# Kummert's Approach to Realization on the Bidisk

## Greg Knese

ABSTRACT. We give a simplified exposition of Kummert's approach to proving that every matrix-valued rational inner function in two variables has a minimal unitary transfer function realization. A slight modification of the approach extends to rational functions which are isometric on the two-torus, and we use this to give a largely elementary new proof of the existence of Agler decompositions for every matrix-valued Schur function in two variables. We use a recent result of Dritschel to prove that two variable matrix-valued rational Schur functions always have finite-dimensional contractive transfer function realizations. Finally, we prove that two variable matrix-valued polynomial inner functions have transfer function realizations built out of special nilpotent linear combinations.

## Contents

## 1. INTRODUCTION

The goal of this paper is to give a simple proof and several applications of the following theorem.

**Theorem 1.1 (Main Theorem).** *Assume $S : \mathbb{D}^2 \to \mathbb{C}^{M \times N}$ is rational with no poles in $\mathbb{D}^2$ and satisfies $S^*S = I_N$ on $\mathbb{T}^2$ away from the zero set of the denominator of $S$.*

*Then, there exist an integer $r$ and an $(M + r) \times (N + r)$ isometric matrix $U = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ such that*

$$(1.1) \qquad S(z) = A + B\Delta(z)(I - D\Delta(z))^{-1}C$$

*where $\Delta(z_1, z_2) = z_1 P_1 + z_2 P_2$ and $P_1, P_2$ are orthogonal projections with $P_1 + P_2 = I_r$.*

Above $\mathbb{D}^2 = \{z = (z_1, z_2) \in \mathbb{C}^2 : |z_1|, |z_2| < 1\}$ is the unit bidisk, and $\mathbb{T}^2 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1| = |z_2| = 1\}$ is the two-torus (or bitorus). We shall call functions that satisfy the hypotheses of this theorem *rational iso-inner functions*. Formulas in the conclusion of this theorem (such as (1.1)), which are built out of block operators, will be called *transfer function realizations* (or TFRs). If the operator is a finite matrix we will call it a finite TFR, and if we have extra information about the operator involved, we will incorporate it into the terminology. For example, the above theorem asserts the existence of a "finite isometric TFR" for two variable rational iso-inner functions.

This theorem is due to Kummert in the square case $M = N$ [33]. Kummert's theorem was ahead of its time and its proof was both ingenious and largely elementary. At the same time, Kummert's argument *seems* complicated and the engineering terminology may obscure the underlying concepts for some, so one of our main goals is to give a simplified, conceptual, and entirely mathematical account of Kummert's approach. We also give an algorithm for constructing the matrix $U$. Motivation for doing so comes from recent interest in the wavelet community in transfer function formulas in one and several variables [14]. We have presented generalizations of our simplified argument in a couple of papers [21, 32], but the generalizations can also potentially obscure the underlying concepts. A minor adjustment allows us to treat the non-square case $M \neq N$, which in turn allows us to give possibly the most elementary and direct proof of the following seminal theorem of Agler.

**Theorem 1.2 (Agler** [1, 2]**).** *Let $f : \mathbb{D}^2 \to \mathbb{C}^{M \times N}$ be holomorphic, and $\|f(z)\| \leq 1$ for all $z \in \mathbb{D}^2$. Then, $f$ has a contractive TFR: there exists a contractive operator $T$ on some Hilbert space with block decomposition $T = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ such that*

$$f(z) = A + B\Delta(z)(I - D\Delta(z))^{-1}C$$

*where $\Delta(z) = z_1 P_1 + z_2 P_2$ and $P_1, P_2$ are pairwise orthogonal projections which sum to the identity on the domain of $D$.*

Perhaps the most important application of this theorem is a Pick interpolation theorem for holomorphic functions on the bidisk. For this and other applications, we refer the reader to the book [4] and the papers [3, 5, 6, 10].

Dritschel has recently proven a strong Fejér-Riesz type of result in two variables (Theorem 6.7) which makes it possible to prove that every two-variable rational function bounded by one in norm on $\mathbb{D}^2$ (with no assumptions on boundary behavior) has a *finite contractive* TFR.

**Theorem 1.3.** *Let* $S : \mathbb{D}^2 \to \mathbb{C}^{M \times N}$ *be rational with no poles in* $\mathbb{D}^2$ *and assume* $\|S(z)\| \leq 1$ *for all* $z \in \mathbb{D}^2$. *Then, there exists a contractive matrix* $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *such that*

$$S(z) = A + B\Delta(z)(I - D\Delta(z))^{-1}C$$

*where* $\Delta(z_1, z_2) = z_1 P_1 + z_2 P_2$, $P_1, P_2$ *are orthogonal projections with* $P_1 + P_2 = I$.

A very important bonus of Kummert's approach is that it constructs the matrix $U$ in Theorem 1.1 with the minimal possible dimensions in a strong way. For a rational iso-inner function $S : \mathbb{D}^2 \to \mathbb{C}^{M \times N}$ we can always make sense of $z_1 \mapsto S(z_1, z_2)$ for each fixed $z_2 \in \mathbb{T}$, and this is a one variable rational iso-inner function (Lemma 4.3). If we have a formula as in Theorem 1.1 where the ranks of $P_1, P_2$ are $r_1, r_2$, then we can construct a transfer function realization for $S(\cdot, z_2)$ with size $r_1$ and a transfer function realization for $S(z_1, \cdot)$ with size $r_2$. In the square case $M = N$, this can be done optimally.

**Theorem 1.4 (Kummert's minimality theorem).** *Suppose* $S : \mathbb{D}^2 \to \mathbb{C}^{N \times N}$ *is rational and inner. Then, one can choose* $U$ *in Theorem 1.1 so that the ranks* $r_1, r_2$ *of* $P_1, P_2$ *are simultaneously minimal:* $r_1$ *is the maximum of the minimal size of a unitary* TFR *for* $z_1 \mapsto S(z_1, z_2)$ *where* $z_2$ *varies over* $\mathbb{T}$ *and* $r_2$ *is the maximum of the minimal size of a unitary* TFR *for* $z_2 \mapsto S(z_1, z_2)$ *where* $z_1$ *varies over* $\mathbb{T}$.

In particular, among all possible unitary TFRs for $S$, neither $r_1$ nor $r_2$ can be smaller than those in Kummert's construction. We will give a conceptual proof of Kummert's minimality theorem, and clarify why this is the best possible result. Before the mathematical community knew of Kummert's results, this result was reproven in the scalar case using the framework of Geronimo-Woerdeman [20] in [30]. Later, Theorem 1.4 was also proven using Hilbert space methods in [12]. The scalar minimality theorem was crucial in giving a characterization of two-variable rational matrix-monotone functions in [5]. It is also useful in proving determinantal representations for certain families of polynomials $p \in \mathbb{C}[z_1, z_2]$ with no zeros in $\mathbb{D}^2$ [29].

We shall present a new application of the minimality theorem which has some relevance to the applications of this theory to wavelets in [13, 14]. In these papers, matrix-valued *polynomial* inner functions are of particular interest.

**Theorem 1.5.** *Let* $S \in \mathbb{C}^{N \times N}[z_1, z_2]$ *and assume* $S^*S = I_N$ *on* $\mathbb{T}^2$. *Then,* $U$ *in Theorem 1.1 can be chosen with* $\det(I - D\Delta(z)) \equiv 1$.

Note this means $D\Delta(z) = z_1 DP_1 + z_2 DP_2$ is nilpotent for every $z$.

**1.1. Guide to the reader.** This paper is structured so that it can hopefully be read by a broad audience. We make no mention of systems theory terminology (except for "transfer function") and we make no use of von Neumann inequalities and related operator theory originally used in the proof of Agler's theorem. (We do discuss some of this for context in Section 6.) Our first goal is to quickly and simply prove Kummert's Theorem 1.1, and explain how this proves Agler's theorem.

Some readers may be satisfied with this quick and mostly constructive approach to these results and can stop after Section 6. After that, we introduce the technicalities necessary to prove Kummert's minimality theorem and give an application to inner polynomials. We include an appendix with extra background.

## 2. FINITE-DIMENSIONAL TRANSFER FUNCTION REALIZATIONS

One of the fundamental things that Agler did in his original proof of Theorem 1.2 was connect TFRs to certain formulas now called *Agler decompositions* which involved positive semi-definite kernels. The following theorem establishes some basic equivalences about *finite* TFRs and *finite*-dimensional Agler decompositions which hold not just on $\mathbb{D}^2$ but on any polydisk $\mathbb{D}^d$. Note that "matrix" below always refers to a finite matrix.

**Theorem 2.1 (Equivalences Theorem).** *Let $S : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ be a function. The following are equivalent:*

(1)  *There exists a contractive matrix $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ such that*

$$S(z) = A + B\Delta(z)(I - D\Delta(z))^{-1}C$$

   *where $\Delta(z) = \sum_j z_j P_j$, for some pairwise orthogonal projections with $\sum_j P_j = I$.*

(2)  *There exist matrix functions $F_j$ and a constant contractive matrix $T$ such that*

$$T \begin{pmatrix} I \\ z_1 F_1(z) \\ \vdots \\ z_d F_d(z) \end{pmatrix} = \begin{pmatrix} S(z) \\ F_1(z) \\ \vdots \\ F_d(z) \end{pmatrix}.$$

(3)  *There exist matrix functions $F_1, \ldots, F_d, G$ such that*

$$I - S(w)^* S(z) = G(w)^* G(z) + \sum_j (1 - \bar{w}_j z_j) F_j(w)^* F_j(z).$$

*We also have the following bonuses:*

(B1)  *Assuming (1)–(3), we have $S, F_1, \ldots, F_d, G$ are all rational and $\|S(z)\| \le 1$ for all $z \in \mathbb{D}^d$. If we assume at the outset that $S$ is holomorphic, then item (3) need only hold initially on an open set in order for it to hold globally.*

(B2)  *The T that works in (1) also works in (2).*

(B3)  *We also get equivalences if we replace "contractive" in (1) and (2) with "iso-metric" and G with 0 in (3). In this case, S is iso-inner and analytic outside the zeros of* $\det(I - D\Delta(z))$.

*Proof.* (2) $\implies$ (1). It helps to define

$$F(z) = \begin{pmatrix} F_1(z) \\ \vdots \\ F_d(z) \end{pmatrix}.$$

Let $P_j$ be the projection matrix for the block corresponding to $F_j$. Then, the equation in (2) can be written as

$$(2.1) \qquad \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & \Delta(z) \end{pmatrix} \begin{pmatrix} I \\ F(z) \end{pmatrix} = \begin{pmatrix} S(z) \\ F(z) \end{pmatrix}$$

for $\Delta(z) = \sum_j z_j P_j$. Block by block, this says

$$A + B\Delta F = S,$$
$$C + D\Delta F = F,$$

which yields $F = (I - D\Delta)^{-1}C$ and then $S = A + B\Delta(I - D\Delta)^{-1}C$.

(1) $\implies$ (2). We simply define $F = (I - D\Delta)^{-1}C$. Then, (2.1) holds because

$$C + D\Delta(I - D\Delta)^{-1}C = (I - D\Delta)^{-1}C.$$

(2) $\implies$ (3). The given equation implies

$$\begin{pmatrix} I \\ \Delta(w)F(w) \end{pmatrix}^* T^*T \begin{pmatrix} I \\ \Delta(z)F(z) \end{pmatrix} = \begin{pmatrix} S(w) \\ F(w) \end{pmatrix}^* \begin{pmatrix} S(z) \\ F(z) \end{pmatrix}.$$

Let $A = \sqrt{I - T^*T}$ and $G(z) = A \begin{pmatrix} I \\ \Delta(z)F(z) \end{pmatrix}$. Then,

$$\begin{pmatrix} I \\ \Delta(w)F(w) \end{pmatrix}^* \begin{pmatrix} I \\ \Delta(z)F(z) \end{pmatrix} = \begin{pmatrix} S(w) \\ F(w) \end{pmatrix}^* \begin{pmatrix} S(z) \\ F(z) \end{pmatrix} + G(w)^*G(z),$$

and this rearranges exactly into the equation in (3).

(3) $\implies$ (2). This is known as a lurking isometry argument. The map

$$\begin{pmatrix} I \\ \Delta(z)F(z) \end{pmatrix} \mapsto \begin{pmatrix} S(z) \\ F(z) \\ G(z) \end{pmatrix}$$

extends linearly and in a well-defined way to an isometric map from the span of the vectors on the left to the span of the vectors on the right as $z$ varies over $\mathbb{D}^d$. We can extend this to an isometric matrix $V$ satisfying

$$V\begin{pmatrix} I \\ \Delta(z)F(z) \end{pmatrix} = \begin{pmatrix} S(z) \\ F(z) \\ G(z) \end{pmatrix}$$

which we can compress to get a contractive matrix satisfying the equation in (2).

