phi\xi\|_2^2 \le \|\Phi\xi_2\|^2 \le (\mathcal{W}\xi, \xi)\,.$$

It follows that $\|\mathbb{P}^-\Phi\xi\|_2^2 \le \|\Phi\xi_2\|^2$, which implies the result. ■

Given an admissible weight $\mathcal{W}$ put

$$\mathcal{E}(\mathcal{W}) = \{\xi \in H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2}) : \ \xi \text{ is maximizing for } \mathcal{W} \text{ or } \xi = \mathbb{O}\}.$$

It is easy to see that $\xi \in \mathcal{E}(\mathcal{W})$ if and only if $\mathbb{P}^+\mathcal{W}\xi - \Gamma_\Phi^*\Gamma_\Phi\xi = \mathbb{O}$, where $\mathbb{P}^+$ is the orthogonal projection onto $H^2(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2})$. Therefore $\mathcal{E}(\mathcal{W})$ is a closed linear subspace. Recall that $\Lambda_a(\mathcal{W})$ does not depend on the choice of superoptimal weight, where Λ_a is defined in (8.1).

Theorem 9.3. *Let Γ be a four block operator satisfying the hypotheses of Theorem 2.1, $\mathcal{W}$ a superoptimal admissible weight, and Φ the superoptimal symbol of Γ. Consider a thematic factorization of Φ of the form (7.1). Let k_j be the indices of the factorization. Then for $a \ge a_l$,*

$$\sum_{\{j: t_j \ge a\}} k_j = \dim \mathcal{E}(\Lambda_a(\mathcal{W}))\,. \tag{9.1}$$

Let us first deduce Theorem 9.1 from Theorem 9.3.

Proof of theorem 9.1. It follows immediately from (9.1) that

$$\nu_0 = \dim \mathcal{E}(\Lambda_{a_0}(\mathcal{W})), \quad \nu_j = \dim \mathcal{E}(\Lambda_{a_j}(\mathcal{W})) \ominus \mathcal{E}(\Lambda_{a_{j-1}}(\mathcal{W})), \ \ 1 \le j \le l,$$

which proves the result. ■

Proof of Theorem 9.3. It is easy to see that $\mathcal{E}(\Lambda_a(\mathcal{W}))$ is constant on $(a_{j+1}, a_j]$. So it is sufficient to prove that for $0 \le s \le l$

$$\sum_{\{j: t_j \ge a_s\}} k_j = \dim \mathcal{E}(\mathcal{W}_s)\,,$$

where $\mathcal{W}_s = \Lambda_{a_s}(\mathcal{W})$.

Let us prove the theorem by induction on d.

If $d = 1$, factorization (7.1) has the form

$$\Phi = W_0^* D V_0^*,$$

where

$$D = \left(\begin{array}{c|c} t_0 u_0 & \\ \mathbb{O} & \\ \vdots & \mathbb{O} \\ \mathbb{O} & \\ \hline \mathbb{O} & \star \end{array}\right),$$

and $\| \star \| < t_0$ (otherwise the essential norm of Γ_Φ would not be less than t_0). Clearly, $\mathcal{W}_0(\zeta) \equiv a_0 I$. If ξ is a maximizing vector for $\mathcal{W}_0$, then it is easy to see that only the first entry of $V_0^*\xi$ is nonzero. Therefore $\xi(\zeta)$ is pointwise orthogonal to all columns of V except for the first one. It follows that $\xi = \chi v$, where χ is a scalar function in L^2. Using the fact that v_1 is a co-outer column function, one can easily deduce that $\chi \in H^2$. It is easy to see that

$$\Phi\xi = t_0 u_0 \chi \overline{w}.$$

Since ξ is a maximizing vector, it follows that $\Phi\xi \in H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$. We can now use the fact that w_1 is a co-outer column function to deduce that $u_0\chi \in H^2_-$ which means that $\chi \in \operatorname{Ker} T_{u_0}$.

Conversely, it is easy to see that if $\chi \in \operatorname{Ker} T_{u_0}$, then $\xi = \chi v$ is a maximizing vector, which proves that $\dim \mathcal{E}(\mathcal{W}_0) = \dim \operatorname{Ker} T_{u_0} = k_0$.

