A FAMILY OF DOMAINS ASSOCIATED WITH μ-SYNTHESIS

GAUTAM BHARALI

ABSTRACT. We introduce a family of domains — which we call the $\mu_{1,n}$-quotients — associated with an aspect of μ-synthesis. We show that the natural association that the symmetrized polydisc has with the corresponding spectral unit ball is also exhibited by the $\mu_{1,n}$-quotient and its associated unit "μ_E-ball". Here, μ_E is the structured singular value for the case $E = \{[w] \oplus (z\mathbb{I}_{n-1}) \in \mathbb{C}^{n\times n} : z, w \in \mathbb{C}\}$, $n = 2, 3, 4, \dots$ Specifically: we show that, for such an E, the Nevanlinna–Pick interpolation problem with matricial data in a unit "μ_E-ball", and in general position in a precise sense, is equivalent to a Nevanlinna–Pick interpolation problem for the associated $\mu_{1,n}$-quotient. Along the way, we present some characterizations for the $\mu_{1,n}$-quotients.

1. Introduction and Main Results

This article is devoted to studying the following infinite family of domains ($\mathbb{D}$ here will denote the open unit disc with centre $0 \in \mathbb{C}$):

$$\mathbb{E}_n := \Big\{(x,y) \in \mathbb{C}^n \times \mathbb{C}^{n-1} : \text{the zero set of } \Big(1 + \sum\nolimits_{j=1}^{n-1} (-1)^j y_j z^j \\ -w\Big(\sum\nolimits_{j=0}^{n-1} (-1)^j x_{j+1} z^j\Big)\Big) \text{ does not intersect } \overline{\mathbb{D}}^2\Big\},\ n \geq 2,$$

which we shall call the $\mu_{1,n}$*-quotients*. These domains are closely associated with an aspect of μ-synthesis. We will provide a couple of characterizations for $\mathbb{E}_n$ that make it easier to work with these domains (that each $\mathbb{E}_n$ is a domain is a classical argument; we defer this matter to Remark 3.6 below). The focus of this work, however, is to establish the connection between the $\mathbb{E}_n$'s and (the relevant aspect of) μ-synthesis.

μ-synthesis is a part of the theory of robust control of systems comprising interconnected electronic or mechanical devices each of whose outputs depend linearly on the inputs. Various performance measures are given by appropriate $\mathbb{R}_+$-homogeneous functionals on the space of matrices associated with such systems — see, for instance, [11]. The "μ" in μ-synthesis refers to such a class of cost functions. Fix $n \in \mathbb{Z}_+$, $n \geq 2$, and let E be a linear subspace of $\mathbb{C}^{n\times n}$. The functional

$$\mu_E(A) := \big(\inf\{\|X\| : X \in E \text{ and } (\mathbb{I} - AX) \text{ is singular}\}\big)^{-1},\ \ A \in \mathbb{C}^{n\times n},$$

is called a *structured singular value*. Here, $\|\cdot\|$ denotes the operator norm relative to the Euclidean norm on $\mathbb{C}^n$. Typically, the subspace E consists of all complex $n \times n$ matrices having a fixed block-diagonal structure. If $E = \mathbb{C}^{n\times n}$, then $\mu_E = \|\cdot\|$, while if E is the space of all scalar matrices, then μ_E is the spectral radius. The motivation for, and the definition of, μ_E comes from the theory of efficient stabilization of systems in which the uncertainties in their governing parameters are highly structured: the subspace E is meant to encode the structure of the perturbations to such systems.

In much the same way that a necessary condition for desigining a controller that stabilizes the aforementioned system (with unstructured uncertainties) is the existence of an interpolant for certain Nevanlinna–Pick data with values in the unit $\|\cdot\|$-ball — see

1991 *Mathematics Subject Classification.* Primary: 30E05, 32F45; Secondary: 47A56, 93D21.
Key words and phrases. Categorical quotients, interpolation, Nevanlinna–Pick problem, μ-synthesis.
This work is supported in part by a Centre for Advanced Study grant.

[10, Chapter 4], for instance — with *structured* uncertainties one needs to understand the Nevanlinna–Pick interpolation problem for the unit "μ_E-ball" for a given E.

At this juncture, we shift our focus entirely to the Nevanlinna–Pick interpolation problem. We refer readers (who aren't already familiar) to the pioneering work of John Doyle [9] for the control-theory motivations behind μ_E. With E as above, let $\Omega_E := \{W \in \mathbb{C}^{n\times n} : \mu_E(W) < 1\}$. The Nevanlinna–Pick interpolation problem for Ω_E is the following:

($*$) *Given M distinct points $\zeta_1, \dots, \zeta_M \in \mathbb{D}$ and matrices $W_1, \dots, W_M$ in Ω_E, find necessary conditions and sufficient conditions on $\{(\zeta_1, W_1), \dots, (\zeta_M, W_M)\}$ for the existence of a holomorphic map $F : \mathbb{D} \longrightarrow \Omega_E$ satisfying $F(\zeta_j) = W_j,\ j = 1, \dots, M$.*

When E is the class of all scalar matrices in $\mathbb{C}^{n\times n}$, Ω_E is the so-called spectral unit ball, which we denote by Ω_n. The problem ($*$) has been studied intensively for Ω_n. Bercovici *et al.* [6] have given a characterization for the interpolation data $\{(\zeta_1, W_1), \dots, (\zeta_M, W_M)\}$ to admit an Ω_n-valued interpolant. However, this characterization involves a non-trivial search over a region in $\mathbb{C}^{n^2M}$. Thus, there is interest in finding alternative characterizations that would at least reduce the dimension of the search-region: see, for instance, [4, 5]. This was one of the motivations behind the ideas in the paper [2] by Agler & Young, wherein they introduced the symmetrized bidisc. Its n-dimensional analogue (the symmetrized polydisc, denoted by $\mathbb{G}_n$) was introduced by Costara in [8]. The importance of $\mathbb{G}_n$ to μ-synthesis is as follows:

a) $\dim(\Omega_n) \gg \dim(\mathbb{G}_n)$, yet, whenever the matrices $W_1, W_2, \dots, W_M \in \Omega_n$ lie off an explicitly defined set $\mathfrak{S}_n \subsetneq \Omega_n$, which is of zero Lebesgue measure, the problem ($*$) is equivalent to an associated Nevanlinna–Pick problem for $\mathbb{G}_n$.

(Also see [12] for an improvement of (*a*) when $n = 2, 3$.)

For most of the systems alluded to above, the associated E comprises matrices whose diagonal blocks are either scalar matrices or rank-one matrices. We address here the next level of complexity in the block structure of E. The domains $\mathbb{E}_n$, $n \geq 2$, introduced above are the analogues of the symmetrized polydiscs $\mathbb{G}_n$ when (for a fixed $n \geq 2$)

$$E \equiv E^{1,n} := \{[w] \oplus (z\mathbb{I}_{n-1}) : z, w \in \mathbb{C}\} \tag{1.1}$$

(here, $\mathbb{I}_{n-1}$ denotes the $(n-1)\times(n-1)$ identity matrix). Theorem 1.2 below is precisely the statement (*a*) with the domains Ω_E, for the above choice of E, replacing Ω_n. For this choice of E, we shall denote Ω_E as $\Omega_{1,n}$.

The feature (*a*) is not the only useful insight that $\mathbb{G}_n$ brings to the Nevanlinna–Pick problem on Ω_n. The set $\mathfrak{S}_n$ (which we have not defined; but see [8, Theorem 2.1]) helps explain certain subtleties of the interpolation problem. We shall elaborate upon these after stating Theorem 1.2, but we mention here that the preceding remark motivates our explicit description of the set $\widetilde{\Omega_{1,n}}$ — the analogue of $(\Omega_n \setminus \mathfrak{S}_n)$ for $\Omega_{1,n}$ — in Theorem 1.2. It is also important to mention that a special case of our domains $\mathbb{E}_n$ is the tetrablock. It was introduced by Abouhajar *et al.* [1] and is the domain $\mathbb{E}_2$.