The bonus results follow. For (B1), $S$ is rational and bounded in operator norm by 1 by (1) and (3). The matrix functions $F_j, G$ are rational by the proofs of (2) $\Longrightarrow$ (1) and (2) $\Longrightarrow$ (3). If we assume $S$ is holomorphic and (3) only holds on an open set, then all of the proofs work on this restricted set but automatically extend holomorphically to $\mathbb{D}^d$ by the matrix formulas. Bonus (B2) follows from the proof of (1) $\Longleftrightarrow$ (2). For bonus (B3), notice that if $T$ is an isometric matrix, then we have $G = 0$ in the proof (2) $\Longrightarrow$ (3), and if we start with $G = 0$ we get $T$ to be isometric in the proof (3) $\Longrightarrow$ (2) since no compression is necessary. Finally, $S$ is iso-inner because we can insert $z = w \in \mathbb{T}^d$ into condition (3) to see $S^*S = I$ at least away from the zero set of $\det(I - D\Delta(z))$ which is a denominator for the $F_j$ and $S$ by the formula in (2) $\Longrightarrow$ (1). □

The next proposition says the conditions of Theorem 2.1 are also equivalent to $S$ being a submatrix of a rational inner function possessing a finite-dimensional unitary transfer function realization. Moreover, the various sizes of the transfer function realizations stay the same. To be more precise, let $r_j$ be the rank of $P_j$ in condition (1) of Theorem 2.1. Then, $r = (r_1, \ldots, r_d)$ will be called the *size breakdown* of the TFR. This terminology is endemic to this paper. The *size* of the TFR will refer to $|r| = r_1 + \cdots + r_d$. Note that $r_j$ also equals the number of rows of $F_j$ in conditions (2) and (3) of Theorem 2.1.

**Proposition 2.2.** *Let $S : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ be a function which has a finite contractive TFR with size breakdown $r$. Then, there exists $n \geq N, M$ and a matrix rational inner function $\Phi : \mathbb{D}^d \to \mathbb{C}^{n \times n}$ with finite unitary TFR with size breakdown $r$ such that $S$ is a submatrix of $\Phi$.*

*As a sort of converse, every submatrix of $S$ has a finite contractive TFR with same size breakdown.*

*Proof.* Suppose $S$ has a finite contractive TFR given via contractive $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Every contractive matrix is a submatrix of a finite unitary, say $U$. If we rearrange rows and columns we may write

$$U = \begin{pmatrix} A & A_{12} & B \\ A_{21} & A_{22} & B_2 \\ C & C_2 & D \end{pmatrix}.$$

If

$$\Phi(z) = \begin{pmatrix} A & A_{12} \\ A_{21} & A_{22} \end{pmatrix} + \begin{pmatrix} B \\ B_2 \end{pmatrix} \Delta(z)(I - D\Delta(z))^{-1} \begin{pmatrix} C & C_2, \end{pmatrix}$$

then $S(z) = (I \ O) \Phi(z) \begin{pmatrix} I \\ O \end{pmatrix}$.

This same type of observation shows that every submatrix of $S$ has a finite contractive *TFR*. $\qquad\square$

The following is referred to as the adjunction formula in [13].

**Proposition 2.3.** *Let* $S : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ *be a function with a finite contractive* TFR *given via a matrix* $T$ *as in* (1), (2) *of Theorem 2.1. Set* $\check{S}(z) = S(\bar{z})^*$. *Then,* $\check{S}$ *has a finite contractive* TFR *given via* $T^*$.

*In particular, if* $T$ *is isometric, then* $\check{S}$ *has a finite coisometric* TFR.

*Proof.* With $S(z) = A + B\Delta(z)(I - D\Delta(z))^{-1}C$, we have

$$\check{S}(z) = A^* + C^*(I - \Delta(z)D^*)^{-1}\Delta(z)B^*$$
$$= A^* + C^*\Delta(z)(I - D^*\Delta(z))^{-1}B^*,$$

which is exactly condition (1) of Theorem 2.1 with $T^*$ in place of $T$. $\qquad\square$

## 3. One-Variable Version of Theorem 1.1

We now prove a detailed one-variable version of the Main Theorem (Theorem 1.1). If $S = Q/p : \mathbb{D} \to \mathbb{C}^{M \times N}$ is a rational iso-inner function, then $S^*S = I$ on $\mathbb{T}$ away from zeros of $p$, but then $|p|^2I = Q^*Q$ on all of $\mathbb{T}$ by continuity.

**Theorem 3.1.** *Assume* $p \in \mathbb{C}[z]$ *has no zeros in* $\mathbb{D}$, $Q \in \mathbb{C}^{M \times N}[z]$, *and* $|p|^2I = Q^*Q$ *on* $\mathbb{T}$. *Let* $n$ *be the maximum of the degrees of* $p$ *and the entries of* $Q$. *Then,*

$$(3.1) \qquad K(w, z) = \frac{\overline{p(w)}p(z)I - Q(w)^*Q(z)}{1 - \bar{w}z}$$
$$= (I, \bar{w}I, \dots, \bar{w}^{n-1}I)T(I, zI, \dots, z^{n-1}I)^t,$$

*where* $T$ *is a positive semi-definite matrix whose entries can be expressed as polynomials in the coefficients of* $p, \bar{p}, Q, Q^*$. *Furthermore,* $K(w, z)$ *is a positive semi-definite kernel whose rank matches the rank of the matrix* $T$.

Positive semi-definite kernels are reviewed in Definition A.5, and the rank of such a kernel is defined in Definition A.6 in Appendix A.

The theorem allows for common zeros of $Q$ and $p$ which is important in using this result in two variables. It immediately follows that $S = Q/p$ possesses

an isometric TFR because we can factor $T = \mathbf{F}^*\mathbf{F}$ where $\mathbf{F}$ is an $r \times nN$ matrix. Then, for $F(z) = \mathbf{F}(I, zI, \ldots, z^{n-1}I)^t$, we have

$$I - S(w)^*S(z) = (1 - \bar{w}z)\left(\frac{F(w)}{p(w)}\right)^* \frac{F(z)}{p(z)}.$$

By Theorem 2.1 we see that $S$ has an isometric TFR. After the proof of Theorem 3.1, we give an explicit way to find a formula for an isometry $U$ out of which a TFR for $S$ can be built. We need a standard lemma to prove Theorem 3.1. We give the short proof in the appendix (see Subsection A.3).

**Lemma 3.2.** *Assume* $S : \mathbb{D} \to \mathbb{C}^{M \times N}$ *is analytic and* $\|S(z)\| \leq 1$ *in* $\mathbb{D}$. *Then, the kernel*

$$(3.2) \qquad\qquad K_S(w, z) = \frac{I - S(w)^*S(z)}{1 - \bar{w}z}$$

*is positive semi-definite.*

The swapping of $z, w$ is deliberate and is discussed in the proof in Appendix A.

*Proof of Theorem 3.1.* By analyticity, we have $\overline{p(1/\bar{z})}p(z)I = Q(1/\bar{z})^*Q(z)$ on $\mathbb{C} \setminus \{0\}$. This implies the polynomial in $z, \bar{w}$

$$\overline{p(w)}p(z)I - Q(w)^*Q(z)$$

is divisible by $(1 - \bar{w}z)$, and hence we can write (3.1) where $T$ is indeed a $nN \times nN$ matrix whose entries are polynomials in the coefficients of $p, \bar{p}, Q, Q^*$. We could solve for them but we do not need to. By Lemma 3.2, $K_S(w, z)$ in (3.2) is positive semi-definite. Multiplying through by $\overline{p(w)}p(z)$, we have that $K(w, z)$ as in (3.1) is a positive semi-definite matrix-valued polynomial function of bounded degree.

To show $T$ is positive semi-definite, take any $z_1, \ldots, z_n \in \mathbb{D}$ and note that

$$(K(z_i, z_j))_{i,j} = \begin{pmatrix} I & \bar{z}_1 I & \cdots & \bar{z}_1^{n-1}I \\ I & \bar{z}_2 I & \cdots & \bar{z}_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ I & \bar{z}_n I & \cdots & \bar{z}_n^{n-1}I \end{pmatrix} T \begin{pmatrix} I & I & \cdots & I \\ z_1 I & z_2 I & \vdots & z_n I \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{n-1}I & z_2^{n-1} & \cdots & z_n^{n-1}I \end{pmatrix} = V^*TV$$

is positive semi-definite where $V = (V_{i,j})$ is the block Vandermonde matrix $V_{i,j} = z_j^{i-1}I$. If the $z_j$ are all distinct then $V$ is invertible which implies that $T$ is positive semi-definite. The above computation also shows that the rank of $K$ equals the rank of $T$, although we omit some details. □

**Remark 3.3.** We now explain how to find an isometry $U$ out of which a TFR for $S = Q/p$ can be built. This will closely parallel our approach in the two-variable setting. We first factor $T = A^*A$ where $A$ is $r \times nN$ with $r = \text{rank}(T)$. Then, $A$ will possess a right inverse $B$, namely, $AB = I$. Set $F(z) = A(I, zI, \ldots, z^{n-1}I)^t$. To find $U$ such that

$$U \begin{pmatrix} p(z)I \\ zF(z) \end{pmatrix} = \begin{pmatrix} Q(z) \\ F(z) \end{pmatrix}$$

we write out $p(z) = \sum_{j=0}^n p_j z^j$, $Q(z) = \sum_{j=0}^n z_j Q_j$, and extracting coefficients, we equivalently need $U$ to satisfy

$$U \begin{pmatrix} p_0 I & [p_1 I, \ldots, p_n I] \\ O & A \end{pmatrix} = \begin{pmatrix} [Q_0, \ldots, Q_{n-1}] & Q_n \\ A & O \end{pmatrix}.$$

The matrix $\begin{pmatrix} p_0 I & [p_1 I, \ldots, p_n I] \\ O & A \end{pmatrix}$ has right inverse

$$\begin{pmatrix} p_0^{-1} I & X \\ O & B \end{pmatrix}$$

where $X = -p_0^{-1}[p_1 I, \ldots, p_n I]B$ so that

$$U = \begin{pmatrix} [Q_0, \ldots, Q_{n-1}] & Q_n \\ A & O \end{pmatrix} \begin{pmatrix} p_0^{-1} I & X \\ O & B \end{pmatrix}.$$

Thus, $U$ can be computed directly from $p, Q, A, B$.

## 4. Two Variables and Theorem 1.1

The basic idea of Kummert's argument is to attempt a parametrized version of the one-variable theorem above. The matrix Fejér-Riesz factorization in one variable, which we now review, then becomes crucial in attempting a parametrized version of the implication (3) $\implies$ (2) in the Equivalences Theorem (Theorem 2.1).

**Theorem 4.1 (Matrix Fejér-Riesz).** Let $T(z) = \sum_{j=-n}^n T_j z^j$ be a matrix Laurent polynomial ($T_j \in \mathbb{C}^{N \times N}$) such that $T(z) \geq 0$ for $z \in \mathbb{T}$. Then, there exist a natural number $r \leq N$, a matrix polynomial $A_0 \in \mathbb{C}^{r \times r}[z]$ with $\det A_0(z) \neq 0$ for $z \in \mathbb{D}$, and a polynomial matrix $V \in \mathbb{C}^{N \times N}[z]$ with polynomial inverse such that for $A = \begin{pmatrix} A_0 & 0_{r \times N-r} \end{pmatrix} V$ we have $T = A^*A$ on $\mathbb{T}$. Furthermore, $A$ has degree at most $n$ and a right rational inverse $B$ which is analytic in $\mathbb{D}$.

The case where $T(z)$ is positive definite at all points of $\mathbb{T}$ is usually attributed to Rosenblatt [36]. If $\det T(z)$ vanishes at a finite number of points, it is possible to factor out these zeros from $T$ (see [16, 17]). If $\det T(z)$ is identically zero, it is possible to use operator-valued versions of this theorem which guarantee an

*outer* factorization of $T$. We explain how to go from the case of $\det T \not\equiv 0$ to the case $\det T \equiv 0$ in Appendix A (Subsection A.2). The factorization above can be computed using semidefinite programming or Riccati equations (see, e.g., [27]).

Theorem 4.1 in particular shows that $T(z)$ has rank $r$ except at the finite number of zeros of $\det A_0$. One nice application of Theorem 4.1 is the one variable version of Theorem 1.3.

**Proposition 4.2.** *Let $S : \mathbb{D} \to \mathbb{C}^{M \times N}$ be rational and $\|S(z)\| \leq 1$ for all $z \in \mathbb{D}$. Then, $S$ has a finite contractive TFR.*

*Proof.* Write $S = Q/p$. Then, $|p|^2 I - Q^* Q$ is positive semi-definite on $\mathbb{T}$. By Theorem 4.1, there exists a matrix polynomial $A$ such that $|p|^2 - Q^* Q = A^* A$ on $\mathbb{T}$. Then, $\Phi = \begin{pmatrix} S \\ A/p \end{pmatrix}$ is iso-inner, and by Theorem 3.1 possesses a finite isometric TFR. By Proposition 2.2, we see that $S$ possesses a finite contractive TFR. $\qquad\square$

The following lemma lets us apply Theorem 3.1 to one variable slices.

**Lemma 4.3.** *Suppose $S : \mathbb{D}^2 \to \mathbb{C}^{M \times N}$ is rational and iso-inner. Write $S = Q/p$ where $Q \in \mathbb{C}^{M \times N}[z_1, z_2]$, $p \in \mathbb{C}[z_1, z_2]$ has no zeros in $\mathbb{D}^2$, and $Q$, $p$ have no common factors. Then, $|p|^2 I = Q^* Q$ on $\mathbb{T}^2$, and for each $z_2 \in \mathbb{T}$, the one-variable polynomial $z_1 \mapsto p(z_1, z_2)$ has no zeros in $\mathbb{D}$.*

*Proof.* As in one variable, $|p|^2 I = Q^* Q$ on $\mathbb{T}^2$ by continuity. For fixed $\tau \in \mathbb{T}$, notice that $z_1 \mapsto p(z_1, \tau)$ either has no zeros in $\mathbb{D}$ or is identically zero by Hurwitz's theorem (by considering $\tau$ as a limit of $t \in \mathbb{D}$). If $p(\cdot, \tau)$ is identically zero, then $Q(\cdot, \tau)$ is identically zero because of $|p|^2 I = Q^* Q$ on $\mathbb{T}^2$. Hence, both polynomials are divisible by $z_2 - \tau$, contradicting the assumption of no common factors. Thus, for every $z_2 \in \mathbb{T}$, $z_1 \mapsto p(z_1, z_2)$ has no zeros in $\mathbb{D}$. $\qquad\square$

We are now ready to prove the Main Theorem (Theorem 1.1).