Suppose now that the theorem is proved for $d - 1$. We have

$$\Phi = W^* \begin{pmatrix} t_0 u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{pmatrix} V^*,$$

where $\Phi^{(1)}$ is the superoptimal symbol of $\Gamma_{\Phi^{(1)}}$, $W \stackrel{\text{def}}{=} W_0$, and $V \stackrel{\text{def}}{=} V_0$. The induction hypothesis implies that the theorem holds for $\Gamma_{\Phi^{(1)}}$.

Let $0 \le s \le l$ and $a = a_s$. Suppose that $\mathcal{W}'$ is a superoptimal weight for $\Gamma_{\Phi^{(1)}}$. By the induction hypothesis

$$\dim \mathcal{E}(\Lambda_a(\mathcal{W}')) = N \stackrel{\text{def}}{=} \sum_{\{j \ge 1: t_j \ge a\}} k_j,$$

where the k_j are the indices of the thematic factorization (7.1). By Lemma 9.2, $\xi \in E(\Lambda_a(\mathcal{W}'))$ if and only if

$$\Phi^{(1)}\xi \in H^2_-(\mathbb{C}^{m_1-1}) \oplus L^2(\mathbb{C}^{m_2}) \qquad \text{and} \qquad \|\Phi^{(1)}\xi\|^2 = (\Lambda_a(\mathcal{W}')\xi, \xi).$$

Let $\xi_1, \xi_2, \cdots, \xi_N$ be a basis in $\mathcal{E}(\Lambda_a(\mathcal{W}'))$ and let $\eta_\iota = \Gamma_{\Phi^{(1)}}\xi_\iota$. By Lemma 6.2 there exist scalar functions χ_ι, $1 \le \iota \le N$, such that

$$W^* \begin{pmatrix} \chi_\iota \\ \eta_\iota \end{pmatrix} \in H^2_-(\mathbb{C}^{m_1-1}) \oplus L^2(\mathbb{C}^{m_2}).$$

As in the proof of Theorem 6.3 we define the functions $\xi_\iota^\#$ as

$$\xi_\iota^\# = A^t \xi_\iota + q_\iota v,$$

where q_ι is a scalar functions in H^2 satisfying

$$\mathbb{P}_+(t_0 u_0 q_\iota + t_0 u_0 v^* A^t \xi_\iota) = \chi_\iota$$

(recall that the matrix function A is defined in after Lemma 6.1). We have

$$\eta_\iota^\# \stackrel{\mathrm{def}}{=} \mathbb{P}^- \Phi \xi_\iota^\# = \mathbb{P}^- W^* \begin{pmatrix} t_0 u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{pmatrix} V^* \xi_\iota^\# = \mathbb{P}^- W^* \begin{pmatrix} \chi_\iota + \omega_\iota \\ \eta_\iota \end{pmatrix},$$

where as in the proof of Theorem 6.3

$$\omega_\iota = \mathbb{P}_-(t_0 u_0 q_\iota + t_0 u_0 v^* A^t \xi_\iota).$$

As we have explained in the proof of Theorem 6.3

$$\mathbb{P}^- W^* \begin{pmatrix} \chi_\iota + \omega_\iota \\ \eta_\iota \end{pmatrix} = W^* \begin{pmatrix} \chi_\iota + \omega_\iota \\ \eta_\iota \end{pmatrix}$$

and so $\eta_\iota^\# = \Phi \xi_\iota^\#$.

Since the matrix function W is unitary-valued, we have

$$\begin{aligned} \|\Phi \xi_\iota^\#\|^2 &= \left\| W^* \begin{pmatrix} t_0 u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{pmatrix} \begin{pmatrix} q_\iota + v^* A^t \xi_\iota \\ \xi_\iota \end{pmatrix} \right\|^2 \\ &= t_0^2 \|u_0(q_\iota + v^* A^t \xi_\iota)\|^2 + \|\Phi^{(1)} \xi_\iota\|^2 \\ &= t_0^2 \|q_\iota + v^* A^t \xi_\iota\|^2 + (\Lambda_a(\mathcal{W}')\xi_\iota, \xi_\iota) \end{aligned}$$

(the last equality holds because $\xi_\iota \in \mathcal{E}(\Lambda_a(\mathcal{W}'))$, where $\mathcal{W}' = (\Phi^{(1)})^* \Phi^{(1)}$ is a superoptimal weight for $\Gamma_{\Phi^{(1)}}$).

