J. OPERATOR THEORY
70:2(2013), 531–571
doi: 10.7900/jot.2011sep21.1940

CONVEXITY ANALYSIS AND THE MATRIX-VALUED SCHUR CLASS OVER FINITELY CONNECTED PLANAR DOMAINS

JOSEPH A. BALL and MOISÉS D. GUERRA HUAMÁN

Dedicated to the memory of William B. Arveson, a true friend and inspiration for the field

Communicated by Kenneth B. Davidson

ABSTRACT. We identify the set of extreme points and apply Choquet theory to a normalized matrix-measure ball subject to finitely many linear side constraints. As an application we obtain integral representation formulas for the Herglotz class of matrix-valued functions on a finitely-connected planar domain and associated continuous Agler decompositions for the matrix-valued Schur class over the domain. The results give some additional insight into the negative answer to the spectral set problem over such domains recently obtained by Agler–Harland–Raphael and Dritschel–McCullough.

KEYWORDS: *Choquet theory, positive operator measures, Schur class, finitely connected planar domain, C^*-convex combination, interior point of the C^*-convex hull.*

MSC (2010): Primary 46A55; Secondary 47A48, 47A56, 47A20.

INTRODUCTION

We define the classical Schur class (operator-valued version) $\mathcal{S}(\mathcal{U},\mathcal{Y})$ to be the class of holomorphic functions $z \mapsto S(z)$ from the unit disk $\mathbb{D}$ into contraction operators between two coefficient Hilbert spaces $\mathcal{U}$, $\mathcal{Y}$. This class has been an object of much study and a source of much inspiration over the last several decades due to its central role in a number of applications but also due to the rich commingling of function theory, operator theory and engineering system theory ideas in the description of its structure. Let us mention several equivalent characterizations/points of view toward the Schur class: (1) the operator M_S of multiplication by S defines a contraction operator on the Hardy space over $\mathbb{D}$, (2) the de Branges–Rovnyak kernel

$$K_S(z,w) = \frac{I - S(z)S(w)^*}{1 - z\overline{w}} \tag{0.1}$$

is a positive kernel over $\mathbb{D}$, (3) S can be realized as the transfer function of a conservative discrete-time input/state/output linear system:

$$S(z) = D + zC(I - zA)^{-1}B \quad \text{with } \mathbf{U} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \to \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix} \text{ unitary.} \tag{0.2}$$

A major step forward in developing an analogous theory in several-variable settings was made by Agler [3] where what we now call the *Schur–Agler class* over the polydisk was introduced. This class is defined as the class of operator-valued functions on the polydisk $\mathbb{D}^d = \{z = (z_1, \ldots, z_d) : |z_k| < 1 \text{ for } k = 1, \ldots, d\}$ such that not only $\|S(z)\| \leqslant 1$ for each $z \in \mathbb{D}^d$ but also $\|S(T)\| \leqslant 1$ for all commutative d-tuples $T = (T_1, \ldots, T_d)$ of operators on some Hilbert space $\mathcal{K}$, where e.g. $S(T)$ can be defined as

$$S(T) = \sum_{n \in \mathbb{Z}_+^d} S_n \otimes T^n \quad \text{if } S(z) = \sum_{n \in \mathbb{Z}_+^d} S_n z^n.$$

A new feature for this class is the analogue of positivity of the kernel (0.1): rather than a characterization in terms of the positivity of a single kernel, the characterization is in terms of being able to solve for d positive kernels $K_1, \ldots, K_d$ on the polydisk so that the so-called *Agler decomposition* holds:

$$I - S(z)S(w)^* = \sum_{k=1}^{d} (1 - z_k \overline{w}_k) K_k(z, w). \tag{0.3}$$

Also the realization formula (0.2) for the multivariable Schur–Agler class takes the form

$$S(z) = D + C(I - Z(z)A)^{-1} Z(z) B \quad \text{with } \mathbf{U} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \to \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \text{ unitary} \tag{0.4}$$

with $Z(z) = \sum_{k=1}^{d} z_k P_k$ where $P_1, \ldots, P_d$ form a spectral family of projection operators ($P_i = P_i^*$, $P_i P_j = \delta_{i,j} I_{\mathcal{X}}$, $\sum_{k=1}^{d} P_k = I_{\mathcal{X}}$) on the state space $\mathcal{X}$.