*Proof of Theorem 1.1.* Assume the setup of Theorem 1.1 and write $S = Q/p$ as in Lemma 4.3. We can essentially follow a parametrized version of Remark 3.3 but we use the matrix Fejér-Riesz theorem to deal with certain matrix factorizations.

*Step 1.* Fix $z_2 = w_2 \in \mathbb{T}$, divide $\overline{p(w)}p(z)I - Q(w)^* Q(z)$ by $(1 - \bar{w}_1 z_1)$, and then extract the coefficients of $\bar{w}_1^j z_1^k$ to obtain

(4.1)
$$\frac{\overline{p(w)}p(z)I - Q(w)^* Q(z)}{1 - \bar{w}_1 z_1}$$

$$= \sum_{j,k} \bar{w}_1^j z_1^k T_{jk}(z_2) = (I, \bar{w}_1 I, \ldots, \bar{w}_1^{n_1-1} I) T(z_2) \begin{pmatrix} I \\ z_1 I \\ \vdots \\ z_1^{n_1-1} I \end{pmatrix},$$

where $T(z_2) = (T_{jk}(z_2))_{jk}$ is a positive semi-definite $(n_1N \times n_1N)$ matrix Laurent polynomial. This follows from Theorem 3.1 applied to $p(\cdot, z_2), Q(\cdot, z_2)$. Here, $n_1$ is the maximum of the degree of $p, Q$ with respect to $z_1$.

*Step 2.* Apply the matrix Fejér-Riesz theorem (Theorem 4.1) to $T(z_2)$ to get an $r \times n_1N$ matrix polynomial $A(z_2)$ and an analytic (in $\mathbb{D}$) rational matrix function $B(z_2)$ such that $A^*A = T$ on $\mathbb{T}$ and $AB = I$ in $\mathbb{D}$. For convenience, we define

$$\Lambda(z_1) = (I_N, z_1 I_N, \dots, z_1^{n_1-1} I_N)^t \in \mathbb{C}^{n_1 N \times N}[z_1].$$

Then, for $z_2 = w_2 \in \mathbb{T}$ and $z_1, w_1 \in \mathbb{C}$,

$$\overline{p(w)}p(z)I_N - Q(w)^*Q(z) = (1 - \bar{w}_1 z_1)\Lambda(w_1)^* A(w_2)^* A(z_2)\Lambda(z_1).$$

By Lemma 4.3, for each fixed $z_2 \in \mathbb{T}$ the map $z_1 \mapsto Q(z_1, z_2)/p(z_1, z_2)$ is an iso-inner rational function and Theorem 2.1 guarantees the existence of an isometric matrix $U(z_2)$ such that

$$(4.2) \qquad U(z_2) \begin{pmatrix} p(z)I_N \\ z_1 A(z_2)\Lambda(z_1) \end{pmatrix} = \begin{pmatrix} Q(z) \\ A(z_2)\Lambda(z_1) \end{pmatrix}.$$

*Step 3.* In this step we find a formula for $U(z_2)$ and show it extends to $\bar{\mathbb{D}}$ as a rational iso-inner function in one variable. We can rewrite (4.2) in terms of the coefficients of the powers of $z_1$ by writing $p(z) = \sum_j p_j(z_2)z_1^j$ and $Q(z) = \sum_j Q_j(z_2)z_1^j$, defining $\vec{p}(z_2) = (p_0(z_2)I_N, p_1(z_2)I_N, \dots, p_{n_1}(z_2)I_N)$, and also $\vec{Q}(z_2) = (Q_0(z_2), \dots, Q_{n_1}(z_2))$. Then,

$$(4.3) \qquad U(z_2) \begin{pmatrix} \vec{p}(z_2) \\ O_{r \times N} \ A(z_2) \end{pmatrix} = \begin{pmatrix} \vec{Q}(z_2) \\ A(z_2) \ O_{r \times N}, \end{pmatrix}$$

using $O_{r \times N}$ to denote the $r \times N$ zero matrix. Since $p(0, z_2) = p_0(z_2)$ has no zeros in $\mathbb{D}$, the matrix

$$\begin{pmatrix} \vec{p}(z_2) \\ 0 \ A(z_2) \end{pmatrix}$$

has a rational matrix right inverse of the form

$$\begin{pmatrix} p_0(z_2)^{-1}I \ X(z_2) \\ 0 \qquad B(z_2) \end{pmatrix}.$$

The exact formula for $X(z_2)$ is $(-1/p_0)(p_1I, \dots, p_{n_1}I)B$.
Then,

$$(4.4) \qquad U(z_2) = \begin{pmatrix} \vec{Q}(z_2) \\ A(z_2) \ 0 \end{pmatrix} \begin{pmatrix} p_0(z_2)^{-1}I \ X(z_2) \\ 0 \qquad B(z_2) \end{pmatrix}$$

extends to a rational function holomorphic in $\mathbb{D}$ and isometry-valued on $\mathbb{T}$ away from any singularities. Thus, not only is $U$ uniquely determined (by $A, B$) and iso-inner, but both sides of (4.3) are now holomorphic, so (4.3) extends to $\mathbb{D}$. (We caution that the blocks in (4.4) do not line up as written. There is no need to multiply this out, so there is no real concern.)

*Step 4.* In this step we find an isometric matrix $V$ such that $S$ has a TFR built out of $V$. It turns out $U(z_2)$ as a one variable function has a TFR built out of the same isometry $V$. Indeed, by Theorem 3.1 and Theorem 2.1 there exist a constant isometric matrix $V$ and matrix function $F(z_2)$ such that

$$V \begin{pmatrix} I \\ z_2 F(z_2) \end{pmatrix} = \begin{pmatrix} U(z_2) \\ F(z_2) \end{pmatrix}.$$

A formula for $V$ can be found via Remark 3.3. As we now show, $V$ is the isometry we are looking for. If we multiply on the right by

$$\begin{pmatrix} p(z)I \\ z_1 A(z_2)\Lambda(z_1) \end{pmatrix}$$

and define

$$H(z) := F(z_2) \begin{pmatrix} p(z)I \\ z_1 A(z_2)\Lambda(z_1) \end{pmatrix}, \quad G(z) := A(z_2)\Lambda(z_1),$$

we get

$$V \begin{pmatrix} p(z)I \\ z_1 G(z) \\ z_2 H(z) \end{pmatrix} = \begin{pmatrix} Q(z) \\ G(z) \\ H(z) \end{pmatrix}.$$

By Theorem 2.1, this means $S$ has a finite-dimensional isometric transfer function realization built out of the isometry $V$. This proves Theorem 1.1. $\square$

When we prove the minimality theorem (Theorem 1.4) we will pick up where this proof leaves off. We will later refer to $G^*G$ as the *dominant $z_1$-term* associated with $S$, while we will refer to $H^*H$ as the *sub-dominant $z_2$-term*. We write $G^*G := G(w)^*G(z)$, $H^*H := H(w)^*H(z)$ instead of $G, H$ because the former are uniquely determined while $G, H$ are determined up to left multiplication by isometric matrices. By symmetry we could also construct a dominant $z_2$-term with associated sub-dominant $z_1$-term.

## 5. DETAILED EXAMPLE

In this section we give a detailed example of the four steps presented in the proof of Theorem 1.1. The $N \times N$ identity matrix is written $I_N$, the $N \times N$ zero matrix is written $O_N$, and the $N \times M$ zero matrix is written $O_{N \times M}$.

Consider the following simple rational inner function:

$$S(z) = \frac{1}{2}\begin{pmatrix} z_1(z_1+z_2) & z_1 z_2(z_1-z_2) \\ z_1-z_2 & z_2(z_1+z_2) \end{pmatrix} = \begin{pmatrix} z_1 & 0 \\ 0 & 1 \end{pmatrix} X \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} X \begin{pmatrix} 1 & 0 \\ 0 & z_2, \end{pmatrix}$$

where $X = (1/\sqrt{2})\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ is a unitary. The right expression shows $S$ is a product of inner functions and is therefore inner itself. Since $S$ is a polynomial, the process below will be simpler than the general case but still illustrative. Note then that, referring to the proof of Theorem 1.1, we have $p = 1$ and $Q = S$.

*Step 1.* Set $|z_2| = 1$, divide $I - S(w_1,z_2)^*S(z_1,z_2)$ by $1 - \bar{w}_1 z_1$, and extract coefficients of the monomials $\bar{w}_1^j z_1^k$ in order to write

$$\frac{I - S(w_1,z_2)^*S(z_1,z_2)}{1-\bar{w}_1 z_1} = \sum_{j,k=0,1} \bar{w}_1^j z_1^k T_{jk}(z_2) = (I_2, \bar{w}_2 I_2)T(z_2)\begin{pmatrix} I_2 \\ z_1 I_2 \end{pmatrix}$$

where $T(z_2)$ is the matrix Laurent polynomial

$$T(z_2) = \frac{1}{4}\begin{pmatrix} 3 & z_2 & z_2^{-1} & 1 \\ z_2^{-1} & 3 & -z_2^{-2} & -z_2^{-1} \\ z_2 & -z_2^2 & 1 & z_2 \\ 1 & z_2 & z_2^{-1} & 1 \end{pmatrix}.$$

Necessarily, $T$ is positive semi-definite on $\mathbb{T}$.

*Step 2.* Factor $T$ according to the one-variable matrix Fejér-Riesz theorem. There exist algorithms for doing this ([27]), and it can also be essentially reduced to polynomial algebra and one variable Fejér-Riesz factorizations (see [17] where this is done in a more general setup). We get $T(z_2) = A(z_2)^*A(z_2)$ on $\mathbb{T}$ where

$$A(z_2) = \frac{1}{2}\begin{pmatrix} \sqrt{2} & (\sqrt{2})z_2 & 0 & 0 \\ z_2 & -z_2^2 & 1 & z_2 \end{pmatrix} =: (A_0(z_2), A_1(z_2))$$

has right inverse

$$B(z_2) = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \\ (-\sqrt{2})z_2 & 2 \\ 0 & 0 \end{pmatrix} =: \begin{pmatrix} B_0(z_2) \\ B_1(z_2) \end{pmatrix}.$$

We use the equations above to define the 2×2 matrix polynomials $A_0(z_2)$, $A_1(z_2)$, $B_0(z_2)$, $B_1(z_2)$. Note that the right inverse in general could be rational.

*Step 3.* We find our parametrized unitary $U(z_2)$ in this step. Form the "vectors" of coefficients $\vec{p}(z_2) = (I_2, O_2, O_2)$ and $\vec{Q}(z_2) = (Q_0(z_2), Q_1(z_2), Q_2(z_2))$, where

$$Q_0(z_2) = \frac{1}{2}\begin{pmatrix} 0 & 0 \\ -z_2 & z_2^2 \end{pmatrix}, \quad Q_1(z_2) = \frac{1}{2}\begin{pmatrix} z_2 & -z_2^2 \\ 1 & z_2 \end{pmatrix}, \quad Q_2(z_2) = \frac{1}{2}\begin{pmatrix} 1 & z_2 \\ 0 & 0, \end{pmatrix}$$

and then compute the one variable rational inner function $U(z_2)$ as in (4.3):

$$U(z_2) = \begin{pmatrix} Q_0(z_2) & Q_1(z_2) & Q_2(z_2) \\ A_0(z_2) & A_1(z_2) & O_2 \end{pmatrix} \begin{pmatrix} I_2 & O_2 \\ O_2 & B_0(z_2) \\ O_2 & B_1(z_2) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 1 \\ -\dfrac{z_2}{2} & \dfrac{z_2^2}{2} & \dfrac{1}{\sqrt{2}} & 0 \\ \dfrac{1}{\sqrt{2}} & \dfrac{z_2}{\sqrt{2}} & 0 & 0 \\ \dfrac{z_2}{2} & \dfrac{z_2^2}{2} & \dfrac{1}{\sqrt{2}} & 0 \end{pmatrix}.$$

The fourth step is to find a TFR for $U(z_2)$. To do this we apply Remark 3.3. Let us emphasize the steps. Divide $I_4 - U(w_2)^*U(z_2)$ by $1 - \bar{w}_2 z_2$ and extract coefficients of $\bar{w}_2^j z_2^k$ to write

$$\frac{I_4 - U(w_2)^*U(z_2)}{1 - \bar{w}_2 z_2} = (I_4, \bar{w}_2 I_4) Y \begin{pmatrix} I_4 \\ z_2 I_4 \end{pmatrix}$$

where

$$Y = \begin{pmatrix} \begin{pmatrix} \dfrac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} & O_2 & \begin{pmatrix} 0 & -\dfrac{1}{2} \\ 0 & 0 \end{pmatrix} & O_2 \\ O_2 & O_2 & O_2 & O_2 \\ \begin{pmatrix} 0 & 0 \\ -\dfrac{1}{2} & 0 \end{pmatrix} & O_2 & \begin{pmatrix} 0 & 0 \\ 0 & \dfrac{1}{2} \end{pmatrix} & O_2 \\ O_2 & O_2 & O_2 & O_2 \end{pmatrix}.$$