Consider the weight $\mathcal{V}$,

$$\mathcal{V} = \begin{pmatrix} t_0^2 & \mathbb{O} \\ \mathbb{O} & \mathcal{W}' \end{pmatrix}.$$

Bearing in mind that

$$V^*\xi_\iota^\# = \begin{pmatrix} q_\iota + v^* A^t \xi_\iota \\ \xi_\iota \end{pmatrix},$$

we can continue the above chain of inequalities:

$$\|\Phi\xi_\iota^\#\|^2 = (\Lambda_a(\mathcal{V})V^*\xi_\iota^\#, V^*\xi_\iota^\#) = (\Lambda_a(\mathcal{W})\xi_\iota^\#, \xi_\iota^\#) \tag{9.2}$$

(the last equality holds because V is unitary-valued). Since $\Gamma_\Phi \xi_\iota^\# = \mathbb{P}^- \Phi \xi_\iota^\# = \Phi \xi_\iota^\#$, it follows from (9.2) that $\xi_\iota^\# \in \mathcal{E}(\Lambda_a(\mathcal{W}))$.

We can add now another k_0 linear independent vectors of $\mathcal{E}(\Lambda_a(\mathcal{W}))$. Let $x_1, \cdots, x_{k_0}$ be a basis of $\operatorname{Ker} T_{u_0}$. Obviously, $x_\iota v \in \mathcal{E}(\Lambda_a(\mathcal{W}))$. Let us show that the vectors $\xi_1^\#, \cdots, \xi_N^\#$, $x_1 v, \cdots, x_{k_0} v$ are linearly independent. It is sufficient to prove that if $x \in \operatorname{Ker} T_{u_0}$ and $xv + \sum_{\iota=1}^N c_\iota \xi_\iota^\# = \mathbb{O}$, then $x = \mathbb{O}$ and $c_\iota = 0$, $1 \le \iota \le N$. We have

$$V^*(xv + \sum_{\iota=1}^N c_\iota \xi_\iota^\#) = \begin{pmatrix} x \\ \mathbb{O} \end{pmatrix} + \sum_{\iota=1}^N c_\iota \begin{pmatrix} v^* A^t \xi_\iota + q_\iota \\ \xi_\iota \end{pmatrix} = \mathbb{O}. \tag{9.3}$$

Since the ξ_ι are linearly independent, it follows that $c_\iota = 0$, $1 \le \iota \le N$, which in turn implies that $x = \mathbb{O}$.

This proves that

$$\sum_{\{j: t_j \ge a\}} k_j \le \dim \mathcal{E}(\Lambda_a(\mathcal{W})).$$

Let us prove the opposite inequality.

Denote by $\mathcal{E}_0$ the set of vectors in $\mathcal{E}(\Lambda_a(\mathcal{W}))$ of the form xv such that x is a scalar function in H^2. It is easy to see that $xv \in \mathcal{E}_0$ if and only if $x \in \operatorname{Ker} T_{u_0}$. It remains to show that there exists at most $\sum_{\{j>0: t_j \ge a\}} k_j$ vectors $\breve{\xi}_\iota$ that are linearly independent modulo $\mathcal{E}_0$. Let $\breve{\eta}_\iota \stackrel{\text{def}}{=} \Gamma_\Phi \breve{\xi}_\iota$. By Lemma 9.2, $\breve{\eta}_\iota = \Phi \breve{\xi}_\iota$. Put

$$V^*\breve{\xi}_\iota = \begin{pmatrix} \gamma_j \\ \xi_\iota \end{pmatrix}, \qquad W^*\breve{\eta}_\iota = \begin{pmatrix} \delta_\iota \\ \eta_\iota \end{pmatrix},$$

where γ_ι, δ_ι are scalar functions in L^2. Since the vectors $\breve{\xi}_\iota$ are linearly independent modulo $\mathcal{E}_0$, the vectors ξ_ι are linearly independent. To complete the proof, it is sufficient to show that $\xi_\iota \in \mathcal{E}(\Lambda_a(\mathcal{W}'))$.