To describe $\widetilde{\Omega_{1,n}}$, we shall need the following:

Definition 1.1. A matrix $A \in \mathbb{C}^{n\times n}$ is said to be *non-derogatory* if A admits a cyclic vector. Therefore, A being non-derogatory is equivalent to A being similar to the companion matrix of its characteristic polynomial — i.e., if $z^n + \sum_{j=1}^n s_j z^{n-j}$ denotes the characteristic polynomial, then

$$A \text{ is non-derogatory} \iff A \text{ is similar to} \begin{bmatrix} 0 & & & -s_n \\ 1 & 0 & \mathbf{0} & -s_{n-1} \\ & \ddots & \ddots & \vdots \\ \mathbf{0} & & 1 & -s_1 \end{bmatrix}_{n\times n}.$$

We shall make use of some notations throughout this work. For a matrix $A \in \mathbb{C}^{n\times n}$, $A^\bullet$ will denote the $(n-1)\times(n-1)$ matrix obtained by deleting the first row and column of A. For any pair of integers $m \leq n$, $[m\,..\,n]$ will denote the integer subset $\{m, m+1, \dots, n\}$. Assume that $n \in \mathbb{Z}_+$ is fixed; for any j such that $1 \leq j \leq n$, $\mathscr{I}^j$ will denote the set of all increasing j-tuples in $[1\,..\,n]^j$. Finally, for n and j as described, for $I \in \mathscr{I}^j$, and for any $A \in \mathbb{C}^{n\times n}$, A_I will denote the $j \times j$ submatrix of A whose rows and columns are indexed by I. Having defined these notations, we can state our first result:

Theorem 1.2. *Let $n \geq 2$, write any $A \in \mathbb{C}^{n\times n}$ as $A = [a_{j,k}]$, and let $\Omega_{1,n}$ be as defined above. Define:*

$$\widetilde{\Omega_{1,n}} := \big\{A \in \Omega_{1,n} : A^\bullet \text{ is non-derogatory, and } (a_{2,1}, \dots, a_{n,1}) \text{ is a cyclic vector of } A^\bullet\big\}.$$

Then:

1) $\Omega_{1,n} \setminus \widetilde{\Omega_{1,n}}\ (=: \mathfrak{S}_{1,n})$ has zero Lebesgue measure.
2) Define the map $\pi_n : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{2n-1}$ by

$$\pi_n(A) := \Bigg(a_{1,1}, \sum_{I\in\mathscr{I}^2 : i_1=1} \det(A_I), \dots, \sum_{I\in\mathscr{I}^n : i_1=1} \det(A_I); \sum_{I\in\mathscr{I}^1 : i_1\geq 2} \det(A_I), \dots, \sum_{I\in\mathscr{I}^{n-1} : i_1\geq 2} \det(A_I)\Bigg).$$

π_n is holomorphic and maps $\Omega_{1,n}$ onto $\mathbb{E}_n$.
3) Let $\zeta_1, \dots, \zeta_M$ be distinct points in $\mathbb{D}$ and let $W_1, \dots, W_M$ belong to $\widetilde{\Omega_{1,n}}$. Then, there exists a holomorphic map $F : \mathbb{D} \longrightarrow \Omega_{1,n}$ satisfying $F(\zeta_j) = W_j$ for every $j \leq M$ if and only if there exists a holomorphic map $f : \mathbb{D} \longrightarrow \mathbb{E}_n$ satisfying $f(\zeta_j) = \pi_n(W_j)$ for every $j \leq M$.

Engineers have had some success in numerically computing solutions to the problem $(*)$. These methods are based on iterative schemes that are supported by convincing, but largely *heuristic*, arguments. However, we now know that the problem $(*)$ is ill-conditioned in a specific sense. The set $\mathfrak{S}_n$ in (a) (and its analogue $\mathfrak{S}_{1,n}$, given by Theorem 1.2) gives us a precise description of this problem:

- b) (following [3, Example 2.3] by Agler–Young) There exist continuous one-parameter families of Nevanlinna–Pick data $\{(\zeta_1, W_{1,\alpha}), (\zeta_2, W_{2,\alpha})\}_{\alpha\in\mathbb{D}}$ with $(W_{1,\alpha}, W_{2,\alpha}) \in (\Omega_n \setminus \mathfrak{S}_n)^2$ $\forall \alpha \neq 0$ such that there exist Ω_n-valued interpolants $\forall \alpha \neq 0$, but **none** for $\alpha = 0$. In this case, either $W_{1,0} \in \mathfrak{S}_n$ or $W_{2,0} \in \mathfrak{S}_n$.

This provides useful information for testing the stability of some of the numerical algorithms used. It is the information that (b) provides that is our second motivation for constructing analogues of $\mathbb{G}_n$ for the case of $E^{1,n}$.

Indeed, Abouhajar *et al.* have shown [1, Remark 9.5-(iv)] that the problem $(*)$ for $\Omega_{1,2}$ is also ill-conditioned, exactly as described in (b) with $\Omega_{1,2}$ replacing Ω_n therein. This pathology extends to $\Omega_{1,n}$ for all $n \geq 2$. It turns out that, analogous to (b), the problem lies in either $W_{1,0}$ or $W_{2,0}$ belonging to $\mathfrak{S}_{1,n}$ (as defined in Theorem 1.2-(1)). In fact, it is [1, Remark 9.5-(iv)] that led us to intuit what $\mathfrak{S}_{1,n}$ must be for general n.

Our second main result provides a necessary condition for the existence of an interpolant that solves the problem $(*)$ for $E^{1,n}$. For this, we must give some definitions. For each

$(x,y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$, let us define:

$$\begin{aligned} P_n(z;x) &:= \sum\nolimits_{j=0}^{n-1} (-1)^j x_{j+1} z^j, \\ Q_n(z;y) &:= 1 + \sum\nolimits_{j=1}^{n-1} (-1)^j y_j z^j, \quad z \in \mathbb{C}, \\ \Psi_n(z;x,y) &:= \begin{cases} \dfrac{P_n(z;x)}{Q_n(z;y)} \\ \text{(with the understanding that, in evaluating } \Psi_n(\cdot\,;x,y)\text{, any} \\ \text{common linear factors of } P_n(\cdot\,;x) \text{ and } Q_n(\cdot\,;y) \text{ are first cancelled).} \end{cases} \end{aligned}$$

With these definitions, we can state our next result.

Theorem 1.3. *Let $\zeta_1, \dots, \zeta_M$ be distinct points in $\mathbb{D}$ and let $W_1, \dots, W_M$ in $\Omega_{1,n}$, $n \geq 2$. Express the map π_n as $\pi_n = (X,Y) : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^n \times \mathbb{C}^{n-1}$. If there exists a holomorphic map $F : \mathbb{D} \longrightarrow \Omega_{1,n}$ satisfying $F(\zeta_j) = W_j$ for every j, then, for each $z \in \overline{\mathbb{D}}$, the matrix*

$$M_z := \left[\frac{1 - \overline{\Psi_n(z; X(W_j), Y(W_j))}\, \Psi_n(z; X(W_k), Y(W_k))}{1 - \overline{\zeta_j}\zeta_k} \right]_{j,k=1}^{M}$$

is positive semi-definite.

Remark 1.4. Implicit in the statement of Theorem 1.3 is that if $(x,y) \in \mathbb{E}_n$, then the rational function $\Psi_n(\cdot\,;x,y)$ has no poles in $\mathbb{D}$. In fact, much more can be said about $\Psi_n(\cdot\,;x,y)$, as we shall see in Section 3.

Theorem 1.3 is an easy corollary to a certain characterization of the set $\mathbb{E}_n$ in terms of the functions $\Psi_n(\cdot\,;x,y)$. It also turns out that the sets $\mathbb{E}_n$, $n \geq 2$, form a certain hierarchy in the sense that membership in $\mathbb{E}_{n+1}$ can be characterized in terms of membership in $\mathbb{E}_n$, $n \geq 2$. The precise results (Theorems 3.3 and 3.5) will be presented in Section 3.

We ought to state that the theorems presented in this section address only a *small part* of what control engineers need. The chief utility to engineers is that, in view of (b) above and the paragraph that follows it, the set $(\Omega_{1,n} \setminus \widetilde{\Omega_{1,n}})$ raises a very specific flag in testing numerical methods for constructing Nevanlinna–Pick interpolants that rely on limit processes. The question arises: given that, in real-world stabilization problems (with structured uncertainties) one encounters other forms of the space E, what can one say about Theorems 1.2 and 1.3 for general E? We make some remarks on this issue, and on the subject of *categorical quotients*—of which the reader gets a *very* fleeting glimpse in Section 2—in Section 4 (Remarks 4.1 and 4.2) below.

2. A Few Preliminary Lemmas

This section is devoted to a few lemmas that we will need in the subsequent sections.

In the following lemma, we shall follow the notation introduced in Section 1 and the standard multi-index notation. A diagonal $n \times n$ matrix having the number a_j as the entry in its jth row and column will be denoted by $\mathrm{diag}(a_1, \dots, a_n)$.