Then, we factor $Y = C^*C$ where

$$C = \left[ \begin{pmatrix} \dfrac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{pmatrix}, O_2, \begin{pmatrix} 0 & -\dfrac{1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}, O_2 \right].$$

Note that

$$D = \begin{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix} \\ O_2 \\ O_2 \\ O_2 \end{pmatrix}$$

is a right inverse for $C$ (i.e., $CD = I_2$). Set

$$F(z_2) = C \begin{pmatrix} I_4 \\ z_2 I_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} z_2 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

We need to compute the unitary (or isometry in general) $V$ such that

$$V \begin{pmatrix} I_4 \\ z_2 F(z_2) \end{pmatrix} = \begin{pmatrix} U(z_2) \\ F(z_2) \end{pmatrix}.$$

After equating coefficients of powers of $z_2$, this is equivalent to

$$V \begin{pmatrix} I_4 & O_{4\times 8} \\ O_{2\times 4} & C \end{pmatrix} = \begin{pmatrix} \overset{U_0}{\left[ \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{pmatrix}, O_2 \right]} & \overset{U_1}{\left[ \begin{pmatrix} 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}, O_2 \right]} & \overset{U_2}{O_{2\times 4}} \end{pmatrix}$$

where $U(z_2) = U_0 + z_2 U_1 + z_2^2 U_2$. Using the right inverse $D$, we have

$$V = \begin{pmatrix} \overset{U_0}{\left[ \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{pmatrix}, O_2 \right]} & \overset{U_1}{\left[ \begin{pmatrix} 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}, O_2 \right]} & \overset{U_2}{O_{2\times 4}} \end{pmatrix} \begin{pmatrix} I_4 & O_{4\times 2} \\ O_{8\times 4} & D \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This is the desired unitary out of which we build our TFR. Setting

$$V_{11} = O_2, \qquad V_{12} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0, \end{pmatrix}$$

$$V_{21} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{pmatrix}, \qquad V_{22} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

we have

$$S(z_1, z_2) = V_{11} + V_{12}\Delta(z)(I - V_{22}\Delta(z))^{-1}V_{21}$$

where

$$\Delta(z_1, z_2) = \begin{pmatrix} z_1 I_2 & O_2 \\ O_2 & z_2 I_2 \end{pmatrix}.$$

This is easy to verify since $(V_{22}\Delta(z))^3 = O$ so that the formula reduces to

$$S(z) = V_{12}\Delta(z)(I + V_{22}\Delta(z) + (V_{22}\Delta(z)^2)V_{21},$$

which can be verified by hand.

While the above method involves several steps it is entirely systematic. Since $S$ is a product of simple inner functions, there are *ad hoc* ways of coming up with a TFR which might be shorter.

## 6. MATRIX AGLER DECOMPOSITIONS IN TWO VARIABLES

Theorem 1.1 makes it possible to prove Agler's theorem (Theorem 1.2). Cole-Wermer [15] showed that in the scalar case it is enough to prove Agler's theorem for rational inner functions because holomorphic $f : \mathbb{D}^2 \to \mathbb{D}$ can be approximated locally uniformly by a rational inner function (Theorem 5.5.1 of Rudin [37]). This approximation argument does not seem to transfer to the matrix-valued function setting, but there is a workaround.

**Lemma 6.1.** *Let $f : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ be holomorphic and $\|f(z)\| \leq 1$ for all $z \in \mathbb{D}^d$. Suppose $\|f(z_0)\| = 1$ for some $z_0 \in \mathbb{D}^d$. Then, there exist unitary matrices $U_1, U_2$ such that $U_1 f U_2$ is a direct sum of a constant unitary matrix and a matrix valued holomorphic function $g$ on $\mathbb{D}^d$ with $\|g(z)\| < 1$ for all $z \in \mathbb{D}^d$.*

*Proof.* If $\|f(z_0)\| = 1$, then there exists $v \in \mathbb{C}^N$ with $|v| = 1$ such that $|f(z_0)v| = 1$. By the maximum principle, $\langle f(z)v, f(z_0)v \rangle$ is constant and equal to one. Then, by equality in Cauchy-Schwarz, $f(z)v \equiv f(z_0)v$. Since $f(z)$ has at most norm one, $v$ is reducing for $f(z)$, meaning $f(z)w \perp f(z)v$ whenever $v \perp w$. Thus, $f(z)$ can be written in the form

$$\begin{pmatrix} 1 & 0 \\ 0 & g(z) \end{pmatrix}$$

using the block decomposition $\mathbb{C}f(z_0)v \oplus (f(z_0)v)^\perp \times (\mathbb{C}v) \oplus v^\perp$. We can, of course, iterate this argument until we are left with the claimed decomposition. □

This lets us reduce to the case of $f$ with $\|f(z)\| < 1$ for all $z \in \mathbb{D}^d$. The following is found in Rudin's book [37] in the scalar case (see Theorem 5.5.1 of [37]). Define

$$\|f\|_{\mathbb{D}^d} = \sup_{z \in \mathbb{D}^d} \|f(z)\|.$$

**Lemma 6.2.** *Suppose* $f : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ *is holomorphic and* $\|f(z)\| < 1$ *for all* $z \in \mathbb{D}^d$. *Then, for any* $r \in (0,1)$ *and* $\varepsilon > 0$ *there exists* $P \in \mathbb{C}^{M \times N}[z_1, \ldots, z_d]$ *such that* $\|P\|_{\mathbb{D}^d} < 1$ *and* $\|f - P\|_{r\mathbb{D}^d} < \varepsilon$.

*Consequently, every such* $f$ *is a local uniform limit of matrix polynomials with supremum norm strictly less than 1.*

*Proof.* Set $f_r(z) = f(rz)$ for $r \in (0,1)$. For fixed $r \in (0,1)$ there exists $s \in (0,1)$ such that $\|f_r - f_{rs}\|_{\mathbb{D}^d} < \varepsilon/2$ since $f_r$ is uniformly continuous on $\bar{\mathbb{D}}^d$. Note $\|f_s\|_{\mathbb{D}^d} < 1$. Choose a Taylor polynomial $P$ of $f_s$ such that $\|f_s - P\|_{\mathbb{D}^d} < \min(1 - \|f_s\|_{\mathbb{D}^d}, \varepsilon/2)$. Then,

$$\|P\|_{\mathbb{D}^d} < 1 \quad \text{and} \quad \|f_r - P_r\|_{\mathbb{D}^d} \leq \|f_r - f_{rs}\|_{\mathbb{D}^d} + \|f_{rs} - P_r\|_{\mathbb{D}^d} < \varepsilon. \qquad \square$$

We need the following Fejér-Riesz type theorem of Dritschel.

**Theorem 6.3 (Dritschel [18]).** *Let* $T(z) = \sum_{j \in \mathbb{Z}^d} T_j z^j$ *be a matrix-valued Laurent polynomial in* $d$ *variables; that is,* $T_j \in \mathbb{C}^{N \times N}$ *for* $j \in \mathbb{Z}^d$ *and at most finitely many* $T_j \neq 0$. *If there is a* $\delta > 0$ *such that* $T(z) \geq \delta I$ *on* $\mathbb{T}^d$, *then there exists a matrix polynomial* $A \in \mathbb{C}^{M \times N}[z_1, \ldots, z_d]$ *such that* $T = A^*A$ *on* $\mathbb{T}^d$.

We sketch a simple proof with some new elements in the appendix (see Subsection A.2).

**Lemma 6.4.** *If* $P : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ *is a matrix polynomial such that* $\|P\|_{\mathbb{D}^d} < 1$ *then there exists a matrix polynomial* $A$ *such that* $\begin{pmatrix} P \\ A \end{pmatrix}$ *is iso-inner. If* $d = 1, 2$, *then* $P$ *has a finite contractive TFR.*

*Proof.* On $\mathbb{T}^d$, $I - P^*P$ is a positive definite matrix Laurent polynomial. By Theorem 6.3 we can factor $I - P^*P = A^*A$. Then, $S = \begin{pmatrix} P \\ A \end{pmatrix}$ is isometry-valued on $\mathbb{T}^d$. If $d = 1, 2$, then $S$ has a finite isometric TFR by Theorem 1.1, and hence $P$ possesses a finite contractive TFR by Proposition 2.2. $\qquad \square$

Positive semi-definite kernels are defined in Definition A.5. Notice that an expression of the form $F(w)^*F(z)$ will always be positive semi-definite. By the above lemma and Theorem 2.1, any matrix polynomial $P \in \mathbb{C}^{M \times N}[z_1, z_2]$ with $\|P\|_{\mathbb{D}^2} < 1$ will satisfy a formula of the form

$$I - P(w)^*P(z) = k_0(w, z) + \sum_{j=1}^{2} (1 - \bar{w}_j z_j) k_j(w, z)$$

where $k_0, k_1, k_2$ are positive semi-definite kernels. The term $k_0$ can be absorbed into $k_1$ since

$$\frac{k_0(w, z)}{1 - \bar{w}_1 z_1}$$

is positive semi-definite by the Schur product theorem. Thus, the following corollary holds for such strictly contractive matrix polynomials in two variables. Such formulas are called *Agler decompositions*.

**Corollary 6.5.** *Let $f : \mathbb{D}^2 \to \mathbb{C}^{M \times N}$ be holomorphic with $\|f(z)\| \leq 1$ for $z \in \mathbb{D}^2$. Then, there exist positive semi-definite kernels $k_1, k_2 : \mathbb{D}^2 \times \mathbb{D}^2 \to \mathbb{C}^{N \times N}$ such that*

$$I - f(w)^* f(z) = \sum_{j=1}^{2} (1 - \bar{w}_j z_j) k_j(w, z).$$

*Sketch of Proof.* The hard work has already been done while the general out-line and some technicalities are essentially in [15], so we only sketch the proof. We can assume that $f$ is pointwise strictly contractive by Lemma 6.1. Then, $f$ is a local uniform limit of matrix polynomials with supremum norm strictly less than one by Lemma 6.2. Each of these possesses an Agler decomposition by the discussion above.

The final part of the argument is the piece found in [15]. The kernels in the Agler decomposition are locally bounded because of the estimate

$$\frac{1}{(1 - |z_1|^2)(1 - |z_2|^2)} I \geq \frac{I - f(z)^* f(z)}{(1 - |z_1|^2)(1 - |z_2|^2)} \geq \frac{k_1(z, z)}{1 - |z_2|^2} \geq k_1(z, z).$$

This shows the kernels in Agler decompositions form a normal family. Subsequences converge locally uniformly to form positive semi-definite kernels in an Agler decomposition for $f$. □

The above corollary proves Theorem 1.2. The proof is essentially the same as (3) $\implies$ (1) in the equivalences theorem (Theorem 2.1) since positive semi-definite kernels can be factored as $F(w)^* F(z)$ for some possibly operator valued function $F$. Readers who have ventured this far (and are not in the cognoscenti of this material) may benefit from some context at this point. The fundamental contribution of Agler can perhaps be encapsulated in the following result.

**Theorem 6.6** (*Agler* [1, 2]). *Let $f : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ be holomorphic. Assume $\|f(z)\| \leq 1$ for $z \in \mathbb{D}^d$. Then, the following are equivalent:*

(1) *$f$ satisfies a von Neumann inequality $\|f(T)\| = \|(f_{j,k}(T))_{j,k}\| \leq 1$ for every $d$-tuple $T = (T_1, \ldots, T_d)$ of pairwise commuting strictly contractive operators (on some underlying Hilbert space).*

(2) *$f$ has an Agler decomposition: that is, there exist positive semi-definite kernels $k_1, \ldots, k_d : \mathbb{D}^d \times \mathbb{D}^d \to \mathbb{C}^{N \times N}$ such that*

$$I - f(w)^* f(z) = \sum_{j=1}^{d} (1 - \bar{w}_j z_j) k_j(w, z).$$

(3) *$f$ has a contractive transfer function realization: there exists a contractive operator with block decomposition $\left( T = \begin{smallmatrix} A & B \\ C & D \end{smallmatrix} \right)$ on some Hilbert space such that $f(z) = A + B\Delta(z)(I - D\Delta(z))^{-1} C$, where $\Delta(z) = \sum_{j=1}^{d} z_j P_j$ and the $P_j$ are pairwise orthogonal projections which sum to the identity on the domain of $D$.*

Theorem 1.2 was originally proven via Andô's inequality [7] which gives item (1) above. The approach we have given sidesteps the use of von Neumann's inequality and the implication (1) $\implies$ (2) in Theorem 6.6. The proof of (1) $\implies$ (2) is possibly the hardest part of the theorem, and is non-constructive since it uses a Hahn-Banach cone separation argument. On the other hand, (2) $\implies$ (1) is a relatively straightforward matter of "plugging" the $d$-tuple $T$ into the Agler decomposition in item (2) in an appropriate sense. (See [15] for details.) Ball-Sadosky-Vinnikov [11] have a different way to prove Theorem 1.2 directly using multi-evolution scattering systems. Theorem 1.2's analogue for 3 or more variables fails because the von Neumann inequality fails for 3 or more contractions [39]. Thus, Theorem 6.6 gives the best way of demonstrating that a function does not have a contractive TFR; specifically, showing that it fails the von Neumann inequality. It is probably difficult to directly show that a function fails item (2) or (3) in Theorem 6.6.

We conclude this section by plugging Dritschel's strong Fejér-Riesz type result (stated below) into earlier arguments in order to show rational contractive matrix-valued functions in two variables have a finite contractive TFR (Theorem 1.3).