One of our main motivations for the present paper was to further develop the understanding of the Schur class $\mathcal{S}(\mathcal{R})$ over a bounded finitely-connected planar domain $\mathcal{R}$. Here $\mathcal{S}(\mathcal{R})$ denotes the class of holomorphic functions mapping the planar domain $\mathcal{R}$ into the unit disk. We shall use the notation $\mathcal{S}^N(\mathcal{R})$ for the class of holomorphic functions mapping $\mathcal{R}$ into contractive $N \times N$ matrices. In the course of constructing a counterexample to the spectral set question over $\mathcal{R}$, Dritschel and McCullough [13] obtained a continuous analogue of the Agler decomposition (0.3) for the scalar-valued Schur class $\mathcal{S}(\mathcal{R})$ over $\mathcal{R}$. Specifically, let $\partial_0, \ldots, \partial_m$ denote the $m+1$ connected components of the boundary $\partial\mathcal{R}$ of $\mathcal{R}$ (with ∂_0 equal to the boundary of the unbounded component of the complement of $\mathcal{R}$ in the complex plane) and let $\mathbb{T}_{\mathcal{R}}$ denote the Cartesian product $\mathbb{T}_{\mathcal{R}} = \partial_0 \times \cdots \times \partial_m$. The coordinate functions $z_1, \ldots, z_d$ appearing in the Agler decomposition (0.3) must be replaced by a continuum $\{s_{\mathbf{x}}(z) : \mathbf{x} \in \mathbb{T}_{\mathcal{R}}\}$ of single-valued inner functions on $\mathcal{R}$ (i.e., holomorphic in $\mathcal{R}$ with modulus-1 values on $\partial\mathcal{R}$), each with m zeros in $\mathcal{R}$ (the minimal number possible for a nonconstant single-valued inner function), indexed by the so-called $\mathcal{R}$-torus $\mathbb{T}_{\mathcal{R}}$. Then the result of Dritschel–McCullough can be formulated as follows:

THEOREM 0.1. *Given any $s \in \mathcal{S}(\mathcal{R})$, then there is a family $k_{\mathbf{x}}(z,w)$ of positive kernels on $\mathcal{R}$, indexed by $\mathbb{T}_{\mathcal{R}}$ and measurable on $\mathbb{T}_{\mathcal{R}}$ for each $(z,w) \in \mathcal{R} \times \mathcal{R}$, so that*

$$1 - s(z)\overline{s(w)} = \int_{\mathbb{T}_{\mathcal{R}}} (1 - s_{\mathbf{x}}(z)\overline{s_{\mathbf{x}}(w)}) k_{\mathbf{x}}(z,w)\, \mathrm{d}\nu(\mathbf{x}). \tag{0.5}$$

There is also obtained in [13] a more elaborate version of the realization formula (0.2) or (0.4) for the Schur class $\mathcal{S}(\mathcal{R})$ which we do not go into here. We also mention that these techniques actually lead to interpolation theorems for the various Schur classes: if the function S is initially given only on some (possibly finite) subset of its domain ($\mathbb{D}$, $\mathbb{D}^d$, or $\mathcal{R}$), then a necessary and sufficient condition for there to be an extension to the whole domain which is in the appropriate Schur class is that the decomposition (0.1), (0.3), (0.5) hold for z, w in the subset. A dual version of the interpolation result for the class $\mathcal{S}(\mathcal{R})$, whereby one tests the positivity of each kernel from a collection of kernels $\{(1 - s(z)\overline{s(w)})k^{(\alpha)}(z,w)\}$ (where $k^{(\alpha)}(z,w)$ is a collection of Szegő-type kernels indexed by α from the m-torus $\mathbb{T}^m$), was obtained earlier by Abrahamse [1].