Lemma 2.1. *Fix an integer $n \geq 2$ and let $A \in \mathbb{C}^{n\times n}$. Then:*

$$\det\left(\mathbb{I}_n - A\,\mathrm{diag}(z_1, \dots, z_n)\right) = 1 + \sum_{j=1}^{n} (-1)^j \sum_{I \in \mathscr{I}^j} \det(A_I) z^I. \tag{2.1}$$

Proof. Let us denote the matrix on the left-hand side above by B. As usual, we write $A = [a_{j,k}]$ and S_n for the group of permutations of n distinct objects. We write down the classical expansion of $\det(B^T)$ to see that

$$\begin{aligned}\det(B) &= \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{k=1}^{n}\left(\boldsymbol{\delta}_{\sigma(k),\,k} - a_{\sigma(k),\,k}\,z_k\right)\\ &= 1+\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\sum_{j=1}^{n}(-1)^j\left[\sum_{I\in\mathscr{I}^j}\left(\prod\nolimits_{s\in I}a_{\sigma(s),\,s}\right)\left(\prod\nolimits_{t\in I^{\mathsf{C}}}\boldsymbol{\delta}_{\sigma(t),\,t}\right)z^I\right], \qquad (2.2)\end{aligned}$$

where I^{C} is the abbreviation for $[1\,..\,n]\setminus I$, and with the understanding that a product indexed by the null set equals 1. Clearly, the second product on the right-hand side of (2.2) is non-zero if and only if σ fixes the subset I^{C}. For any subset $J\subseteq[1\,..\,n]$, write

$$\mathsf{Fix}(J) := \{\sigma\in S_n : \sigma \text{ fixes } J\}.$$

Then, from (2.2), we get

$$\det(B) = 1+\sum_{j=1}^{n}(-1)^j\sum_{I\in\mathscr{I}^j}\left(\sum\nolimits_{\sigma\in\mathsf{Fix}(I^{\mathsf{C}})}\operatorname{sgn}(\sigma)\prod\nolimits_{s\in I}a_{\sigma(s),\,s}\right)z^I.$$

Given the definition of the submatrices A_I, the above identity is precisely (2.1). $\square$

For the next lemma, we present a convention that we will follow in this article. The notation $\mathbb{C}^*\oplus GL_{n-1}(\mathbb{C})$, $n\geq 2$, will denote the group (with respect to matrix multiplication) of $n\times n$ matrices G that are block-diagonal, with the $(1,1)$-entry of G being a non-zero complex number and $G^\bullet\in GL_{n-1}(\mathbb{C})$.

Lemma 2.2. *Let $(x,y)\in\mathbb{C}^n\times\mathbb{C}^{n-1}$, $n\geq 2$, and let $\pi_n:\mathbb{C}^{n\times n}\longrightarrow\mathbb{C}^n\times\mathbb{C}^{n-1}$ be the map defined in Theorem 1.2. Let $A\in\pi_n^{-1}\{(x,y)\}$. Then, the congugacy orbit*

$$\{G^{-1}AG : G\in\mathbb{C}^*\oplus GL_{n-1}(\mathbb{C})\} \subseteq \pi_n^{-1}\{(x,y)\}.$$

Proof. We shall denote the conjugacy orbit $\{G^{-1}AG : G\in\mathbb{C}^*\oplus GL_{n-1}(\mathbb{C})\}$ as O_A. It suffices to show that π_n is constant on O_A. As in Section 1, we write $\pi_n=(X,Y)$. We will denote any element $G\in\mathbb{C}^*\oplus GL_{n-1}(\mathbb{C})$ as $g\oplus\Gamma$: g being the $(1,1)$-entry of G, and $G^\bullet=\Gamma$. It is a classical fact that, by definition:

$$Y(G^{-1}AG) = \left(\mathscr{S}_{n-1,1}(\sigma(\Gamma^{-1}A^\bullet\Gamma)),\dots,\mathscr{S}_{n-1,\,n-1}(\sigma(\Gamma^{-1}A^\bullet\Gamma))\right) \qquad (2.3)$$

where

$$\begin{aligned}\mathscr{S}_{n-1,\,j} &:= \text{the } j\text{-th elementary symmetric polynomial in } n-1 \text{ indeterminates},\\ \sigma(B) &:= \text{the list of eigenvalues of the matrix } B\text{, listed according to multiplicity.}\end{aligned}$$

As $\mathscr{S}_{n-1,\,j}$ is a similarity invariant, (2.3) implies that Y is constant on O_A.

Therefore, it suffices to show that X is constant on O_A. For any $G=g\oplus\Gamma$ as above, and $j=2,\dots,n$, we compute:

$$\begin{aligned}\sum_{I\in\mathscr{I}^j}\det(A_I) &= \sum_{I\in\mathscr{I}^j}\det((G^{-1}AG)_I)\\ &= \sum_{I\in\mathscr{I}^j\,:\,i_1=1}\det((G^{-1}AG)_I)+\sum_{I\in\mathscr{I}^j\,:\,i_1\geq 2}\det((G^{-1}AG)_I)\\ &= \sum_{I\in\mathscr{I}^j\,:\,i_1=1}\det((G^{-1}AG)_I)+\mathscr{S}_{n-1,\,j}(\sigma(\Gamma^{-1}A^\bullet\Gamma)). \qquad (2.4)\end{aligned}$$

The left-hand side of (2.4) is a constant. Therefore, it follows from (2.4) that the j-th component of the map $X:\mathbb{C}^{n\times n}\longrightarrow\mathbb{C}^n$, $j=2,\dots n$, is constant on O_A. And, of course, the $(1,1)$ entry of $(g\oplus\Gamma)^{-1}A(g\oplus\Gamma)$ does not vary with G. Hence the lemma. $\square$

The next two lemmas will be essential to the proof of Theorem 1.2.

Lemma 2.3. *Let $(x,y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$, $n \geq 2$. There exist polynomials $p_k \in \mathbb{C}[x,y]$, $k = 1,\dots,(n-1)$, such that, if we define*

$$B(x,y) := \left[\begin{array}{c|cccc} x_1 & p_1(x,y) & \dots & p_{n-2}(x,y) & p_{n-1}(x,y) \\ \hline 1 & 0 & & & (-1)^n y_{n-1} \\ 0 & 1 & 0 & \mathbf{0} & (-1)^{n-1} y_{n-2} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \mathbf{0} & & 1 & y_1 \end{array}\right],$$

then, for each $j = 2,\dots,n$,

$$\sum_{I \in \mathscr{I}^j \,:\, i_1 = 1} \det(B(x,y)_I) \;=\; x_j.$$

Furthermore, for a given (x,y), $p_1(x,y),\dots p_{n-1}(x,y)$ are the unique *numbers for which the above equations hold true.*

Proof. Let $\mathbf{B}$ be the $n \times n$ matrix obtained by replacing the entries $p_k(x,y)$ by the unknowns Z_k, $k = 1,\dots,(n-1)$, in the matrix $B(x,y)$ given above. We shall need some auxiliary objects. First, given a vector $w \in \mathbb{C}^n$, for each integer $m \in [1\,..\,n]$, let us define the matrices

$$M(m;w,y) := \left[\begin{array}{c|cccc} w_{n+1-m} & w_{n+2-m} & \dots & w_{n-1} & w_n \\ \hline 1 & 0 & & & (-1)^m y_{m-1} \\ 0 & 1 & 0 & \mathbf{0} & (-1)^{m-1} y_{m-2} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \mathbf{0} & & 1 & y_1 \end{array}\right]_{m\times m}.$$

For m as above, we shall write:

$$\begin{aligned} {}^m\mathscr{I}^j \;&:=\; \text{the set of all increasing } j\text{-tuples in } [1\,..\,m],\ 1 \leq j \leq m, \\ {}^m\mathscr{I}^j(1) \;&:=\; \{I \in {}^m\mathscr{I}^j : i_1 = 1\},\ 1 \leq j \leq m, \\ {}^m\mathscr{I}^j(1,2) \;&:=\; \{I \in {}^m\mathscr{I}^j : i_1 = 1,\ i_2 = 2\},\ 2 \leq j \leq m. \end{aligned}$$

Finally, we shall define, for m as above, and $1 \leq k \leq m$,

$$\Phi(k,m;w,y) \;:=\; \sum_{I \in {}^m\mathscr{I}^k(1)} \det(M(m;w,y)_I).$$

We begin with an elementary observation. Suppose, *for the moment*, $n \geq 4$. Then, for (m,k) such that $3 \leq k \leq m-1$, we have

$$\begin{aligned} \Phi(k,m;w,y) \;&=\; \sum_{I \in {}^m\mathscr{I}^k(1,2)} \det(M(m;w,y)_I) + \sum_{I \in {}^m\mathscr{I}^k(1)\setminus {}^m\mathscr{I}^k(1,2)} \det(M(m;w,y)_I) \\ &=\; -\Phi(k-1,m-1;w,y) + \sum_{I \in {}^m\mathscr{I}^k(1)\setminus {}^m\mathscr{I}^k(1,2)} \det(M(m;w,y)_I). \end{aligned} \tag{2.5}$$

This follows by expanding each determinant in the first sum with respect to its first column and from the fact that, as $3 \leq k \leq m-1$, the $(1,1)$-cofactor of each submatrix $M(m;w,y)_I$, $I \in {}^m\mathscr{I}^k(1,2)$, has at least one zero-column. As for $\Phi(2,m;w,y)$, it is easy to see, owing to the structure of $M(m;w,y)^\bullet$, that

$$\Phi(2,m;w,y) \;=\; -w_{n+2-m} + y_1 w_{n+1-m}. \tag{2.6}$$

It is possible to simplify the second sum in the equation (2.5) further. We argue along the lines described just after (2.5): we expand each determinant with respect to its first column. However, there is a difference in this case. The $(1,1)$-cofactor of each relevant $M(m;w,y)_I$ will have a zero-column *except* when $I = (1, m-k+2, \dots, m)$. Note that, as