**Theorem 6.7 (Dritschel [19]).** *Let $T(z) = \sum_{j \in \mathbb{Z}^2} T_j z^j$ be a matrix-valued Laurent polynomial in two variables; that is, $T_j \in \mathbb{C}^{N \times N}$ for $j \in \mathbb{Z}^2$ and at most finitely many $T_j \neq 0$. If $T(z) \geq 0$ on $\mathbb{T}^2$, then there exists a matrix polynomial $A \in \mathbb{C}^{M \times N}[z_1, z_2]$ such that $T = A^* A$ on $\mathbb{T}^2$.*

This theorem is considerably deeper than Theorem 6.3, and both theorems also apply to operator-valued functions. An earlier sums-of-squares theorem of Scheiderer, which applied to polynomials on a much more general class of two-dimensional domains (than simply $\mathbb{T}^2$), implies Theorem 6.7 in the scalar case [38].

*Proof of Theorem 1.3.* Apply the proof of Proposition 4.2 with Theorem 6.7 in place of Theorem 4.1. $\qquad\square$

## 7. MORE ON FINITE TFRS

We need to collect one more fact about finite-dimensional TFRs before proving the minimality theorem. If we have an Agler decomposition of an iso-inner function $S = Q/p$ written in lowest terms, then the sums of squares terms are rational with denominator $p$.

**Theorem 7.1.** *Suppose $S : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ is rational and iso-inner. Write $S = Q/p$ in lowest terms with $Q \in \mathbb{C}^{M \times N}[z_1, \ldots, z_d]$ and $p \in \mathbb{C}[z_1, \ldots, z_d]$. Suppose we have an Agler decomposition*

$$I_N - S(w)^* S(z) = \sum_{j=1}^{d} (1 - \bar{w}_j z_j) F_j(w)^* F_j(z),$$

*where the $F_j$ are matrix functions. Then, for $j = 1, \ldots, d$, $p(z)F_j(z)$ is a matrix polynomial.*

The significance of this theorem is that although $S$ has a TFR with denominator $\det(I - D\Delta(z))$, this polynomial may not be the lowest-degree denominator of $S$.

*Proof.* By Theorem 2.1, we already see that each $F_j$ is rational and holomorphic in $\mathbb{D}^d$. To prove that $H_j := pF_j$ is a matrix polynomial, consider

$$\overline{p(w)}p(z)I_N - Q(w)^*Q(z) = \sum_{j=1}^{d}(1 - \bar{w}_j z_j)H_j(w)^*H_j(z).$$

Fix $\tau \in \mathbb{T}^d$ and set $z = \zeta\tau$, $w = \eta\tau$ for $\zeta, \eta \in \mathbb{D}$. Then,

$$\overline{p(\eta\tau)}p(\zeta\tau)I_N - Q(\eta\tau)^*Q(\zeta\tau) = (1 - \bar{\eta}\zeta)\sum_{j=1}^{d}H_j(\eta\tau)^*H_j(\zeta\tau).$$

Because $S^*S = I_N$ on $\mathbb{T}^d$, the lefthand side above is divisible by $(1 - \bar{\eta}\zeta)$, and therefore

$$\sum_{j=1}^{d}H_j(\eta\tau)^*H_j(\zeta\tau)$$

is a polynomial in $\zeta, \bar{\eta}$ of degree in each less than the total degree of $p$ and $Q$. For simplicity, we can regroup $\sum_{j=1}^{d}H_j(w)^*H_j(z) = H(w)^*H(z)$, where now $H(\eta\tau)^*H(\zeta\tau)$ is a polynomial in $\zeta, \bar{\eta}$ for every $\tau \in \mathbb{T}^d$. If we write out the homogeneous expansion of $H$,

$$H(z) = \sum_{j=0}^{\infty}P_j(z),$$

we see that

$$H(\eta\tau)^*H(\zeta\tau) = \sum_{j,k}\bar{\eta}^j\zeta^k P_j(\tau)^*P_k(\tau).$$

In particular, for $j$ greater than the total degrees of $p$ and $Q$, the coefficient of $\bar{\eta}^j\zeta^j$ vanishes for every $\tau$; specifically, we have $P_j(\tau)^*P_j(\tau) \equiv 0$ for all $\tau \in \mathbb{T}^d$. Since $P_j$ is a matrix polynomial, this implies $P_j \equiv 0$ for $j$ greater than the total degrees of $p$ and $Q$. Therefore, $H$ is a polynomial. $\square$

We conclude this short section with a few asides. The Agler norm (sometimes Schur-Agler norm) for holomorphic $f : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ is

$$\|f\|_{\mathcal{A}_d} := \sup_T \|f(T)\|$$

where the supremum is taken over all $d$-tuples $T = (T_1, \ldots, T_d)$ of strictly contractive pairwise commuting operators on some Hilbert space. The Agler class $\mathcal{A}_d$ consists of functions satisfying $\|f\|_{\mathcal{A}_d} \leq 1$.

The argument in the proof above is related to the argument used to prove the following automatic finite-dimensionality result.

**Theorem 7.2.** *Suppose* $S : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ *is rational, iso-inner, or coiso-inner* $(SS^* = I$ *on* $\mathbb{T}^d)$, *and belongs to the Agler class* $\mathcal{A}_d$. *Then, $S$ has a finite-dimensional isometric* (*respectively coisometric* TFR *as in Theorem* 2.1.

The essence of this theorem was first proved in Cole-Wermer [15]. Although it was only stated and proved in the scalar case for $d = 2$, the proof goes through easily to all $d$ and for iso-inner functions. We gave a proof with some bounds on degrees and the numbers of squares involved in the scalar case in [31]. A proof of the square matrix-valued case is in [9]. Extending to the iso-inner (non-square) case causes no difficulties. The coisometric case follows from Proposition 2.3. A proof where $S$ is assumed to be a polynomial is also given in [13]. The next theorem also produces a family of functions with finite TFRs.

**Theorem 7.3** (**Grinshpan et al** [23]). *Suppose* $S : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ *is rational and analytic on a neighborhood of* $\bar{\mathbb{D}}^d$, *and* $\|S\|_{\mathcal{A}_d} < 1$. *Then, $S$ has a finite-dimensional contractive* TFR *as in Theorem* 2.1.

The following question asks about what is still left open.

**Question 7.4.** *For $d > 2$, if $S : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ is rational, $\|S\|_{\mathcal{A}_d} = 1$, and is neither iso-inner nor coiso-inner. Then, does $S$ have a finite-dimensional contractive* TFR?

We also do not know how essential analyticity on $\bar{\mathbb{D}}^d$ is for Theorem 7.3. Note $d = 1, 2$ follows from Theorem 1.3.

## 8. KUMMERT'S MINIMALITY THEOREM

In this section we discuss minimality of size breakdowns for finite TFRs, namely, Theorem 1.4. Minimality in one variable follows directly from Theorem 2.1.

**Proposition 8.1.** *Let* $S : \mathbb{D} \to \mathbb{C}^{M \times N}$ *be rational and iso-inner. Then, the minimal size of an isometric* TFR *for $S$ is the rank of the positive semi-definite kernel*

$$(w, z) \mapsto \frac{I - S(w)^* S(z)}{1 - \bar{w} z}.$$

The definition of the rank of a positive semi-definite kernel is given in Definition A.6 in Appendix A. In two variables, we will frequently refer to the dominant $z_1$-term $G^* G$ and sub-dominant $z_2$-term $H^* H$ associated with $S$ which were constructed in the proof of Theorem 1.1 (see the end of Section 4). Note that the number of rows of $G$ matches the generic rank of the matrix $T(z_2)$ as in equation (4.1). This cannot be reduced because this is the generic or maximal rank of the

positive semi-definite kernels

$$(w_1, z_1) \mapsto \frac{I - S(w_1, z_2)^* S(z_1, z_2)}{1 - \bar{w}_1 z_1} \quad \text{as } z_2 \text{ varies over } \mathbb{T}.$$

Note division of (4.1) by $\overline{p(w_1, z_2)} p(z_1, z_2)$ will not change the rank of the positive semi-definite kernel and does not introduce any poles in $\mathbb{D}$ since $p(\cdot, z_2)$ has no zeros in $\mathbb{D}$ by Lemma 4.3.

We claim that in the inner case the rank of $H^*H$ is also as small as possible. We suspect this happens in the iso-inner case but cannot prove it.

**Question 8.2.** *If $S : \mathbb{D}^2 \to \mathbb{C}^{M \times N}$ is iso-inner (and not inner), does the construction in Section 4 produce a size breakdown $(r_1, r_2)$ with $r_1$ equal to the generic size of a TFR for $S(\cdot, z_2)$ (for $z_2 \in \mathbb{T}$) and $r_2$ equal to the generic size of a TFR for $S(z_1, \cdot)$ (for $z_1 \in \mathbb{T}$)?*

This question is subtle because every iso-inner function $S$ is a submatrix of an inner function $\Phi$ with the same size breakdown. We have built a size breakdown with $r_1$ minimal so $r_1$ must also be minimal for $\Phi$. We could then build a TFR with size breakdown $(r_1, r_2^*)$ where $r_2^*$ is minimal for $\Phi$. Is it minimal for the restriction to $S$?

The next result characterizes $G^*G$ and $H^*H$.

**Proposition 8.3.** *Assume $S : \mathbb{D}^2 \to \mathbb{C}^{M \times N}$ is rational and iso-inner. Write $S = Q/p$ in lowest terms. Suppose we had a formula*

$$\overline{p(w)} p(z) I - Q(w)^* Q(z) = (1 - \bar{w}_1 z_1) \Gamma_1(w)^* \Gamma_1(z) + (1 - \bar{w}_2 z_2) \Gamma_2(w)^* \Gamma_2(z)$$

*where $\Gamma_1, \Gamma_2$ are matrix polynomials. Then,*

$$(8.1) \quad (w, z) \mapsto \frac{G(w)^* G(z) - \Gamma_1(w)^* \Gamma_1(z)}{1 - \bar{w}_2 z_2} = \frac{\Gamma_2(w)^* \Gamma_2(z) - H(w)^* H(z)}{1 - \bar{w}_1 z_1}$$

*is a positive semi-definite polynomial kernel. Here again, $G^*G$ is the dominant $z_1$-term and $H^*H$ is the sub-dominant $z_2$-term.*

This result characterizes $G^*G$ as maximal and $H^*H$ as minimal in the above sense. Indeed, if some other kernel $L^*L$ satisfied the same property as $G^*G$, then both

$$\frac{G^*G - L^*L}{1 - \bar{w}_2 z_2} \quad \text{and} \quad \frac{L^*L - G^*G}{1 - \bar{w}_2 z_2}$$

would be positive semi-definite forcing $G^*G = L^*L$.

*Proof of Proposition 8.3.* If we set $z_2 = w_2 \in \mathbb{T}$ we get

$$\frac{\overline{p(w)} p(z) I - Q(w)^* Q(z)}{1 - \bar{w}_1 z_1} = \Gamma_1(w)^* \Gamma_1(z) = G(w)^* G(z).$$

The left side has degree at most $n_1 - 1$ in $z_1$. We claim $\Gamma_1(z)$ has degree at most $n_1 - 1$ in $z_1$. Consider $\Gamma_1$'s top degree term $\gamma(z_2)z_1^k$ where $\gamma(z_2)$ is a matrix polynomial. Then, the term $\bar{w}_1^k z_1^k$ appears on the righthand side with coefficient $\gamma(z_2)^*\gamma(z_2)$ for $z_2 \in \mathbb{T}$. If $k > n_1 - 1$ then $\gamma(z_2)^*\gamma(z_2) \equiv 0$ on $\mathbb{T}$ implying $\gamma(z_2) \equiv 0$ on $\mathbb{T}$ and also on $\mathbb{C}$ by analyticity. Thus, $\Gamma_1$ has degree at most $n_1 - 1$ in $z_1$.

Just as we have factored $G(z) = A(z_2)\Lambda(z_1)$ we can also factor $\Gamma_1(z) = C(z_2)\Lambda(z_1)$. Recall $\Lambda(z_1) = (I, z_1 I, \dots, z_1^{n_1-1} I)^t$. Upon extracting coefficients of $\bar{w}_1^j z_1^k$ we see that

$$A(z_2)^* A(z_2) = C(z_2)^* C(z_2) \quad \text{for } z_2 \in \mathbb{T}.$$

This is related to characterizing uniqueness in the matrix Fejér-Riesz theorem. We address this in the appendix in Theorem A.4. By Theorem A.4, since $A$ has a left inverse, there exists a one variable iso-inner function $\Phi$ such that $C = \Phi A$.

Thus,

$$\frac{A(w_2)^* A(z_2) - C(w_2)^* C(z_2)}{1 - \bar{w}_2 z_2} = A(w_2)^* \left( \frac{I - \Phi(w_2)^* \Phi(z_2)}{1 - \bar{w}_2 z_2} \right) A(z_2),$$

which is positive semi-definite. Applying $\Lambda(w_1)^*$ on the left and $\Lambda(z_1)$ on the right, we get that

$$\frac{G(w)^* G(z) - \Gamma_1(w)^* \Gamma_1(z)}{1 - \bar{w}_2 z_2} = \frac{\Gamma_2(w)^* \Gamma_2(z) - H(w)^* H(z)}{1 - \bar{w}_1 z_1}$$

is positive semi-definite. It is a polynomial kernel because $A^* A = C^* C$ on $\mathbb{T}$.  □

We now switch to the square/inner case and show that the Kummert construction gives the best possible size breakdown $r = (r_1, r_2)$. We need to show $H(w)^* H(z)$ has the minimal rank possible in the sense that it matches the generic size of a TFR for $S(z_1, \cdot)$ for $z_1 \in \mathbb{T}$. To do this, we show that we can "reflect" an Agler decomposition of $S$ to get an Agler decomposition for $\check{S}$, and this reflection reverses the dominant and sub-dominant properties of $G^* G$ and $H^* H$. This is not the original approach of Kummert; instead, it more closely resembles the Hilbert space approach in [12]. Recall $\check{S}(z) = S(\bar{z})^*$.