Since $\breve{\eta}_\iota = \Phi \breve{\xi}_\iota$, we have that $\eta_\iota = \Phi^{(1)} \xi_\iota$ and $\delta_\iota = t_0 u_0 \gamma_\iota$. It follows from the block structure of V and W that $\xi_\iota \in H^2(\mathbb{C}^{n_1 - 1}) \oplus L^2(\mathbb{C}^{n_2})$ and $\eta_\iota \in H^2_-(\mathbb{C}^{m_1-1}) \oplus L^2(\mathbb{C}^{m_2})$. So $\eta_\iota = \Phi^{(1)} \xi_\iota = \Gamma_{\Phi^{(1)}} \xi_\iota$.

To show that $\xi_\iota \in \mathcal{E}(\Lambda_a(\mathcal{W}'))$, consider the following chain of equalities

$$\begin{aligned}\left(\Lambda_a(\mathcal{W})\breve{\xi}_\iota, \breve{\xi}_\iota\right) &= \left(\Lambda_a(V\mathcal{W}V^*)V\breve{\xi}_\iota, V\breve{\xi}_\iota\right) \\ &= \left(\left(\begin{array}{cc} t_0^2 & \mathbb{O} \\ \mathbb{O} & \Lambda_a(\mathcal{W}') \end{array}\right)\left(\begin{array}{c} \gamma_\iota \\ \xi_\iota \end{array}\right), \left(\begin{array}{c} \gamma_\iota \\ \xi_r \end{array}\right)\right) \\ &= t_0^2\|\gamma_\iota\|^2 + \left(\Lambda_a(\mathcal{W}')\xi_\iota, \xi_\iota\right).\end{aligned}$$

On the other hand

$$\left(\Lambda_a(\mathcal{W})\breve{\xi}_\iota, \breve{\xi}_\iota\right) = \|\Phi\breve{\xi}_\iota\| = \|\breve{\eta}_\iota\| = \|\eta_\iota\|^2 + \|\delta_\iota\|^2 = \|\Phi^{(1)}\xi_\iota\|^2 + t_0^2\|\gamma_\iota\|^2.$$

Therefore $(\Lambda_a(\mathcal{W}')\xi_\iota, \xi_\iota) = \|\Phi^{(1)}\xi_\iota\|^2$, which implies $\xi_\iota \in \Lambda_a(\mathcal{W}')$. ■

10. Singular values of $\mathbf{\Gamma_\Phi}$ and superoptimal singular values

Let Γ be a four block operator satisfying the hypotheses of Theorem 2.1. Denote by Φ its unique superoptimal symbol and consider a thematic factorization of Φ of the form (7.1). Let $\{t_j\}$ be the superoptimal singular values and $\{k_j\}$ the indices of the factorization. Consider the extended t-sequence for Γ:

$$t_0, t_0, \cdots, t_1, t_1, \cdots, t_1, \cdots$$

in which t_j is repeated k_j times. We denote the terms of the extended sequence by $t'_0, t'_1, t'_2, \cdots$. Although the indices k_j depend on the choice of thematic factorization, it follows from Theorem 9.1 that the extended t-sequence is uniquely determined by Γ.

In [PY2] it was shown in the case of Nehari's problem with $\Phi \in H^\infty + C$ that $t'_j \le s_j(H_\Phi)$, $0 \le j \le k-1$. In this section we are going to prove the same inequality in the case of the four block problem under the hypotheses of Theorem 2.1. Moreover, we prove in this section a stronger result which is also new in the case of Nehari's problem with an $H^\infty + C$ symbol. To prove the results we use in this section the same machinery as we used in Section 6.