$k \leq m-1$, $m-k+2 \neq 2$. The $(1,1)$-cofactor of $M(m;w,y)_{(1,m-k+2,\dots,m)}$ is the companion matrix of the polynomial $X^{k-1} - y_1 X^{k-2} + \dots + (-1)^{k-2} y_{k-2} X + (-1)^{k-1} y_{k-1}$. Thus:

$$\sum_{I \in {}^m\mathscr{I}^k(1) \setminus {}^m\mathscr{I}^k(1,2)} \det(M(m;w,y)_I) \;=\; w_{n+1-m} y_{k-1}.$$

Combining this with (2.5), we get

$$\Phi(k,m;w,y) \;=\; -\Phi(k-1,m-1;w,y) + w_{n+1-m} y_{k-1}, \quad 3 \leq k \leq m-1. \tag{2.7}$$

The conclusions of the lemma can easily be established for $n = 2,3$ (we leave it to the reader to check this). We shall establish the lemma for $n \geq 4$. Recall the definition of the matrix $\mathbf{B}$. Treating $(Z_1, \dots, Z_{n-1})$ as unknowns, the following:

$$\sum_{I \in \mathscr{I}^j \,:\, i_1 = 1} \det(\mathbf{B}_I) \;=\; x_j, \quad j = 2, \dots, n, \tag{2.8}$$

is a system of $(n-1)$ algebraic equations in $(n-1)$ unknowns.

Observe that the matrix $\mathbf{B}$ is the matrix $M(n;w,y)$ with $w = (x_1, Z_1, \dots, Z_{n-1})$. Thus, taking $w = (x_1, Z_1, \dots, Z_{n-1})$ in (2.6) and (2.7) and applying (2.7) recursively, we see that the system (2.8) is a lower-triangular system of linear equations in $(Z_1, \dots, Z_{n-1})$. From the recursion relation (2.7), we get that the coefficient of the unknown Z_j in the j-th equation of (2.8) (which concerns the sum of the $(j+1)$-st principal minors of $\mathbf{B}$) is $(-1)^j$, $1 \leq j \leq (n-2)$. Finally, expanding $\det(\mathbf{B})$ along the first *row*, we see that the coefficient of Z_{n-1} in the last equation of (2.8) is $(-1)^{n-1}$. It follows from Cramer's rule that each Z_j is a polynomial p_j in (x,y). By our definition of $\mathbf{B}$, these polynomials, $p_1, \dots, p_{n-1}$, are the required polynomials. The uniqueness statement follows from the fact that, for a fixed (x,y), the system (2.8) has a unique solution $\square$

We continue to follow the notation presented just before the statement of Theorem 1.2. Further notation: if S is a square matrix, then $\mathbf{C}_S$ will denote the companion matrix of its characteristic polynomial (normalized as in Definition 1.1).

Lemma 2.4. *Fix an integer $n \geq 2$, and write any $A \in \mathbb{C}^{n \times n}$ as $A = [a_{j,k}]$. Define*

$$\begin{aligned} \mathscr{G}_{1,n} \;:=\; \big\{ A \in \mathbb{C}^{n \times n} : A^\bullet \text{ is non-derogatory, and} \\ (a_{2,1}, \dots, a_{n,1}) \text{ is a cyclic vector of } A^\bullet \big\}. \end{aligned}$$

Let $A, B \in \mathscr{G}_{1,n}$. Suppose $A, B \in \pi_n^{-1}\{(x,y)\}$ for some $(x,y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$. If $A^\bullet = B^\bullet$ and $(a_{2,1}, \dots, a_{n,1}) = (b_{2,1}, \dots, b_{n,1})$, then $(a_{1,2}, \dots, a_{1,n}) = (b_{1,2}, \dots, b_{1,n})$.

Remark 2.5. In the proof of the above lemma — as elsewhere in this article — a vector in $\mathbb{C}^k$, $1 \leq k < \infty$, will also be treated (without any change in notation) as a $k \times 1$ complex matrix.

Proof. By assumption, $A^\bullet$ is non-derogatory. It is well-known that any matrix $G \in GL_{n-1}(\mathbb{C})$ such that $G^{-1} A^\bullet G = \mathbf{C}_{A^\bullet}$ must be of the form.

$$G \;=\; [\boldsymbol{c} \;\; A^\bullet \boldsymbol{c} \, \dots \, (A^\bullet)^{n-2} \boldsymbol{c}],$$

where $\boldsymbol{c}$ is some cyclic vector of $A^\bullet$. Thus, the matrix

$$\Gamma \;:=\; [\mathbf{a} \;\; A^\bullet \mathbf{a} \, \dots \, (A^\bullet)^{n-2} \mathbf{a}], \tag{2.9}$$

where $\mathbf{a} := [a_{2,1} \dots a_{n,1}]^T$, is the *unique* element in $GL_{n-1}(\mathbb{C})$ with the two properties

$$\begin{aligned} \Gamma^{-1} A^\bullet \, \Gamma \;&=\; \mathbf{C}_{A^\bullet}, \\ \Gamma \, [1 \; 0, \, \dots \, 0]^T \;&=\; \mathbf{a}. \end{aligned}$$

We will denote elements $X \in \mathbb{C}^* \oplus GL_{n-1}(\mathbb{C})$ using the abbreviated notation introduced in the proof of Lemma 2.2. By what we have just discussed:

$$(1 \oplus \Gamma)^{-1} A (1 \oplus \Gamma) \;=\; \left[\begin{array}{c|c} a_{1,1} & [a_{1,2} \;\; a_{1,3} \; \dots \; a_{1,n}]\Gamma \\ \hline 1 & \\ 0 & \\ \vdots & \mathsf{C}_{A^\bullet} \\ 0 & \end{array}\right].$$

Call the above matrix $\widehat{A}$. By Lemma 2.2, $\widehat{A} \in \pi_n^{-1}\{(x,y)\}$. Thus, by Lemma 2.3, it follows — compare the matrix above with the matrix $B(x,y)$ in Lemma 2.3 — that

$$[a_{1,2}, \;\; a_{1,3} \; \dots \; a_{1,n}] \;=\; [p_1(x,y) \;\; p_2(x,y) \; \dots \; p_{n-1}(x,y)]\Gamma^{-1}.$$

However, the argument above applies to B as well, and as $A^\bullet = B^\bullet$ and $(a_{2,1}, \dots, a_{n,1}) = (b_{2,1}, \dots, b_{n,1})$, the matrix Γ given by (2.9) works for B as well. And as $A, B \in \pi_n^{-1}\{(x,y)\}$, we can conclude that

$$\begin{aligned} [a_{1,2}, \;\; a_{1,3} \; \dots \; a_{1,n}] \;&=\; [p_1(x,y) \;\; p_2(x,y) \; \dots \; p_{n-1}(x,y)]\Gamma^{-1} \\ &=\; [b_{1,2}, \;\; b_{1,3} \; \dots \; b_{1,n}]. \end{aligned}$$

□

3. Two Characterizations of $\mathbb{E}_n$

As hinted in Section 1, Theorem 1.3 follows from a certain characterization of $\mathbb{E}_n$. This characterization is the focus of this section. We begin with a proposition that explains the origins of the (somewhat odd-looking) sets $\mathbb{E}_n$. Readers familiar with [1] will notice that the following proposition is a generalization of [1, Theorem 9.1].

Proposition 3.1. *A point $(x,y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$ belongs to $\mathbb{E}_n$ if and only if there exists a matrix $A \in \Omega_{1,n}$ such that $\pi_n(A) = (x,y)$. Furthermore, if $(x,y) \in \mathbb{E}_n$, then the matrix $B(x,y)$ defined in the statement of Lemma 2.3 belongs to $\Omega_{1,n}$.*

Remark 3.2. The first part of the above is, essentially, part (2) of Theorem 1.2.

Proof. Let $E^{1,n}$ be as in (1.1). Given $r > 0$ and a matrix $A \in \mathbb{C}^{n\times n}$, $\mu_{E^{1,n}}(A) \leq 1/r$ if and only if for, any matrix $M \in E^{1,n}$ that satisfies

$$\det(\mathbb{I} - AM) \;=\; 0,$$

$\|M\| \geq r$. Let us write π_n as $(X,Y) : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^n \times \mathbb{C}^{n-1}$. It follows from Lemma 2.1 that if the M above is written as $M = [w] \oplus (z\mathbb{I}_{n-1})$, then

$$\det(\mathbb{I} - AM) \;=\; \Big(1 + \sum_{j=1}^{n-1} (-1)^j Y_j(A) z^j\Big) - w\left(\sum_{j=0}^{n-1} (-1)^j X_{j+1}(A) z^j\right). \tag{3.1}$$

The preceding discussion is summarized as follows:

- (•) $\mu_{E^{1,n}}(A) \leq 1/r$, $r > 0$, if and only if the zero set of the polynomial on the right-hand side of (3.1) is disjoint from $(r\mathbb{D})^2$.