**Proposition 8.4.** *Suppose $S : \mathbb{D}^2 \to \mathbb{C}^{N \times N}$ is rational and inner. Write $S = Q/p$ in lowest terms. Suppose we had a formula*

$$(8.2) \qquad \overline{p(w)} p(z) I_N - Q(w)^* Q(z) = (1 - \bar{w}_1 z_1)\Gamma_1(w)^* \Gamma_1(z)$$
$$+ (1 - \bar{w}_2 z_2)\Gamma_2(w)^* \Gamma_2(z)$$

*where $\Gamma_1, \Gamma_2$ are matrix polynomials. Then,*

$$(8.3a) \qquad \bar{\Gamma}_1(z) := \frac{1}{z_1 p(1/z)} \Gamma_1(1/z)\check{S}(z),$$

*and*

(8.3b) $$\bar{\Gamma}_2(z) := \frac{1}{z_2 p(1/z)} \Gamma_2(1/z) \check{S}(z)$$

*are matrix polynomials, and*

(8.4) $$\overline{p(w)} \check{p}(z) I - \check{Q}(w)^* \check{Q}(z) = (1 - \bar{w}_1 z_1) \bar{\Gamma}_1(w)^* \bar{\Gamma}_1(z)$$
$$+ (1 - \bar{w}_2 z_2) \bar{\Gamma}_2(w)^* \bar{\Gamma}_2(z).$$

*The sub-dominant $z_2$-term of $S$ reflects to the dominant $z_2$-term of $\check{S}$.*

When we say *reflects* above we mean the operations

$$\Gamma_1 \mapsto \bar{\Gamma}_1 \quad \text{and} \quad \Gamma_2 \mapsto \bar{\Gamma}_2$$

listed in the proposition statement equations (8.3a), (8.3b). Notice that reflection of the $\Gamma_1$ term is slightly different from the reflection of the $\Gamma_2$ term.

*Proof of Proposition 8.4.* Since $S(z)^* S(z) = I$ on $\mathbb{T}^2$ (where defined), we have $I = S(1/\bar{z})^* S(z) = S(z) S(1/\bar{z})^*$ for $z \in \mathbb{C}^2$ where defined. (This is where $M = N$ gets used.) Thus, $Q(1/z)\check{Q}(z) = p(1/z)\check{p}(z)I$. Now, take equation (8.2), replace $z, w$ with $1/z, 1/w$, multiply on the right by $\check{Q}(z)$ and left by $\check{Q}(w)^*$, and finally divide through by $-\overline{p(1/w)}p(1/z)$ to get (8.4) after applying various simplifications. Of course, we have the caveat that the formula only holds where all of the operations are defined. Fortunately, (8.4) only needs to hold on an open set for the proof of (3) $\implies$ (1), (2) in Theorem 2.1 to go through (bonus (B1) of Theorem 2.1 addresses this). We automatically obtain that $\bar{\Gamma}_1, \bar{\Gamma}_2$ are polynomials by Theorem 7.1, since if $Q/p$ is in lowest terms then $\check{Q}/\check{p}$ is too.

If we reflect equation (8.1) in the sense of replacing $z, w$ with $1/z, 1/w$ and conjugating by $\check{Q}$, we obtain

$$\bar{w}_1 z_1 \frac{\bar{G}(w)^* \bar{G}(z) - \bar{\Gamma}_1(w)^* \bar{\Gamma}_1(z)}{1 - (\bar{w}_2 z_2)^{-1}} = \bar{w}_2 z_2 \frac{\bar{\Gamma}_2(w)^* \bar{\Gamma}_2(z) - \bar{H}(w)^* \bar{H}(z)}{1 - (\bar{w}_1 z_1)^{-1}},$$

which can be rearranged into

$$\frac{\bar{\Gamma}_1(w)^* \bar{\Gamma}_1(z) - \bar{G}(w)^* \bar{G}(z)}{1 - \bar{w}_2 z_2} = \frac{\bar{H}(w)^* \bar{H}(z) - \bar{\Gamma}_2(w)^* \bar{\Gamma}_2(z)}{1 - \bar{w}_1 z_1}.$$

This is still a positive semi-definite polynomial kernel. Thus, $\bar{H}^* \bar{H}$ dominates an arbitrary $z_2$-term, making it the dominant $z_2$-term for $\check{S}$. □

*Proof of Theorem 1.4.* By Proposition 8.4 the subdominant $z_2$-term $H^* H$ of $S$ reflects to the dominant $z_2$-term of $\check{S}$, $\bar{H}^* \bar{H}$. Note that this reflection does not

change the rank of a positive semi-definite kernel. The rank of $\check{H}^*\check{H}$ is then the generic rank of

$$(w_2, z_2) \mapsto \frac{I - \check{S}(z_1, w_2)^* \check{S}(z_1, z_2)}{1 - \bar{w}_2 z_2}$$

for $z_1 \in \mathbb{T}$. This matches the generic size of a TFR for $\check{S}(z_1, \cdot)$ which matches the generic size of a TFR for $S(z_1, \cdot)$ by the adjunction formula, Proposition 2.3. Thus, the rank of $H^*H$ matches the generic rank of

$$(w_2, z_2) \mapsto \frac{I - S(z_1, w_2)^* S(z_1, z_2)}{1 - \bar{w}_2 z_2}. \qquad \square$$

## 9. APPLICATION TO INNER POLYNOMIALS

Of special interest in the papers connecting wavelets to TFRs is the case of iso-inner and inner polynomials [13, 14]. In one variable, we have the following well-known result.

**Proposition 9.1.** *Let $S \in \mathbb{C}^{M \times N}[z]$ be iso-inner. Then, every isometric TFR of minimal size for $S$ is built out of an isometric matrix $T = \left( \begin{smallmatrix} A & B \\ C & D \end{smallmatrix} \right)$ where $D$ is nilpotent.*

We prove this using the following also well-known characterization of minimality.

**Proposition 9.2.** *Let $S : \mathbb{D} \to \mathbb{C}^{M \times N}$ be rational and iso-inner with minimal isometric TFR built out of the isometric matrix $T = \left( \begin{smallmatrix} A & B \\ C & D \end{smallmatrix} \right)$. Then,*

$$\mathrm{span}\{\mathrm{range}(D^j C) : j = 0, 1, \dots\} = \mathrm{domain}(D) \text{ and } \bigcap_{j \geq 0} \mathrm{kernel}(B D^j) = \{0\}.$$

*Proof.* First note that if $S$ has a TFR via $T$, meaning

$$S(z) = A + zB(I - zD)^{-1}C,$$

then it also has a TFR via

$$\begin{pmatrix} I & 0 \\ 0 & U^* \end{pmatrix} T \begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix} = \begin{pmatrix} A & BU \\ U^*C & U^*DU \end{pmatrix}$$

where $U$ is a unitary matrix with the same dimensions as $D$. This is apparent from the formula $A + zBU(I - zU^*DU)^{-1}U^*C = S(z)$. We can now apply a unitary change of coordinates and break up the domain/codomain of $D$ into $\mathcal{H} = \mathrm{span}\{D^j C : j = 0, 1, \dots\}$ and its orthogonal complement $\mathcal{H}^\perp$. In these new coordinates, $T$ takes the form

$$\begin{array}{c} \qquad \mathbb{C}^N \;\; \mathcal{H} \;\; \mathcal{H}^\perp \\ \begin{array}{c} \mathbb{C}^M \\ \mathcal{H} \\ \mathcal{H}^\perp \end{array} \begin{pmatrix} A & B_1 & B_2 \\ C & D|_{\mathcal{H}} & * \\ 0 & 0 & * \end{pmatrix}, \end{array}$$

since $D$ maps $\mathcal{H}$ to itself and range$(C) \subset \mathcal{H}$. Since the formula for $S$ is only determined by $D|_{\mathcal{H}}$, we see that $S$ has an isometric TFR via the matrix

$$\begin{pmatrix} A & B_1 \\ C & D|_{\mathcal{H}} \end{pmatrix}$$

which has a smaller size unless $\mathcal{H}^{\perp} = \{0\}$ or rather $\mathcal{H} = \text{domain}(D)$.

For the second identity, we break up the domain of $D$ into

$$\mathcal{L} = \bigcap_{j \geq 0} \text{kernel}(BD^j)$$

and its orthogonal complement $\mathcal{L}^{\perp}$. Using this orthogonal decomposition we can write $T$ in new coordinates as

$$\begin{array}{c} \qquad\quad \mathbb{C}^N \;\; \mathcal{L}^{\perp} \;\; \mathcal{L} \\ \begin{array}{c} \mathbb{C}^M \\ \mathcal{L}^{\perp} \\ \mathcal{L} \end{array} \begin{pmatrix} A & B & 0 \\ C_1 & D_{11} & 0 \\ C_2 & D_{21} & D|_{\mathcal{L}}, \end{pmatrix} \end{array}$$

since $B$ maps $\mathcal{L}$ to 0 while $D$ maps $\mathcal{L}$ into itself. But since this is an isometry we must have $D|_{\mathcal{L}}$, a unitary which forces $C_2, D_{21} = 0$. This means $S$ is given by the TFR with isometry $\begin{pmatrix} A & B \\ C_1 & D_{11} \end{pmatrix}$. This has smaller size unless $\mathcal{L} = \{0\}$. $\qquad\square$

*Proof of Proposition 9.1.* If $S(z) = A + zB(I - zD)^{-1}C$ is a polynomial, then necessarily $BD^jC = 0$ for all $j$ large enough. By Proposition 9.2, $BD^n = 0$ for $n$ large enough. Then,

$$\text{range}(D^n) \subset \bigcap_{j \geq 0} \text{kernel}(BD^j),$$

implying range$(D^n) = 0$ or rather $D^n = 0$. $\qquad\square$

Minimality of TFR representations in the rational inner case in two variables makes it possible to prove an analogous result for inner matrix-valued polynomials in two variables. Our approach uses determinants to count the size of minimal TFRs. The following is a standard result in one variable. We provide a proof in Subsection A.3.

**Proposition 9.3.** *Let* $S : \mathbb{D} \to \mathbb{C}^{N \times N}$ *be a rational inner function. Then,* $\deg \det S$ *equals the size of a minimal TFR for* $S$.

Since $S$ is rational inner, $\det S$ is a scalar rational inner function in one variable which is a finite Blaschke product. Thus, the $\deg \det S$ refers to the degree of the numerator when written in lowest terms. This immediately yields a method using determinants to calculate the optimal size breakdown for rational inner functions in two variables. (This is another place where it helps to have square matrices.)

**Theorem 9.4 (Kummert).** *If $S : \mathbb{D}^2 \to \mathbb{C}^{N \times N}$ is rational inner, then the minimal size breakdown $r = (r_1, r_2)$ of a TFR for $S$ is*

$$r_j = \deg_j \det S(z_1, z_2) \quad \text{for } j = 1, 2.$$

Similarly, for all but finitely many $\zeta \in \mathbb{T}$, the degree of

$$z \mapsto \det S(z, \zeta z)$$

is $r_1 + r_2$. Therefore, the generic size of a TFR for $z \mapsto S(z, \zeta z)$ is $r_1 + r_2$. This shows that generic restrictions to slices of our two variable minimal TFRs yield minimal TFRs for restricted functions.

*Proof of Theorem 1.5.* The above argument shows that if a polynomial inner function $S$ has a minimal TFR via the unitary $U = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and projections $P_1, P_2$ as in Theorem 1.1, then $z \mapsto S(z, \zeta z)$ has minimal unitary TFR via the unitary

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & P_1 + \zeta P_2 \end{pmatrix}.$$

By Proposition 9.1, $D\Delta(1, \zeta)$ is nilpotent for all but finitely many $\zeta \in \mathbb{T}$. This means $(D\Delta(1, \zeta))^N = 0$ for all but finitely many $\zeta \in \mathbb{T}$. Since this is a polynomial equation we have $(D\Delta(1, \zeta))^N \equiv 0$, and since $D\Delta(z)$ is homogeneous we also have $(D\Delta(z_1, z_2))^N \equiv 0$. Thus, $D\Delta(z)$ is always nilpotent. $\square$

This leads to the interesting question of describing contractions $D$ such that $D\Delta(z)$ is nilpotent for all $z$. An easy way to produce examples would be to make $D$ strictly upper triangular and choose the projections $P_1, P_2$ via projections onto the span of subsets of standard basis vectors. For such examples, note that $D\Delta(z)$ is triangular; however, it is possible to produce matrices $D_1, D_2$ such that $z_1 D_1 + z_2 D_2$ is nilpotent for all $z$ yet is not triangularizable independent of $z$ (see [34]). This could be an interesting source of examples.

## APPENDIX A. AUXILIARY RESULTS

### 1.1. Maximum principle for rational iso-inner functions.

**Proposition A.1.** *Suppose $S : \mathbb{D}^d \to \mathbb{C}^{M \times N}$ is rational, analytic in $\mathbb{D}^d$, and $\|S(z)\| \le 1$ for $z \in \mathbb{T}^d$ where defined. Then, $\|S(z)\| \le 1$ for all $z \in \mathbb{D}^d$.*

Rationality is a key assumption since $f(z) = \exp((1 + z)/(1 - z))$ is unimodular on $\mathbb{T} \setminus \{1\}$ and analytic on $\mathbb{C} \setminus \{1\}$ yet not bounded by 1 in $\mathbb{D}$.