While Dritschel–McCullough indicated some results for the matrix-valued Schur class over $\mathcal{R}$ on their way to constructing a counterexample to the spectral set question over $\mathcal{R}$, the analogue of (0.5) for the matrix-valued case was left rather mysterious. In general, extensions of scalar-valued results to the matrix-valued case for the Schur class over a planar domain $\mathcal{R}$ have led to surprises: it is known for example that the Abrahamse interpolation result does not extend to the matrix-valued case without the addition of additional matrix-valued kernels $k^{(\boldsymbol{\alpha})}$ (see [7], [11], [25], [26]).

One of the main motivations of the present paper was to find an appropriate analogue of the Dritschel–McCullough decomposition (0.5) for the matrix-valued setting; such an analogue appears as Theorem 4.4 below. The basic idea in [13] for getting the decomposition (0.5) is to apply a linear-fractional change of variable on the range of the function to convert the problem to a problem concerning the Herglotz class over $\mathcal{R}$ (holomorphic functions on $\mathcal{R}$ with positive real part). When this class is normalized by the condition that all such functions f have the value 1 at some fixed point $t_0 \in \mathcal{R}$, it becomes a compact convex set. Once one identifies the extreme points for this class, Choquet theory (see e.g. [27] for a thorough account) can be applied to obtain an integral representation for a given Herglotz-class function f in terms of the extreme points $f_{\mathbf{x}}$. The Cayley transforms of these extreme points for the Herglotz class turn out to be unimodular scalar multiples of the inner functions with exactly m zeros appearing in the decomposition (0.5): $s_{\mathbf{x}}(z) = \frac{f_{\mathbf{x}}(z)-1}{f_{\mathbf{x}}(z)+1}$. Explicit identification of the extreme points $f_{\mathbf{x}}$ involves some clever function theory (see [4], [13], [20], [29]). The starting point is the Poisson-kernel representation for positive harmonic functions. This leads to a one-to-one correspondence between normalized Herglotz functions on $\mathcal{R}$ and probability measures on $\partial\mathcal{R}$ which satisfy m additional linear constraints (m equal to the number of holes in $\mathcal{R}$). In this way extremal normalized Herglotz functions correspond to probability measures which are extremal in this set of linearly-constrained probability measures. The problem of characterizing the extreme points of such a set of probability measures can be formulated in the setting of an abstract Borel measure space X (in place of $\partial\mathcal{R}$). We study this general problem and give a geometric characterization of the extreme points in terms of 0 being in the interior of the convex hull of a given collection of vectors in $\mathbb{R}^m$, putting the results of Dritschel–Pickering in [15] into a broader context.

The extension of these ideas to the matrix-valued setting leads to new issues to be understood. Each $N \times N$-matrix valued Herglotz function normalized to be the identity I_N at the fixed point $t_0 \in \mathcal{R}$ corresponds to a quantum probability measure, i.e., a positive matrix-valued measure μ on $\partial\mathcal{R}$ with total mass $\mu(\partial\mathcal{R})$ equal to the identity matrix I_N, subject to m linear side constraints (given by integration against m continuous real-valued functions on $\partial\mathcal{R}$). This problem in turn can be considered more generally, where $\partial\mathcal{R}$ is replaced by a general Borel space X. The problem then is to characterize the set of extreme points of the compact convex set of quantum probability measures subject to m linear side constraints. It turns out that the special case of this problem where there are no side constraints has been analyzed and solved by Arveson [5]: extremal measures μ are characterized by the condition that $\mu = \sum_{k=1}^{n} W_k \delta_{x_k}$ (where δ_{x_k} is the scalar unit point-mass measure at the point x_k and $W_k \geqslant 0$ is a matrix weight) where the family of subspaces $\{\operatorname{Ran} W_k : 1 \leqslant k \leqslant n\}$ should satisfy a condition called *weak independence* which, as suggested by the terminology, is somewhat weaker than the standard linear algebra notion of linear independence of subspaces (i.e., any collection of nonzero vectors $x_1, \dots, x_d$ with $x_k \in W_k$ should be a linear independent set of vectors in the standard sense). We obtain an extension of Arveson's result to the constrained case which has a geometric interpretation analogous to that in [15] for the scalar-case, namely: the 0 vector must be in the interior of the C^*-convex hull of a given set of matrix-tuples (see Remark 2.11 below), thereby providing links with the general area of noncommutative convexity as in [16], [17], [19]. Finally we apply this general result on extreme points to obtain a characterization (although not quite as explicit as in the scalar-valued case) of the extreme points of the normalized matrix-valued Herglotz class over a planar domain $\mathcal{R}$.