Now, suppose $A \in \Omega_{1,n}$. Then there exists an $r_0 > 1$ such that $\mu_{E^{1,n}}(A) \leq 1/r_0$. It follows from (•) that the zero set of the polynomial on the right-hand side of (3.1) is disjoint from $(r_0\mathbb{D})^2$, whence it is disjoint from $\overline{\mathbb{D}}^2$. Thus $(X,Y)(A) = \pi_n(A) \in \mathbb{E}_n$.

Let $(x,y) \in \mathbb{E}_n$. Let $p_1, \dots, p_{n-1}$ be the polynomials provided by Lemma 2.3 and let A be the matrix $B(x,y)$ defined in Lemma 2.3. Since $A^\bullet$ is a companion matrix, it follows by examination of its last column that $Y(A) = (y_1, \dots, y_{n-1})$. Thus, from the definition of π_n and by Lemma 2.3, we have

$$\pi_n(A) = (x,y). \tag{3.2}$$

As $(x, y) \in \mathbb{E}_n$, it follows that there exists a small positive constant ε_0 such that the zero set of the polynomial

$$\left(1+\sum\nolimits_{j=1}^{n-1}(-1)^j y_j z^j\right) - w\left(\sum\nolimits_{j=0}^{n-1}(-1)^j x_{j+1} z^j\right)$$

is disjoint from $((1+\varepsilon_0)\mathbb{D})^2$. From ($\bullet$) and (3.2), we have $\mu_{E^{1,n}}(A) \leq 1/(1+\varepsilon_0) < 1$. This completes the proof. $\square$

The first theorem of this section is a consequence of Proposition 3.1. In order to state it, we need a definition. Fix an integer $n \geq 2$ and let $(x, y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$. Let $P_n(\cdot\,; x)$ and $Q_n(\cdot\,; y)$ be the polynomials defined just prior to Theorem 1.3, and define

$$\mathscr{R}_n(x, y) \;:=\; \mathsf{Res}(P_n(\cdot\,; x), Q_n(\cdot\,; y)),$$

where Res denotes the resultant of a pair of univariate polynomials.

Theorem 3.3. *Fix an integer $n \geq 2$, and, for $(x, y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$, let $\Psi_n(\cdot\,; x, y)$ be the rational function defined in Section 1. The point $(x, y) \in \mathbb{E}_n$ if and only if the following two conditions are satisfied:*

(I) *$\Psi_n(\cdot\,; x, y)|_{\overline{\mathbb{D}}} \in \mathcal{O}(\mathbb{D}) \cap \mathcal{C}(\overline{\mathbb{D}})$, and*

$$\sup_{z\in\partial\mathbb{D}} |\Psi_n(z; x, y)| \;<\; 1;$$

(II) *If $\mathscr{R}_n(x, y) = 0$, then every common zero of $P_n(\cdot\,; x)$ and $Q_n(\cdot\,; y)$ lies outside $\overline{\mathbb{D}}$.*

Proof. In this proof, for any polynomial $p \in \mathbb{C}[z, w]$, $\mathbf{Z}(p)$ will denote its zero set in $\mathbb{C}^2$. Let us fix $(x, y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$ and write:

$${}^n p(z, w; x, y) \;:=\; Q_n(z; y) - wP_n(z; x).$$

We will begin with some basic observations. First:

$$z_0 \text{ is a common zero of } P_n(\cdot\,; x) \text{ and } Q_n(\cdot\,; y) \iff \{z_0\} \times \mathbb{C} \subset \mathbf{Z}({}^n p(\cdot\,; x, y)), \tag{3.3}$$

$$P_n(z_0; x) = 0 \text{ and } Q_n(z_0; y) \neq 0 \iff (\{z_0\} \times \mathbb{C}) \cap \mathbf{Z}({}^n p(\cdot\,; x, y)) = \varnothing. \tag{3.4}$$

Secondly: in view of (3.3) and (3.4), it follows that for any $z_0 \in \mathbb{C}$:

$$\begin{aligned} &(\{z_0\} \times \mathbb{C}) \cap \mathbf{Z}({}^n p(\cdot\,; x, y)) \subset \{z_0\} \times (\mathbb{C} \setminus \overline{\mathbb{D}}) \\ &\qquad\Longrightarrow \begin{cases} \{z_0\} \text{ is not a common zero of } P_n(\cdot\,; x) \text{ and } Q_n(\cdot\,; y), \\ Q_n(z_0; y) \neq 0, \text{ and} \\ (\{z_0\} \times \mathbb{C}) \cap \mathbf{Z}(\widetilde{{}^n p}(\cdot\,; x, y)) \subset \{z_0\} \times \mathbb{D}, \end{cases} \end{aligned} \tag{3.5}$$

where the polynomial $\widetilde{{}^n p}(\cdot\; x, y)$ is defined by

$$\widetilde{{}^n p}(z, w; x, y) \;:=\; wQ_n(z; y) - P_n(z; x).$$

Claim. *For any z_0, the converse of* (3.5) *holds true.*
To see this, let us abbreviate the statement (3.5) as $\mathcal{P}(z_0) \implies \mathcal{Q}(z_0)$. Now fix a z_0 and suppose that it satisfies the three conditions in $\mathcal{Q}(z_0)$. If $P(z_0; x) = 0$, then by (3.4) $\mathcal{P}(z_0)$ is vacuously true. Hence, let us assume that $P_n(z_0; x) \neq 0$. Then:

$$(\{z_0\} \times \mathbb{C}) \cap \mathbf{Z}(\widetilde{{}^n p}(\cdot\,; x, y)) \;=\; \{(z_0, P_n(z_0; x)/Q_n(z_0; y))\} \;\equiv\; \{(z_0, w_0)\}$$

and, by assumption, $0 < |w_0| < 1$. Thus $(\{z_0\} \times \mathbb{C}) \cap \mathbf{Z}({}^n p(\cdot\,; x, y)) = \{(z_0, 1/w_0)\} \subset \{z_0\} \times (\mathbb{C} \setminus \overline{\mathbb{D}})$. This establishes the claim.

The condition for membership of (x, y) in $\mathbb{E}_n$ can be stated as:

$$(x, y) \in \mathbb{E}_n \iff \text{for each } z \in \overline{\mathbb{D}},\ (\{z\} \times \mathbb{C}) \cap \mathbf{Z}({}^n p(\cdot\,; x, y)) \subset \{z\} \times (\mathbb{C} \setminus \overline{\mathbb{D}}).$$

In view of (3.3), (3.5) and its converse, and (3.4), the above statement is rephrased as:

$$\begin{aligned}(x,y)\in\mathbb{E}_n \iff & \text{ for each } z\in\overline{\mathbb{D}},\\ & z \text{ is not a common zero of } P_n(\,\boldsymbol{\cdot}\,;x) \text{ and } Q_n(\,\boldsymbol{\cdot}\,;y),\\ & Q_n(z;y)\neq 0, \text{ and } |P_n(z;x)/Q_n(z;y)|<1. \end{aligned} \tag{3.6}$$

Finally, we make use the following two facts. First: for any fixed (x,y), the polynomials $P_n(\,\boldsymbol{\cdot}\,;x)$ and $Q_n(\,\boldsymbol{\cdot}\,;y)$ have a common zero if and only if $\mathscr{R}_n(x,y)=0$—see, for instance, [15]. Second: since $\Psi_n(\,\boldsymbol{\cdot}\,;x,y)$ (as defined in Section 1) is a **rational** function,

$$\Psi_n(\,\boldsymbol{\cdot}\,;x,y)\in\mathcal{C}(\overline{\mathbb{D}}) \iff \Psi_n(\,\boldsymbol{\cdot}\,;x,y) \text{ is bounded on } \mathbb{D}.$$

In view of these two facts, the theorem follows from (3.6) after an application of the Maximum Modulus Theorem. □

For our next theorem we shall need the following result by Costara:

Result 3.4 (Costara, [8], Corollary 3.4)**.** *For any $(s_1,\dots,s_n)\in\mathbb{C}^n$, $n\geq 2$, the following assertions are equivalent:*

(*i*) *The element $(s_1,\dots,s_n)$ belongs to the symmetrized polydisc $\mathbb{G}_n$.*

(*ii*) *For each $z\in\overline{\mathbb{D}}$, $(\widetilde{s}_1(z),\dots,\widetilde{s}_{n-1}(z))\in\mathbb{G}_{n-1}$, where*

$$\widetilde{s}_j(z) := n^{-1}\frac{(n-j)s_j-(j+1)zs_{j+1}}{1-n^{-1}zs_1},\quad j=1,\dots,n-1.$$