*Proof.* We can reduce to the scalar case by considering arbitrary unit vectors $v, w$ and the function $F(z) = w^* S(z) v$. Fix $\omega \in \mathbb{T}^d$ and consider the one-variable rational function $f(\zeta) = F(\zeta \omega)$. This function is bounded by 1 on $\mathbb{T}$ away from its potential finite number of poles. But, $f$ must be unbounded near a

pole, so any singularities on the boundary are removable. Hence, $f$ is analytic on $\bar{\mathbb{D}}$ and bounded by 1 by the maximum principle. This implies $F$ is bounded by 1 at any point of $r\mathbb{T}^d$ for $r < 1$. Given any $z \in \mathbb{D}^d$, we can calculate $F(z)$ as a Poisson integral of $F$ on $r\mathbb{T}^d$ for $\|z\|_\infty < r < 1$ to see that $|F(z)| \le 1$. $\quad\square$

**1.2. Fejér-Riesz proofs.** A more traditional and well-known version of the matrix Fejér-Riesz theorem is as follows. (See [16] for a proof.)

**Theorem A.2.** *Let $T(z) = \sum_{j=-n}^n T_j z^j$ be a matrix Laurent polynomial (with $T_j \in \mathbb{C}^{N\times N}$) such that $T(z) \ge 0$ for $z \in \mathbb{T}$ and $\det T(z)$ is not identically zero.*

*Then, there exists a matrix polynomial $A \in \mathbb{C}^{N\times N}[z]$ of degree at most $n$ such that $T = A^*A$ on $\mathbb{T}$ and $\det A(z) \ne 0$ for $z \in \mathbb{D}$.*

We think it is worthwhile to show how to go from this theorem to the degenerate version, Theorem 4.1, using ideas from [17]. The key tool is the Smith normal form.

**Theorem A.3 (Smith normal form).** *Let $P \in \mathbb{C}^{M\times N}[z]$ be a matrix polynomial. Then, there exist $T_1 \in \mathbb{C}^{M\times M}[z]$, $T_2 \in \mathbb{C}^{N\times N}[z]$ with matrix polynomial inverses (equivalently, with constant determinants) and $D \in \mathbb{C}^{M\times N}[z]$ such that $P = T_1 D T_2$. The matrix $D$ has the following form: every entry off the main diagonal of $D$ is zero and the main diagonal consists of polynomials $d_1, \ldots, d_k$ such that $d_j$ divides $d_{j+1}$. Here, $k = \min\{N, M\}$ and the $d_j$ may be zero for $j$ large enough.*

(See Hoffman-Kunze [28].)

*Proof of Theorem 4.1.* The function $G(z) = z^n T(z)$ is a polynomial matrix and therefore has Smith normal form decomposition

$$G(z) = T_1(z) \begin{pmatrix} D(z) & 0 \\ 0 & 0 \end{pmatrix} T_2(z).$$

Here, $T_1, T_2$ are matrix polynomials with matrix polynomial inverses while

$$D(z) = \text{diag}(d_1(z), \ldots, d_r(z))$$

is an $r \times r$ diagonal matrix with only non-zero polynomials on the diagonal. Notice that $T(z)$ has rank $r$ whenever $\det D(z) \ne 0$, $z \ne 0$. Since $T$ is self-adjoint on $\mathbb{T}$, we have $T(z) = T(1/\bar{z})^*$ for $z \ne 0$, and so

(A.1) $\quad T_2^{-1}(1/\bar{z})^* T(z) T_2^{-1}(z) = z^{-n} T_2^{-1}(1/\bar{z})^* T_1(z) \begin{pmatrix} D(z) & 0 \\ 0 & 0 \end{pmatrix}$

$$= z^n \begin{pmatrix} D(1/\bar{z})^* & 0 \\ 0 & 0 \end{pmatrix} T_1(1/\bar{z})^* T_2^{-1}(z)$$

is a matrix Laurent polynomial which is positive semi-definite on $\mathbb{T}$ and with 0 in the last $N - r$ columns and rows. Thus, (A.1) has the form $\begin{pmatrix} T_0(z) & 0 \\ 0 & 0 \end{pmatrix}$ where $T_0$

is an $r \times r$ matrix Laurent polynomial which is positive semi-definite on $\mathbb{T}$ and crucially satisfying $\det T_0 \not\equiv 0$ since $T$ has rank $r$ outside of a finite set.

By Theorem A.2, there is an $r \times r$ matrix polynomial $A_0$ so that $\det A_0(z) \neq 0$ in $\mathbb{D}$ and $A_0(z)^* A_0(z) = T_0(z)$ on $\mathbb{T}$. If we set $V = T_2$ and

$$A = \begin{pmatrix} A_0 & 0_{r \times (N-r)} \end{pmatrix} V,$$

then $A(z)^* A(z) = T(z)$ on $\mathbb{T}$. Note that $A(1/\bar{z})^* A(z) = T(z)$ holds in $\mathbb{C} \setminus \{0\}$ since both sides are analytic and agree on $\mathbb{T}$.

Our degree bound on $A$ follows from the fact that

$$z^n T(z) V(z)^{-1} \begin{pmatrix} A_0(z)^{-1} \\ 0 \end{pmatrix} = z^n A(1/\bar{z})^*$$

is analytic at $0$. A right rational inverse of $A$ is given by

$$V^{-1} \begin{pmatrix} A_0^{-1} \\ 0 \end{pmatrix}. \qquad \square$$

The matrix Fejér-Riesz factorization described is maximal in the sense of the following theorem. One can also describe all other factorizations. There is nothing essentially new about this result, but it is probably difficult to attribute. It could be deduced from inner-outer factorizations.

**Theorem A.4.** *Assuming the setup and notation of Theorem 4.1. For any other factorization $T = C^* C$ on $\mathbb{T}$ with a matrix polynomial $C$, there exists a rational iso-inner function $\Phi$ such that $C = \Phi A$ (necessarily, $\Phi = CB$). If $C$ has a right rational inverse holomorphic in $\mathbb{D}$, then $\Phi$ is a constant unitary matrix.*

*Proof.* Suppose $T = C^* C$ on $\mathbb{T}$. Then, we may write $CV^{-1} = \begin{pmatrix} C_0 & C_1 \end{pmatrix}$ where $C_0$ has $r$ columns. Since

$$(V^{-1})^* A^* A V^{-1} = \begin{pmatrix} A_0^* A_0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} C_0^* C_0 & C_0^* C_1 \\ C_1^* C_0 & C_1^* C_1 \end{pmatrix} = (V^{-1})^* C^* C V^{-1},$$

we see that $C_0^* C_0 = A_0^* A_0$, $C_1^* C_1 = 0$ on $\mathbb{T}$. This implies that $C_1 \equiv 0$. Then, $\Phi := C_0 A_0^{-1}$ is analytic on $\mathbb{D}$ and isometry valued on $\mathbb{T}$. Any poles on $\mathbb{T}$ are necessarily removable because $\Phi$ is rational and bounded on $\mathbb{T}$. We also have $\Phi A = C$. If $C$ has right rational inverse $C'$, then $\Phi A C' = I$. An isometry can only have a right inverse if it is square, so $\Phi$ must be square (hence unitary on $\mathbb{T}$) and $AC'$ must be unitary-valued on $\mathbb{T}$. By the maximum principle, $\Phi$ and $AC'$ are contractive in the disk; however, since they are inverses of each other they must be unitary-valued in the disk. Such analytic functions are constant. (Lemma 6.1 proves something more general than this.) $\qquad \square$

We now sketch a simple proof of Dritschel's positive definite multivariable Fejér-Riesz result (Theorem 6.3). Although it borrows elements from the original proof, we think it has some nice efficiencies in exposition.

*Proof of Theorem 6.3.* Let $n$ be a positive integer and define the multivariable Cesaro summation operator $C_n$ which we apply to $N \times N$ matrix Laurent polynomials $L(z) = \sum_{k \in \mathbb{Z}^d} L_k z^k$:

$$(C_n L)(z) = \sum_{k \in \mathbb{Z}^d} c_k^n L_k z^k = \int_{\mathbb{T}^d} F_n(z, \zeta) L(\zeta) \, d\sigma(\zeta)$$

where

$$c_k^n = \begin{cases} \prod_{j=1}^d \dfrac{n - |k_j|}{n} & \text{for } |k_1|, \ldots, |k_d| \leq n, \\ 0 & \text{otherwise,} \end{cases}$$

$$F_n(z, \zeta) = \frac{1}{n^d} \prod_{j=1}^d \left| \frac{1 - z_j^n \bar{\zeta}_j^n}{1 - z_j \bar{\zeta}_j} \right|^2 = \sum_{k \in \mathbb{Z}^d} c_k^n z^k$$

is the Fejér kernel and $d\sigma$ is normalized Lebesgue measure on $\mathbb{T}^d$.

Let $\mathcal{L}_m$ be the vector space of $N \times N$ Laurent polynomials of degree at most $m$ in each variable separately. We shall consider $C_n^m := C_n|_{\mathcal{L}_m} : \mathcal{L}_m \to \mathcal{L}_m$. By basic properties of Cesaro summation, $C_n^m L \to L$ uniformly on $\mathbb{T}^d$ as $n \to \infty$ for $L \in \mathcal{L}_m$. Since the set of linear operators $B(\mathcal{L}_m)$ on $\mathcal{L}_m$ is finite dimensional, $C_n^m$ tends to the identity as $n \to \infty$ with respect to any norm on $B(\mathcal{L}_m)$. In particular, for $n$ large enough, $C_n^m$ is invertible and $(C_n^m)^{-1}$ tends to the identity as $n \to \infty$.

We next point out that if $L \in \mathcal{L}_m$ is positive semi-definite on $\mathbb{T}^d$ then $C_n^m L$ is a sum of squares. The reason is that on $\mathbb{T}^d$, $F_n(z, \zeta) L(\zeta)$ is a Laurent polynomial of degree at most $n + m$ with respect to $\zeta$. Then, the integral representation of $C_n L$ can be computed via "quadrature." Indeed, for any $M$, if $H \in \mathcal{L}_M$ and $\mu = e^{2\pi i/(M+1)}$, then

$$\int_{\mathbb{T}^d} H(\zeta) \, d\sigma(\zeta) = \frac{1}{(M+1)^d} \sum_{0 \leq j_1, \ldots, j_d \leq M} H(\mu^{j_1}, \ldots, \mu^{j_d}).$$

This can be proven by testing on monomials. This means that $C_n L(z)$ is a positive finite linear combination of the terms

$$F_n(z, (\mu^{j_1}, \ldots, \mu^{j_d})) L(\mu^{j_1}, \ldots, \mu^{j_d}).$$

Since $F_n$ is evidently a squared polynomial and each value of $L$ on $\mathbb{T}^d$ is assumed positive semi-definite, we see that $C_n L$ is a sum of squares of polynomials.

Now, let $T \in \mathcal{L}^m$ be strictly positive on $\mathbb{T}^d$; that is, there exists $\delta > 0$ such that $T(z) \geq \delta I$ for $z \in \mathbb{T}^d$. For $n$ large enough, $T_n := (C_n^m)^{-1} T$ is also strictly

positive. Then, $T = C_n T_n$ is a Cesaro sum of a positive Laurent polynomial which was already shown to be a sum of squares. $\qquad\square$

**1.3. PSD kernels.** We now discuss the proof of Lemma 3.2 which claims that for $S : \mathbb{D} \to \mathbb{C}^{M \times N}$ analytic and $\|S(z)\| \le 1$ in $\mathbb{D}$, we have that

$$K_S(w, z) = \frac{I_N - S(w)^* S(z)}{1 - \bar{w}z}$$

is positive semi-definite (PSD). Let us recall the abstract definition of PSD for matrix or operator-valued kernels.

**Definition A.5.** Let $X$ be a set, let be $\mathcal{L}$ a complex Hilbert space, and let $K : X \times X \to B(\mathcal{L})$ be a function; here, $B(\mathcal{L})$ is the set of bounded linear self-maps of $\mathcal{L}$. We say that $K$ is a *PSD kernel* if for any $x_1, \dots, x_n \in X$ and $v_1, \dots, v_n \in \mathcal{L}$ we have

$$\sum_{i,j} \langle K(x_i, x_j) v_j, v_i \rangle \ge 0.$$

Notice that if $(x, y) \mapsto K(x, y)$ is a PSD kernel, then $(x, y) \mapsto K(y, x)$ is not necessarily PSD except in the scalar case $\mathcal{H} = \mathbb{C}$.

**Definition A.6.** The *rank* of $K$ is the maximum of the ranks of the block operators $(K(x_i, x_j))_{i,j}$ as we vary over $n$ and $x_1, \dots, x_n \in X$.

*Proof of Lemma 3.2.* Our proof uses rudiments of vector-valued Hardy spaces on the unit disk. (See Agler-McCarthy [4] for details.)