Let Γ be a four block operator that satisfies the hypotheses of Theorem 2.1. Then (see Section 3)

$$\Phi = W^*\left(\begin{array}{cc} t_0 u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{array}\right)V^*,$$

where the unitary-valued matrix functions V and W are defined in Section 3. The following inequality is the main result of the section.

Theorem 10.1. *Let Γ be a four block operator such that $\|\Gamma_\Phi\|_{\mathrm{e}} < \|\Gamma_\Phi\|$ and let Φ be its superoptimal symbol. Then*

$$s_j(\Gamma_{\Phi^{(1)}}) \le s_{j+k_0}(\Gamma_\Phi), \quad j \in \mathbb{Z}_+.$$

Recall that $k_0 = \dim \operatorname{Ker} T_{u_0}$.

Let us first derive from Theorem 10.1 the desired inequality between the singular values of Γ_Φ and the superoptimal singular values.

Theorem 10.2. *Under the hypotheses of Theorem 2.1*

$$t'_j \le s_j(\Gamma_\Phi), \quad j \ge 0.$$

Proof of Theorem 10.2. Let $x \in \operatorname{Ker} T_{u_0}$. Clearly,

$$\Gamma_\Phi x v = W^* \begin{pmatrix} t_0 u_0 x & \mathbb{O} & \cdots & \mathbb{O} \end{pmatrix}^t = t_0 u_0 x \overline{w} \in H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2}).$$

It follows that $\|\Gamma_\Phi x v\|_2 = t_0 \|xv\|_2$ which proves that

$$s_j(\Gamma_\Phi) = t_0, \quad 0 \le j \le k_0 - 1. \tag{10.1}$$

We can now proceed by induction on d. Clearly, the result holds for $d = 1$. It is also obvious that if the theorem holds for $\Gamma_{\Phi^{(1)}}$, then by Theorem 10.1

$$t'_j \le s_j(\Gamma_\Phi), \quad k_0 \le j \le d-1,$$

which together with (10.1) proves the theorem. ■

Proof of Theorem 10.1. Clearly, it is sufficient to prove the following fact. Let $\mathcal{L}$ be a subspace of $H^2_-(\mathbb{C}^{n_1-1}) \oplus L^2(\mathbb{C}^{n_2})$ such that $\|\Gamma_{\Phi^{(1)}}\xi\|_2 \ge s\|\xi\|_2$, for every $\xi \in \mathcal{L}$, where $0 < s \le t_0$, then there exists a subspace $\mathcal{M}$ of $H^2_-(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2})$ such that $\dim \mathcal{M} \ge \dim \mathcal{L} + k_0$ and $\|\Gamma_\Phi \rho\|_2 \ge s\|\rho\|_2$ for every $\rho \in \mathcal{M}$.

Let ξ_ι, $1 \le \iota \le N$, be a basis in $\mathcal{L}$. Put $\eta_\iota = \Gamma_{\Phi^{(1)}}\xi_\iota$. By Lemma 6.2 there exist scalar functions χ_ι in H^2 such that $W^* \begin{pmatrix} \chi_\iota \\ \eta_\iota \end{pmatrix} \in H^2_-(\mathbb{C}^{m_1}) \oplus L^2(\mathbb{C}^{m_2})$. We define the functions $\xi^\#_\iota \in H^2_-(\mathbb{C}^{n_1}) \oplus L^2(\mathbb{C}^{n_2})$ by

$$\xi^\#_\iota = A^t \xi_\iota + q_\iota v,$$

where q_ι is a scalar function in H^2 satisfying

$$\mathbb{P}_+(t_0 u_0 q_\iota + t_0 u_0 v^* A^t \xi_\iota) = \chi_\iota$$

(see the proof of Theorem 6.3).

We can now define $\mathcal{M}$ by

$$\mathcal{M} = \operatorname{span}\{\xi_\iota^\# + xv : \ 1 \le \iota \le N, \ x \in \operatorname{Ker} T_{u_0}\}.$$

Let us show that $\dim \mathcal{M} = N + k_0$. Since $\dim \operatorname{Ker} T_{u_0} = k_0$, it is sufficient to prove that if $xv + \sum_{\iota=1}^{N} c_\iota \xi_\iota^\# = \mathbb{O}$, then $x = \mathbb{O}$ and $c_\iota = 0$, $1 \le \iota \le N$. This follows immediately from (9.3).