As in [8], implicit in the phrase "$(s_1,\dots,s_n)\in\mathbb{G}_n$" is the sign-convention of the definition:

$$\mathbb{G}_n := \left\{(s_1,\dots,s_n)\in\mathbb{C}^n : \text{the roots of } z^n+\sum\nolimits_{j=1}^n(-1)^js_jz^{n-j}=0 \text{ lie in } \mathbb{D}\right\},\ n\in\mathbb{Z}_+.$$

Theorem 3.5. *For any $(x,y)\in\mathbb{C}^n\times\mathbb{C}^{n-1}$, $n\geq 3$, the following assertions are equivalent:*

(*i*) *The point (x,y) belongs to the $\mu_{1,n}$-quotient $\mathbb{E}_n$.*

(*ii*) *For each $\xi\in\overline{\mathbb{D}}$, the point $(\widetilde{x}(\xi),\widetilde{y}(\xi))\in\mathbb{E}_{n-1}$, where*

$$\begin{aligned}\widetilde{x}_j(\xi) &:= \frac{(n-j)x_j-j\xi x_{j+1}}{(n-1)-\xi y_1},\quad j=1,\dots,n-1,\\ \widetilde{y}_j(\xi) &:= \frac{(n-1-j)y_j-(j+1)\xi y_{j+1}}{(n-1)-\xi y_1},\quad j=1,\dots,n-2.\end{aligned}$$

Proof. Fix an integer $N\geq 2$ (this N is unrelated to the n in the theorem above). For $(x,y)\in\mathbb{C}^N\times\mathbb{C}^{N-1}$, let $P_N(\,\boldsymbol{\cdot}\,;x)$ and $Q_N(\,\boldsymbol{\cdot}\,;y)$ be as in the proof of the previous theorem. Note that the following statements are equivalent:

a) The point $(x,y)\in\mathbb{C}^N\times\mathbb{C}^{N-1}$ belongs to $\mathbb{E}_N$.

b) For each fixed $w\in\overline{\mathbb{D}}$, the zeros of the polynomial $(Q_N(z;y)-wP_N(z;x))$ lie in $(\mathbb{C}\setminus\overline{\mathbb{D}})$.

c) For each fixed $w\in\overline{\mathbb{D}}$, the zeros of the polynomial $(Q_N(z;y)-wP_N(z;x))$ lie in $(\mathbb{C}\setminus\overline{\mathbb{D}})$ and $(1-wx_1)\neq 0$.

d) For each fixed $w\in\overline{\mathbb{D}}$, the zeros of the polynomial

$$\frac{z^{N-1}}{1-wx_1}\left(Q_N\left(\frac{1}{z};y\right)-wP_N\left(\frac{1}{z};x\right)\right) = z^{N-1}+\sum_{j=1}^{N-1}(-1)^j\frac{y_j-wx_{j+1}}{1-wx_1}z^{N-(j+1)}$$

lie in $\mathbb{D}$.

e) For each fixed $w\in\overline{\mathbb{D}}$,

$$\left(\frac{y_1-wx_2}{1-wx_1},\dots,\frac{y_{N-1}-wx_N}{1-wx_1}\right)\in\mathbb{G}_{N-1}. \tag{3.7}$$

Except, perhaps, for the implication $(b) \Rightarrow (c)$, it is either self-evident or follows from definitions that each statement in the above list is equivalent to the one that follows it. As for the implication $(b)\Rightarrow(c)$: it follows from (b) that if the polynomial in (b) is nonconstant, then the product of its zeros must be non zero, and if it is constant (for a fixed $w \in \overline{\mathbb{D}}$), then this constant must be non-zero. In either case, this gives $(1 - wx_1) \neq 0$.

Now consider $n \geq 3$ as given. From the equivalence $(a) \iff (e)$ with $N = n$, and from Costara's theorem, we get:

(▲) The point $(x, y) \in \mathbb{E}_n \iff$ for each $(w, \xi) \in (\overline{\mathbb{D}})^2$, $(\widetilde{s}_1(\xi, w; x, y), \ldots, \widetilde{s}_{n-2}(\xi, w; x, y))$ belongs to $\mathbb{G}_{n-2}$, where

$$\widetilde{s}_j(\xi, w; x, y) := \frac{\left(\dfrac{n-j-1}{n-1}\right)\left(\dfrac{y_j - wx_{j+1}}{1 - wx_1}\right) - \xi\left(\dfrac{j+1}{n-1}\right)\left(\dfrac{y_{j+1} - wx_{j+2}}{1 - wx_1}\right)}{1 - \left(\dfrac{\xi}{n-1}\right)\left(\dfrac{y_1 - wx_2}{1 - wx_1}\right)}, \quad j = 1, \ldots, n-2.$$

Observe that the expressions for $\widetilde{s}_j(\xi, w; x, y)$ can be rewritten as

$$\widetilde{s}_j(\xi, w; x, y) := \frac{\left[\dfrac{(n-1-j)y_j - (j+1)\xi y_{j+1}}{(n-1) - \xi y_1}\right] - w\left[\dfrac{(n-j-1)x_{j+1} - (j+1)\xi x_{j+2}}{(n-1) - \xi y_1}\right]}{1 - w\left[\dfrac{(n-1)x_1 - \xi x_2}{(n-1) - \xi y_1}\right]}, \quad j = 1, \ldots, n-2. \tag{3.8}$$

For $N \geq 2$, it follows from the equivalence $(a) \iff (e)$ that we established above that (just take $w = 0$ in (3.7))

$$(x, y) \in \mathbb{E}_N \implies (y_1, \ldots, y_{N-1}) \in \mathbb{G}_{N-1} \implies |y_1| < N - 1.$$

From this, we get

$$(x, y) \in \mathbb{E}_N \iff \big((x, y) \in \mathbb{E}_N \text{ and, for each } w \in \overline{\mathbb{D}},\ (1 - wy_1) \neq 0\big). \tag{3.9}$$

We now apply the equivalence $(a) \iff (e)$ taking $N = (n - 1)$ (which is valid, since, by hypothesis, $(n - 1) \geq 2$). From (3.9), the equivalence (▲), and by comparing (3.8) with (3.7), we see that for any $(x, y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$, each assertion in the list below is equivalent to the one following it:

A) The point $(x, y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$ belongs to $\mathbb{E}_n$.
B) The point $(x, y) \in \mathbb{C}^n \times \mathbb{C}^{n-1}$ belongs to $\mathbb{E}_n$ and, for each $w \in \overline{\mathbb{D}}$, $(1 - wy_1) \neq 0$.
C) For each $w \in \overline{\mathbb{D}}$, $(1 - wy_1) \neq 0$ and, for each $\xi \in \overline{\mathbb{D}}$, $(\widetilde{s}_1(\xi, w; x, y), \ldots, \widetilde{s}_{n-2}(\xi, w; x, y))$ belongs to $\mathbb{G}_{n-2}$, where $\widetilde{s}_j(\xi, w; x, y)$, $j = 1, \ldots, (n - 2)$, is given by (▲).
D) The assertion (ii) in the statement of Theorem 3.5.

This completes the proof. □

Remark 3.6. In Section 1, we mentioned that the sets $\mathbb{E}_n$ are domains. That each $\mathbb{E}_n$, $n \geq 2$, is open can be established by a classical argument. It can be deduced from the fact that the condition defining $\mathbb{E}_n$ is an open condition; that $\mathbb{C}$-affine algebraic hypersurfaces of degree $\leq n$ vary continuously — in an appropriate sense; see [7, Chapter 1, § 1.2] — with respect to the coefficients of their defining functions; and that the varieties occurring in the definition of $\mathbb{E}_n$ have a rather simple form. **However**, the proof of the previous theorem provides a slick way of establishing the openness of $\mathbb{E}_n$. Fix an $n \geq 2$ and let $(x_0, y_0) \in \mathbb{E}_n$. By the implication $(a) \Rightarrow (e)$ in the above proof, we get, for each $w \in \overline{\mathbb{D}}$:

$$(Z_1^w, \ldots, Z_{n-1}^w) := \left(\frac{y_{0,1} - wx_{0,2}}{1 - wx_{0,1}}, \ldots, \frac{y_{0,\,n-1} - wx_{0,\,n}}{1 - wx_{0,1}}\right) \in \mathbb{G}_{n-1}.$$

As $\mathbb{G}_{n-1}$ is open and w varies through a compact set, there exists an $\varepsilon > 0$ such that the polydiscs $\boldsymbol{\Delta}_w := D(Z_1^w;\varepsilon) \times \cdots \times D(Z_{n-1}^w;\varepsilon) \subset \mathbb{G}_{n-1}$. By the implication $(a) \Rightarrow (c)$, the set $(-x_{0,1}\overline{\mathbb{D}} + 1) \not\ni 0$, whence by equicontinuity we can find a $\delta > 0$ such that

$$\left(\frac{y_1 - wx_2}{1 - wx_1}, \dots, \frac{y_{n-1} - wx_n}{1 - wx_1}\right) \in \boldsymbol{\Delta}_w \quad \forall (x,y) \in \mathbb{B}^{2n-1}((x_0, y_0);\delta)$$

(where $\mathbb{B}^N(a;r)$ denotes the open Euclidean ball centered at $a \in \mathbb{C}^N$ of radius r), for *each* $w \in \overline{\mathbb{D}}$. This time, by the implication $(e) \Rightarrow (a)$, we get $\mathbb{B}^{2n-1}((x_0,y_0);\delta) \subset \mathbb{E}_n$. It follows that $\mathbb{E}_n$ is open. The connectedness of $\mathbb{E}_n$ is a consequence of part (2) of Theorem 1.2.