Let $H_M = H^2(\mathbb{D}) \otimes \mathbb{C}^M$ be the set of $M$-dimensional column vectors with entries in the Hardy space on the unit disk $H^2(\mathbb{D})$. Left multiplication by $S$, $M_S : H_N \to H_M$ is contractive. If $k_w(z) = k(z, w) := 1/(1 - \bar{w}z)$ is the Szegő kernel, then by a fundamental formula in reproducing kernel Hilbert space theory

$$M_S^*(k_w \otimes v) = k_w \otimes S(w)^* v \quad \text{for } v \in \mathbb{C}^M,$$

we see that

$$\langle (I - M_S M_S^*)(k_w \otimes v_1), k_z \otimes v_2 \rangle_{H_M} = k(z, w) \langle (I - S(z)S(w)^*) v_1, v_2 \rangle_{\mathbb{C}^M},$$

which after a short calculation using the fact that $I - M_S M_S^* \ge 0$ shows that $(z, w) \mapsto (I - S(z)S(w)^*)/(1 - z\bar{w})$ is PSD. We could apply the same argument to $\check{S}(z) := S(\bar{z})^*$ to see that $(z, w) \mapsto (I - S(\bar{z})^* S(\bar{w}))/(1 - z\bar{w})$ is PSD. Replace $z, w$ with their conjugates and relabel the variables to see that $K_S(w, z)$ is PSD. $\qquad\square$

*Proof of Proposition 9.3.* Assuming $S : \mathbb{D} \to \mathbb{C}^{N \times N}$ is rational inner we need to compute the rank of the positive semi-definite kernel

$$(w, z) \mapsto \frac{I - S(w)^* S(z)}{1 - \bar{w}z}.$$

We shall use notation from the proof of Lemma 3.2 above. As in the said proof, it is notationally easier to deal with the kernel

$$K(z, w) = \frac{I - S(z)S(w)^*}{1 - z\bar{w}},$$

and we can reduce to this case by replacing $S$ with $S(\bar{z})^*$.

Now, $K$ is the reproducing kernel for $H_N \ominus SH_N$. This follows from the fact that $S$ is inner: $SH_N$ is a closed subspace of $H_N$ and has reproducing kernel

$$\frac{S(z)S(w)^*}{1 - z\bar{w}},$$

which can be verified by the calculation

$$\langle Sf, k_w SS(w)^* v \rangle_{H_N} = \langle f, k_w S(w)^* v \rangle_{H_N} = \langle S(w)f(w), v \rangle_{\mathbb{C}^N}$$

for $f \in H_N$. The rank of $K$ is the dimension of $H_N \ominus SH_N$.

To count this dimension we write $S = Q/p$ in lowest terms. Since $S$ is bounded on $\mathbb{T}$ it can have no poles on $\mathbb{T}$, and therefore $p$ has no zeros in $\bar{\mathbb{D}}$. Let $Q(z) = T_1(z)D(z)T_2(z)$ be the Smith normal form decomposition for $Q$ (Theorem A.3 above). Notice that $D$ has full rank on $\mathbb{T}$ since $S$ is inner. Write $D = \text{diag}(d_1, \ldots, d_N)$. Then, $\det Q = c \det D = c \prod_j d_j$ where $c = \det T_1 \det T_2$ is a constant because $T_1, T_2$ have polynomial inverses. Since $S$ is inner, $\det S = \det Q/p^N$ is a finite Blaschke product. Its degree equals its number of zeros in $\mathbb{D}$, which equals the number of zeros of $\det Q$ in $\mathbb{D}$ since $p$ has none.

The vector space $H_N \ominus SH_N$ is isomorphic to the vector space quotient

$$H_N/SH_N = H_N/(T_1DT_2)H_N = H_N/(T_1DH_N) \cong H_N/DH_N.$$

The first equality holds because $p$ has no zeros in $\bar{\mathbb{D}}$, the second holds because $T_2$ has a polynomial inverse, and the last isomorphism holds because $T_1$ has a polynomial inverse. Recalling $D = \text{diag}(d_1, \ldots, d_N)$ we note the dimension of $H^2/d_j H^2$ is the number of zeros of $d_j$ in $\mathbb{D}$ and therefore the dimension of $H_N/DH_N$ is the number of zeros of $\prod_{j=1}^N d_j$ inside $\mathbb{D}$ (counting multiplicities). $\qquad\square$

This proof appears in [12].

## References

[1] JIM AGLER, *On the representation of certain holomorphic functions defined on a polydisc*, Topics in operator theory: Ernst D. Hellinger memorial volume, Oper. Theory Adv. Appl., vol. 48, Birkhäuser, Basel, 1990, pp. 47–66. MR1207393

[2] ———, *Some interpolation theorems of Nevanlinna-Pick type* (1988), unpublished.

[3] JIM AGLER and JOHN E. MCCARTHY, *Nevanlinna-Pick interpolation on the bidisk*, J. Reine Angew. Math. **506** (1999), 191–204. http://dx.doi.org/10.1515/crll.1999.004. MR1665697

[4] ———, *Pick Interpolation and Hilbert Function Spaces*, Graduate Studies in Mathematics, vol. 44, American Mathematical Society, Providence, RI, 2002. http://dx.doi.org/10.1090/gsm/044. MR1882259

[5] JIM AGLER, JOHN E. MCCARTHY, and N. J. YOUNG, *Operator monotone functions and Löwner functions of several variables*, Ann. of Math. (2) **176** (2012), no. 3, 1783–1826. http://dx.doi.org/10.4007/annals.2012.176.3.7. MR2979860

[6] ———, *A Carathéodory theorem for the bidisk via Hilbert space methods*, Math. Ann. **352** (2012), no. 3, 581–624. http://dx.doi.org/10.1007/s00208-011-0650-7. MR2885589

[7] T. ANDÔ, *On a pair of commutative contractions*, Acta Sci. Math. (Szeged) **24** (1963), 88–90. MR0155193

[8] JOSEPH A. BALL, *Multidimensional circuit synthesis and multivariable dilation theory*, Multidimens. Syst. Signal Process. **22** (2011), no. 1–3, 27–44. http://dx.doi.org/10.1007/s11045-010-0123-2. MR2771469

[9] JOSEPH A. BALL and DMITRY S. KALIUZHNYI-VERBOVETSKYI, *Rational Cayley inner Herglotz-Agler functions*: *Positive-kernel decompositions and transfer-function realizations*, Linear Algebra Appl. **456** (2014), 138–156. http://dx.doi.org/10.1016/j.laa.2013.10.022. MR3223895

[10] JOSEPH A. BALL and TAVAN T. TRENT, *Unitary colligations, reproducing kernel Hilbert spaces, and Nevanlinna-Pick interpolation in several variables*, J. Funct. Anal. **157** (1998), no. 1, 1–61. http://dx.doi.org/10.1006/jfan.1998.3278. MR1637941

[11] JOSEPH A. BALL, CORA SADOSKY, and VICTOR VINNIKOV, *Scattering systems with several evolutions and multidimensional input/state/output systems*, Integral Equations Operator Theory **52** (2005), no. 3, 323–393. http://dx.doi.org/10.1007/s00020-005-1351-y. MR2184571

[12] KELLY BICKEL and GREG KNESE, *Inner functions on the bidisk and associated Hilbert spaces*, J. Funct. Anal. **265** (2013), no. 11, 2753–2790. http://dx.doi.org/10.1016/j.jfa.2013.08.002. MR3096989

[13] MARIA CHARINA, MIHAI PUTINAR, CLAUS SCHEIDERER, and JOACHIM STÖCKLER, *An algebraic perspective on multivariate tight wavelet frames. II*, Appl. Comput. Harmon. Anal. **39** (2015), no. 2, 185–213. http://dx.doi.org/10.1016/j.acha.2014.09.003. MR3352013

[14] MARIA CHARINA, COSTANZA CONTI, MARIANTONIA COTRONEI, and MIHAI PUTINAR, *System theory and orthogonal multi-wavelets*, J. Approx. Theory **238** (2019), 85–102. http://dx.doi.org/10.1016/j.jat.2017.09.004. MR3912669

[15] BRIAN J. COLE and JOHN WERMER, *Ando's theorem and sums of squares*, Indiana Univ. Math. J. **48** (1999), no. 3, 767–791. http://dx.doi.org/10.1512/iumj.1999.48.1716. MR1736979

[16] PHILIPPE DELSARTE, Y. GENIN, and Y. KAMP, *A simple approach to spectral factorization*, IEEE Trans. Circuits and Systems **25** (1978), no. 11, 943–946. http://dx.doi.org/10.1109/TCS.1978.1084403. MR508983

[17] DRAGOMIR Ž. DJOKOVIĆ, *Hermitian matrices over polynomial rings*, J. Algebra **43** (1976), no. 2, 359–374. http://dx.doi.org/10.1016/0021-8693(76)90119-8. MR0437565

[18] MICHAEL A. DRITSCHEL, *On factorization of trigonometric polynomials*, Integral Equations Operator Theory **49** (2004), no. 1, 11–42. http://dx.doi.org/10.1007/s00020-002-1198-4. MR2057766

[19] ———, *Factoring non-negative operator-valued trigonometric polynomials in two variables* (2018), preprint, available at 1811.06005.

[20] JEFFREY S. GERONIMO and HUGO J. WOERDEMAN, *Positive extensions, Fejér-Riesz factorization and autoregressive filters in two variables*, Ann. of Math. (2) **160** (2004), no. 3, 839–906. http://dx.doi.org/10.4007/annals.2004.160.839. MR2144970

[21] JEFFREY S. GERONIMO, PLAMEN ILIEV, and GREG KNESE, *Polynomials with no zeros on a face of the bidisk*, J. Funct. Anal. **270** (2016), no. 9, 3505–3558. http://dx.doi.org/10.1016/j.jfa.2016.02.002. MR3475462

[22] ANATOLII GRINSHPAN, DMITRY S. KALIUZHNYI-VERBOVETSKYI, VICTOR VINNIKOV, and HUGO J. WOERDEMAN, *Classes of tuples of commuting contractions satisfying the multivariable von Neumann inequality*, J. Funct. Anal. **256** (2009), no. 9, 3035–3054. http://dx.doi.org/10.1016/j.jfa.2008.09.012. MR2502431

[23] ———, *Matrix-valued Hermitian Positivstellensatz, lurking contractions, and contractive determinantal representations of stable polynomials*, Operator theory, function spaces, and applications, Oper. Theory Adv. Appl., vol. 255, Birkhäuser/Springer, Cham, 2016, pp. 123–136. MR3617004

[24] ———, *Contractive determinantal representations of stable polynomials on a matrix polyball*, Math. Z. **283** (2016), no. 1–2, 25–37. http://dx.doi.org/10.1007/s00209-015-1587-4. MR3489057

[25] ———, *Stable and real-zero polynomials in two variables*, Multidimens. Syst. Signal Process. **27** (2016), no. 1, 1–26. http://dx.doi.org/10.1007/s11045-014-0286-3. MR3441374

[26] ———, *Rational inner functions on a square-matrix polyball*, Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory. Vol. 2, Assoc. Women Math. Ser., vol. 5, Springer, Cham, 2017, pp. 267–277. MR3688146

[27] YVAN HACHEZ and HUGO J. WOERDEMAN, *The Fischer-Frobenius transformation and outer factorization*, Operator Theory, Structured Matrices, and Dilations, Theta Ser. Adv. Math., vol. 7, Theta, Bucharest, 2007, pp. 181–203. MR2389624

[28] KENNETH HOFFMAN and RAY KUNZE, *Linear Algebra*, Second edition, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971. MR0276251

[29] GREG KNESE, *Polynomials defining distinguished varieties*, Trans. Amer. Math. Soc. **362** (2010), no. 11, 5635–5655. http://dx.doi.org/10.1090/S0002-9947-2010-05275-4. MR2661491

[30] ———, *Polynomials with no zeros on the bidisk*, Anal. PDE **3** (2010), no. 2, 109–149. http://dx.doi.org/10.2140/apde.2010.3.109. MR2657451

[31] ———, *Rational inner functions in the Schur-Agler class of the polydisk*, Publ. Mat. **55** (2011), no. 2, 343–357. http://dx.doi.org/10.5565/PUBLMAT_55211_04. MR2839446

[32] ———, *Schur-Agler class rational inner functions on the tridisk*, Proc. Amer. Math. Soc. **139** (2011), no. 11, 4063–4072. http://dx.doi.org/10.1090/S0002-9939-2011-10975-4. MR2823051

[33] ANTON KUMMERT, *Synthesis of two-dimensional lossless m-ports with prescribed scattering matrix*, Circuits Systems Signal Process. **8** (1989), no. 1, 97–119. http://dx.doi.org/10.1007/BF01598747. MR998029

[34] BEN MATHES, MATJAŽ OMLADIČ, and HEYDAR RADJAVI, *Linear spaces of nilpotent matrices*, Linear Algebra Appl. **149** (1991), 215–225. http://dx.doi.org/10.1016/0024-3795(91)90335-T. MR1092879

[35] VERN I. PAULSEN and MRINAL RAGHUPATHI, *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*, Cambridge Studies in Advanced Mathematics, vol. 152, Cambridge University Press, Cambridge, 2016.
http://dx.doi.org/10.1017/CBO9781316219232. MR3526117

[36] MURRAY ROSENBLATT, *A multi-dimensional prediction problem*, Ark. Mat. **3** (1958), 407–424.
http://dx.doi.org/10.1007/BF02589495. MR92332

[37] WALTER RUDIN, *Function Theory in Polydiscs*, W. A. Benjamin, Inc., New York-Amsterdam, 1969. MR0255841

[38] CLAUS SCHEIDERER, *Sums of squares on real algebraic surfaces*, Manuscripta Math. **119** (2006), no. 4, 395–410. http://dx.doi.org/10.1007/s00229-006-0630-5. MR2223624

[39] NICOLAS TH. VAROPOULOS, *On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory*, J. Functional Analysis **16** (1974), 83–100. MR0355642

Washington University in St. Louis
Department of Mathematics & Statistics
St. Louis, MO 63130, USA
E-MAIL: geknese@wustl.edu