To complete the proof it remains to show that $\|\Gamma_\Phi \rho\|_2 \ \ge \ s\|\rho\|_2$ for $\rho \ = \ xv + \sum_{\iota=1}^{r} c_j \xi_\iota^\#$. Let $\xi \ = \ \sum_{\iota=1}^{r} c_\iota \xi_\iota$, $\eta \ = \ \Gamma_\Phi^{(1)} \xi$, $q \ = \ \sum_{\iota=1}^{r} c_\iota q_\iota$, and $\xi_\iota^\# = \sum_{\iota=1}^{r} c_\iota \xi_\iota^\#$.

We have

$$\begin{aligned}
W^* \begin{pmatrix} t_0 u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{pmatrix} \rho &= W^* \begin{pmatrix} t_0 u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{pmatrix} (xv + qv + A^t \xi) \\
&= W^* \begin{pmatrix} t_0 u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{pmatrix} \begin{pmatrix} x + q + v^* A^t \xi \\ \xi \end{pmatrix} \\
&= W^* \begin{pmatrix} t_0 u_0 x + t_0 u_0 q + t_0 u_0 v^* A^t \xi \\ \Phi^{(1)} \xi \end{pmatrix}.
\end{aligned}$$

It follows (see the proof of Theorem 6.3) that

$$\Gamma_\Phi \rho = W^* \begin{pmatrix} t_0 u_0 x + t_0 u_0 q + t_0 u_0 v^* A^t \xi \\ \Gamma_{\Phi^{(1)}} \xi \end{pmatrix}.$$

Therefore

$$\|\Gamma_\Phi \rho\|_2^2 = |t_0|^2 \|x + q + v^* A^t \xi\|_2^2 + \|\eta\|_2^2.$$

We have

$$\|\rho\|_2^2 = \|V^* \rho\|_2^2 = \|x + q + v^* A^t \xi\|_2^2 + \|\xi\|_2^2.$$

Since $s \le t_0$ and $\|\eta\|_2 \ge s\|\xi\|_2$, it follows that $\|\Gamma_\Phi \rho\|_2^2 \ge s^2 \|\rho\|_2^2$. ■

Theorem 10.1 certainly applies to the case of Nehari's problem. Recall that a matrix function Φ is called *very badly approximable* (see [PY]) if the zero function is a superoptimal approximant of Φ.

Recall that under the condition $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$ the function Φ admits a factorization

$$\Phi = W^* \begin{pmatrix} t_0 u_0 & \mathbb{O} \\ \mathbb{O} & \Phi^{(1)} \end{pmatrix} V^*, \tag{10.2}$$

where V and W are unitary matrix functions of the form

$$V = \begin{pmatrix} v & \overline{V}_c \end{pmatrix}, \quad W = \begin{pmatrix} w & \overline{W}_c \end{pmatrix}^t,$$

and u_0 is a unimodular function such that $k_0 \stackrel{\text{def}}{=} \dim \operatorname{Ker} T_{u_0} > 0$.

The following result is certainly a partial case of Theorem 10.1.

Theorem 10.3. *Let Φ be a very badly approximable matrix function on $\mathbb{T}$ such that $\|H_\Phi\|_{\mathrm{e}}$ is less that the smallest nonzero superoptimal singular value of Nehari's problem. Then*

$$s_j(H_{\Phi^{(1)}}) \le s_{j+k_0}(H_\Phi), \quad j \in \mathbb{Z}_+,$$

where $\Phi^{(1)}$ and k_0 are given by the factorization (10.2).

References

[AAK1] V.M. Adamyan, D.Z. Arov, and M.G. Krein, *On infinite Hankel matrices and generalized problems of Carathéodory-Fejér and F. Riesz,* Funktsional. Anal. i Prilozhen. **2:1** (1968), 1-19.