4. Proofs of the Main Theorems

The results of the last two sections provide us all the tools needed to prove Theorems 1.2 and 1.3.

The proof of Theorem 1.2. We begin by reminding the reader of the notational comment in Remark 2.5. Recall further: if S is a square matrix, then C_S will denote the companion matrix of its characteristic polynomial (normalized as in Definition 1.1).

1) Let Λ denote the holomorphic identification $\Lambda : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{(n-1)\times(n-1)} \times \mathbb{C}^{n-1} \times \mathbb{C}^n$,

$$\Lambda(A) := \big(A^\bullet,\, (a_{j,1})_{2\le j\le n},\, (a_{1,k})_{1\le k\le n}\big),$$

writing $A = [a_{j,k}]$. Define

$$\begin{aligned}\mathscr{N} &:= \{X \in \mathbb{C}^{(n-1)\times(n-1)} : X \text{ is non-derogatory}\},\\ \mathfrak{S}^1 &:= (\mathbb{C}^{(n-1)\times(n-1)} \setminus \mathscr{N}) \times \mathbb{C}^{n-1} \times \mathbb{C}^n.\end{aligned}$$

Define the function $\Theta : \mathscr{N} \times \mathbb{C}^{n-1} \longrightarrow \mathbb{C}$ as follows:

$$\Theta(X,v) := \det\big([v \;\; Xv \, \dots \, X^{n-2}v]\big). \tag{4.1}$$

Fix some $X^0 \in \mathscr{N}$. As X^0 is non-derogatory, it has a cyclic vector: call it $\boldsymbol{c}_{X^0}$. Clearly, $\Theta(X^0, \boldsymbol{c}_{X^0}) \neq 0$, whence $\Theta(X^0;\,\boldsymbol{\cdot}) \not\equiv 0$, and this is true for any $X^0 \in \mathscr{N}$. By construction, $\Theta(X;\,\boldsymbol{\cdot})$ and Θ are holomorphic functions. Since $\Theta(X;\,\boldsymbol{\cdot}) \not\equiv 0$ (for $X \in \mathscr{N}$) and $\Theta \not\equiv 0$, it is a classical result — see, for instance, [13, Theorem 14.4.9] — that

$$\begin{gathered}\Theta(X;\,\boldsymbol{\cdot})^{-1}\{0\} \times \mathbb{C}^n \subsetneq (\{X\} \times \mathbb{C}^{n-1} \times \mathbb{C}^n) \text{ has zero } ((4n-2)\text{-dim'l.})\\ \text{Lebesgue measure (for each } X \in \mathscr{N}),\\ \Theta^{-1}\{0\} \times \mathbb{C}^n \subsetneq (\mathscr{N} \times \mathbb{C}^{n-1} \times \mathbb{C}^n) \text{ has zero } (2n^2\text{-dim'l.}) \text{ Lebesgue measure.}\end{gathered} \tag{4.2}$$

Note that, for a matrix $X \in \mathscr{N}$, $\Theta(X,v) \neq 0 \iff v$ is a cyclic vector of X. Hence, writing $\Theta^{-1}\{0\} \times \mathbb{C}^n =: \mathfrak{S}^2$, we get

$$\widetilde{\Omega_{1,n}} = \Omega_{1,n} \cap \Lambda^{-1}((\mathscr{N} \times \mathbb{C}^{n-1} \times \mathbb{C}^n) \setminus \mathfrak{S}^2). \tag{4.3}$$

Since $\Lambda^{-1}(\mathfrak{S}^1)$ has zero ($2n^2$-dimensional) Lebesgue measure, it follows from (4.2) and (4.3) that $(\Omega_{1,n} \setminus \widetilde{\Omega_{1,n}})$ has zero Lebesgue measure.

2) Part (2) is essentially the first part of Proposition 3.1. That π_n is holomorphic is trivial as it is a polynomial map.

3) If there exists a holomorphic map $F : \mathbb{D} \longrightarrow \Omega_{1,n}$ that interpolates the given data, then, by part (2), $f := \pi_n \circ F$ has the required properties.

Let us now assume that there exists a holomorphic map $f : \mathbb{D} \longrightarrow \mathbb{E}_n$ such that $f(\zeta_j) = \pi_n(W_j)$ for every j. Let us write $f = (\mathsf{x}, \mathsf{y})$, where $\mathsf{x} =: (\mathsf{x}_1, \dots, \mathsf{x}_n) : \mathbb{D} \longrightarrow \mathbb{C}^n$ and

$\mathsf{y} =: (\mathsf{y}_1, \dots, \mathsf{y}_{n-1}) : \mathbb{D} \longrightarrow \mathbb{C}^{n-1}$. Let $p_1, \dots p_{n-1}$ be the polynomials given by Lemma 2.3. Define the holomorphic map $\phi : \mathbb{D} \longrightarrow \mathbb{C}^{n\times n}$ as follows:

$$\phi(\zeta) := \left[\begin{array}{c|cccc} \mathsf{x}_1(\zeta) & p_1\circ f(\zeta) & \dots & p_{n-2}\circ f(\zeta) & p_{n-1}\circ f(\zeta) \\ \hline 1 & 0 & & & (-1)^n\mathsf{y}_{n-1}(\zeta) \\ 0 & 1 & 0 & \mathbf{0} & (-1)^{n-1}\mathsf{y}_{n-2}(\zeta) \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \mathbf{0} & & 1 & \mathsf{y}_1(\zeta) \end{array}\right]. \tag{4.4}$$

Note that, in the notation of Lemma 2.3, $\phi = B(\mathbf{x}, \mathbf{y})$. Hence, it follows from the second assertion in Proposition 3.1 that $\phi : \mathbb{D} \longrightarrow \Omega_{1,n}$. And it follows from Lemma 2.3 that

$$\pi_n \circ \phi(\zeta_j) = \pi_n(W_j), \quad j = 1, \dots, M. \tag{4.5}$$

The above ϕ is not, in general, the desired F (although the range of ϕ *is* contained in $\Omega_{1,n}$). We must now address this problem. Let $\mathbb{C}^* \oplus GL_{n-1}(\mathbb{C})$ be as introduced just before the statement of Lemma 2.2. The importance of this group to our discussion is the following simple (but powerful):

Fact. *For a matrix $A \in \mathbb{C}^{n\times n}$, $\mu_{E^{1,n}}(A) = \mu_{E^{1,n}}(G^{-1}AG)$ for each $G \in \mathbb{C}^* \oplus GL_{n-1}(\mathbb{C})$.*

So, the idea behind what follows is to construct an appropriate holomorphic $(\mathbb{C}^*\oplus GL_{n-1}(\mathbb{C}))$-valued map ψ, defined on $\mathbb{D}$, such that $F := \psi^{-1}\phi\,\psi$ is the desired interpolant.

To this end, we point out that by (4.5) and by the definition of the map π_n, we get

$$\phi(\zeta_j)^{\bullet} = \mathsf{C}_{W_j^{\bullet}}, \quad j = 1, \dots, M. \tag{4.6}$$

Now refer to the proof of Lemma 2.4. By the fact that $W_1, \dots, W_M \in \widetilde{\Omega_{1,n}}$, there exists a unique matrix $\Gamma_j \in GL_{n-1}(\mathbb{C})$, $j = 1, \dots, M$, such that

$$\Gamma_j^{-1}W_j^{\bullet}\Gamma_j = \mathsf{C}_{W_j^{\bullet}}, \qquad \Gamma_j\,[1\ 0\ \dots\ 0]^T = [{}^jw_{2,1}\ \ {}^jw_{3,1}\ \dots\ {}^jw_{n,1}]^T, \quad j = 1, \dots, M, \tag{4.7}$$

where we write $W_j = [{}^jw_{i,k}]$ for each $j \leq M$. At this point, we know two things:

- by examining (4.4), $(1 \oplus \Gamma_j)\phi(\zeta_j)(1 \oplus \Gamma_j)^{-1}$ belongs to $\mathscr{G}_{1,n}$;
- by the above observation, (4.6), and (4.7), if we set $A := (1 \oplus \Gamma_j)\phi(\zeta_j)(1 \oplus \Gamma_j)^{-1}$ and $B := W_j$, then this choice of (A, B) satisfies the hypothesis of Lemma 2.4;

for each $j = 1, \dots, M$. Here, we have used the abbreviated notation, introduced in Section 2, for an element in $\mathbb{C}^* \oplus GL_{n-1}(\mathbb{C})$. Therefore, Lemma 2.4 tells us:

$$\text{The first row of } (1 \oplus \Gamma_j)\phi(\zeta_j)(1 \oplus \Gamma_j)^{-1} \text{ equals the first row of } W_j \text{ for each } j \leq M. \tag{4.8}$$

As each Γ_j above is an invertible matrix, there exists a matrix $L_j \in \mathbb{C}^{(n-1)\times(n-1)}$ such that $\exp(L_j) = \Gamma_j$. Let $\Psi : \mathbb{D} \longrightarrow \mathbb{C}^{(n-1)\times(n-1)}$ be any matrix-valued holomorphic function such that $\Psi(\zeta_j) = L_j$, $j = 1, \dots M$. Now, let us define the following $(\mathbb{C}^*\oplus GL_{n-1}(\mathbb{C}))$-valued holomorphic map:

$$\psi(\zeta) := 1 \oplus e^{-\Psi(\zeta)} \quad \forall \zeta \in \mathbb{D}.$$

Since we have shown that $\phi(\zeta) \in \Omega_{1,n}$ $\forall \zeta \in \mathbb{D}$, it follows from the Fact stated above that:

$$\psi(\zeta)^{-1}\phi(\zeta)\psi(\zeta) \in \Omega_{1,n} \quad \forall \zeta \in \mathbb{D}.$$

We now write $F := \psi^{-1}\phi\psi$. We have just argued that $F : \mathbb{D} \longrightarrow \Omega_{1,n}$ and is holomorphic. From (4.4), (4.6), (4.7), and (4.8), it follows that this F is the desired interpolant. $\square$

The ideas used in the above proof lead to some observations that would be relevant when dealing with the unit "μ_E-balls" when E is of greater complexity.

Remark 4.1. Probably the most important role in the proof of Theorem 1.2 was played by the fact that the group $\mathbb{C}^* \oplus GL_{n-1}(\mathbb{C})$ acts on $\Omega_{1,n}$. A more abstract look into the relationship between this action and the domain $\mathbb{E}_n$ might suggest the way forward in formulating analogues of Theorem 1.2 for more general cases of E. For both the pairs $(\Omega_n, \mathbb{G}_n)$ and $(\Omega_{1,n}, \mathbb{E}_n)$, $n \geq 2$, it turns out that the relationship of the lower-dimensional domain to its associated unit "μ_E-ball" is analogous to the categorical quotient associated to an affine algebraic variety with a reductive group acting on it. We say "analogous" because $\Omega_{1,n}$ is not an algebraic variety. But there *are* settings — see [14] by Snow, for instance — to which the constructions of classical geometric invariant theory carry over. In this work, owing to the nature of the "structural space" $E^{1,n}$, we did not need to appeal to the abstract theory (which still needs some enhancements to Snow's work). However, in that language, the components of the map π_n are the generators of the ring of G-invariant functions, $\mathbb{E}_n$ is the analogue of the categorical quotient, and $\widetilde{\Omega_{1,n}}$ is the union of all **closed** G-orbits of $\Omega_{1,n}$, $G = \mathbb{C}^* \oplus GL_{n-1}(\mathbb{C})$. For a general E, Lemma 2.1 will *still* give us the generators of the ring of G-invariant functions on Ω_E (for an appropriate G). However, when E is of much greater complexity, the abstract viewpoint hinted at might be helpful in determining the analogue of the set $\widetilde{\Omega_{1,n}}$ without engaging in ever more complex computations.

We now come to the proof of Theorem 1.3. This proof is an easy consequence of part (2) of the previous theorem and Theorem 3.3.

The proof of Theorem 1.3. In view of Theorem 1.2-(2), the function $\pi_n \circ F$ is a holomorphic map and $\pi_n \circ F(\mathbb{D}) \subset \mathbb{E}_n$. Thus, for each $\zeta \in \mathbb{D}$, $\pi_n \circ F(\zeta) \in \mathbb{E}_n$. It follows from this, from condition (I) in Theorem 3.3, and the Maximum Modulus Theorem that if we *fix* a point $z \in \overline{\mathbb{D}}$, then

$$|\Psi_n(z\,;\, X \circ F(\zeta), Y \circ F(\zeta))| \,<\, 1 \quad \text{for each } \zeta \in \mathbb{D}. \tag{4.9}$$

It is obvious that the functions

$$\zeta \longmapsto P_n(z\,;\, X \circ F(\zeta)), \qquad \zeta \longmapsto Q_n(z\,;\, Y \circ F(\zeta)), \quad \zeta \in \mathbb{D},$$

are holomorphic functions. Thus, it follows from the bound (4.9) that f^z defined by

$$f^z(\zeta) \,:=\, \Psi_n(z\,; X \circ F(\zeta), Y \circ F(\zeta)), \quad \zeta \in \mathbb{D},$$

is a holomorphic $\mathbb{D}$-valued function. This function, for each fixed $z \in \overline{\mathbb{D}}$, satisfies

$$f^z(\zeta_j) \,=\, \Psi_n(z\,;\, X(W_j), Y(W_j)), \quad j = 1, \ldots, M.$$

It follows from the classical result by Pick that the $M \times M$ matrix M_z is positive semi-definite. □

We end this article with an observation:

Remark 4.2. Two major effects of the idea introduced by Agler–Young in [2] — of which this work is an extension — are the reduction in the dimensional complexity of the problem $(*)$, and the ability to deduce necessary conditions for Nevanlinna–Pick interpolation such as Theorem 1.3. However, the discussion in Remark 4.1 suggests that the advantage to be gained from the first of those two features has certain limits. As the number of the disparate diagonal blocks determining E increases, the number of generators of the ring of G-invariant functions on Ω_E (for an appropriate reductive group G naturally associated with Ω_E and acting on it by conjugation) would tend to grow; see Lemma 2.1 above. This implies that there would be diminishing advantage, in terms of reduction in dimensional complexity of the problem $(*)$, in working with analogues of $\mathbb{G}_n$ or $\mathbb{E}_n$.

References

[1] A.A. Abouhajar, M.C. White, and N.J. Young, *A Schwarz lemma for a domain related to μ-synthesis*, J. Geom. Anal. **17** (2007), no. 4, 717-750.

[2] J. Agler and N.J. Young, *A commutant lifting theorem for a domain in $\mathbb{C}^2$ and spectral interpolation*, J. Funct. Anal. **161** (1999), 452-477.

[3] J. Agler and N.J. Young, *The two-point spectral Nevanlinna–Pick problem*, Integral Equations Operator Theory **37** (2000), 375-385.
[4] J. Agler and N.J. Young, *The two-by-two spectral Nevanlinna–Pick problem*, Trans. Amer. Math. Soc. **356** (2004), 573-585.
[5] H. Bercovici, *Spectral versus classical Nevanlinna–Pick interpolation in dimension two*, Electron. J. Linear Algebra **10** (2003), 60-64.
[6] H. Bercovici, C. Foias and A. Tannenbaum, *A spectral commutant lifting theorem*, Trans. Amer. Math. Soc. **325** (1991), 741-763.
[7] E.M. Chirka, *Complex Analytic Sets*, Mathematics and its Applications (Soviet Series), volume 46, Kluwer Academic Publishers, Dordrecht, 1989.
[8] C. Costara, *On the spectral Nevanlinna–Pick problem*, Studia Math. **170** (2005), 23-55.
[9] J.C. Doyle, *Analysis of feedback systems with structured uncertainties*, IEE Proc. Control Theory Appl. **129** (1982), no. 6, 242-250.
[10] B.A. Francis, *A Course in H^∞ Control Theory*, Lecture Notes in Control and Information Sciences **88**, Springer-Verlag, Berlin, 1987.
[11] J.W. Helton, *Non-Euclidean functional analysis and electronics*, Bull. Amer. Math. Soc. **7** (1982), no. 1, 1-64.
[12] N. Nikolov, P. Pflug and P.J. Thomas, *Spectral Nevanlinna–Pick and Carathéodory–Fejér problems for* $n \leq 3$, Indiana Univ. Math. J. **60** (2011), no. 3, 883-893.
[13] W. Rudin, *Function Theory in the Unit Ball of* $\mathbb{C}^n$, Springer-Verlag, New York, 1980.
[14] D.M. Snow, *Reductive group actions on Stein spaces*, Math. Ann. **259** (1982), no. 1, 79-97.
[15] B.L. van der Waerden, *Algebra Vol. 1 (translated from the German by F. Blum & J.R. Schulenberger)*, Frederick Ungar Publishing Co., 1970.

Department of Mathematics, Indian Institute of Science, Bangalore – 560 012
E-mail address: `bharali@math.iisc.ernet.in`