[AAK2] V.M. Adamyan, D.Z. Arov, and M.G. Krein, *On infinite Hankel matrices and generalized problems of Carathéodory-Fejér and I,Schur*, Funktsional. Anal. i Prilozhen. **2:2** (1968), 1-17.

[AAK3] V.M. Adamyan, D.Z. Arov, and M.G. Krein, *Infinite Hankel block matrices and some related continuation problems*, Izv. Akad. Nauk Armyan. SSR Ser. Mat. **6** (1971), 87-112.

[Ar] W.B. Arveson, *Interpolation problems in nest algebras*, J. Funct. Anal. **20** (1975), 208-233.

[CJ] L. Carleson and S. Jacobs, *Best uniform approximation by analytic functions*, Ark. Mat. bf 10 (1972), 219-229.

[FT] C. Foias and A. Tannenbaum, *On the four block problem, II: the singular system*, Int. Equat. Oper. Th. **11** (1988), 726-767.

[Fr] B.A. Francis, *A course in H^∞ control theory*, Lecture Notes in Control and Information Sciences **88**, Springer-Verlag, Berlin, 1986.

[H] H. Helson, *Lectures on invariant subspaces*, Acad. Press, NY, 1964.

[Kh] S. Khavinson, *On some extremal problems of the theory of analytic functions*, Uchen. Zapiski Mosk. Universiteta, Matem. **144**:4 (1951), 133-143. English transl.: Amer Math. Soc. Translations (2) **32** (1963), 139-154.

[Ne] Z. Nehari, *On bounded bilinear forms*, Ann. Math. **65** (1957), 153-162.

[Ni] N.K. Nikol'skii, *Treatise on the shift operator*, Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1986.

[Pa] L.B. Page, *Bounded and compact vectorial Hankel operators,* Trans. Amer. Math. Soc. **64** (1970),529-539.

[P] V.V. Peller, *Approximation by analytic operator-valued functions,* Preprint, 1994.

[PKh] V.V. Peller and S.V. Khrushchev, *Hankel operators, best approximation and stationary Gaussian processes*, Uspekhi Mat. Nauk **37:1** (1982), 53-124. English Transl.: Russian Math. Surveys **37:1** (1982), 61-144.

[PY1] V.V. Peller and N.J. Young, *Superoptimal analytic approximations of matrix functions*, J. Funct. Anal. **120** (1994), 300-343.

[PY2] V.V. Peller and N.J. Young, *Superoptimal singular values and indices of matrix functions*, Int. Eq. Op. Theory

[RSh] W. Rogosinski and H.S. Shapiro, *On certain extremum problems for analytic functions*, Acta Math. **90** (1953), 287-318.

[R] Yu.A. Rozanov, *Stationary stochastic processes*, Fizmatgiz, Moscow, 1963. English transl.: Golden-Day, San Francisco, 1967.

[SNF1] B. Sz.-Nagy and C. Foias, *Analyse harmonique des opérateurs de l'espace de Hilbert*, Akadémiai Kiadó, Budapest, 1967.

[SNF2] B. Sz.-Nagy and C. Foias, *On contractions similar to isometries and Toeplitz operators*, Ann. Acad. Sci.. Fenn. A I **2** (1976), 553-564.

[T] S.R. Treil, *On superoptimal approximation by analytic and meromorphic matrix-valued functions*, to appear in J. Funct. Anal.

[TV] S.R. Treil and A.L. Volberg, *A fixed point approach to Nehari's problem and its applications*, Oper. Theory: Adv. Appl., Birkhäuser, **71** (1994), 165-186.

[Y] N.J. Young, *The Nevanlinna-Pick problem for matrix-valued functions*, J. Oper. Theory, **15** (1986), 239-265.

[Va] V. I. Vasyunin, *Formula for multiplicity of contractions with finite defect indices*, Oper. Theory: Adv. Appl., Birkhäuser, **4** (1989), 281-304.

V.V. Peller
Department of Mathematics
Kansas State University
Manhattan, Kansas 66506

S.R.Treil
Department of Mathematics
Michigan State University
East Lansing, Michigan 48824

AMS 1991 Subject Classification: 47B35, 93B36.