We do not treat here the transfer-function realization and interpolation theory for the matrix-valued Schur class $\mathcal{S}^N(\mathcal{R})$. Such results can be obtained as part of a general theory of the matrix-valued Schur class associated with a collection of matrix-valued test functions. We address this topic beyond what already appears in [21] in a separate report [8].

The paper is organized as follows. Following this Introduction, Section 1 sets notation and reviews results from convexity theory (in particular, the Choquet–Bishop–de Leeuw theory on integral representations for points of a compact, convex set) which will be needed in the sequel. Section 2 considers the extreme-point problem for a linearly-constrained normalized set of positive matrix measures in the general measure-theory framework. Section 3 introduces the function-theory setting and applies the theory of Section 2 to obtain characterizations of extreme points and integral representations for normalized matrix-valued Herglotz-class functions over a finitely-connected planar domain $\mathcal{R}$. Section 4 applies the linear-fractional change of variable to convert the results concerning Herglotz-class functions to results concerning Schur-class functions over $\mathcal{R}$. The final Section 5 presents connections with the spectral set question over a region $\mathcal{R}$: it turns out that the recent negative solution of the spectral set question can be partially explained by the lack of a simple transition formula from the extreme points for the scalar-valued normalized Herglotz class to the extreme points for the matrix-valued normalized Herglotz class over $\mathcal{R}$ (see Corollary 5.1 below).

Preliminary versions of many of the results described appear already in the Virginia Tech dissertation of the second author [21].

1. GENERAL CONVEXITY THEORY

A subset $\mathcal{C}$ of a real linear space C is said to be *convex* if, given any collection of vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ in $\mathcal{C}$ and a collection of nonnegative real numbers $\lambda_1, \dots, \lambda_n$ with $\lambda_1 + \cdots + \lambda_n = 1$, it happens that the convex linear combination $\sum_{i=1}^{n} \lambda_i \mathbf{u}_i$ is again in $\mathcal{C}$. Given any subset $\mathcal{S}$ of the linear space E, there is always a smallest subset of E containing $\mathcal{S}$, denoted as $\operatorname{conv} \mathcal{S}$ (the *convex hull* of $\mathcal{S}$).

A vector $\mathbf{v}$ in the convex set $\mathcal{C}$ is said to be an *extreme point* of $\mathcal{C}$ if, whenever it is the case that $\mathbf{v} = \lambda \mathbf{u}_1 + (1-\lambda)\mathbf{u}_2$ for a real λ with $0 < \lambda < 1$ and $\mathbf{u}_1$ and $\mathbf{u}_2$ in $\mathcal{C}$, it follows that $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{v}$. The following characterization of extreme point is often easier to apply than the definition.

LEMMA 1.1. *The point $\mathbf{v} \in \mathcal{C}$ is an extreme point of the convex set $\mathcal{C}$ ($\mathbf{v} \in \partial_{\mathrm{e}}\mathcal{C}$) if and only if the following condition holds: whenever $\mathbf{u} \in E$ is such that $\mathbf{v} \pm \mathbf{u} \in \mathcal{C}$, then $\mathbf{u} = 0$.*

Proof. Suppose $\mathbf{v}$ is extreme and $\mathbf{v} \pm \mathbf{u} \in \mathcal{C}$ for some $\mathbf{u} \in E$. Since $\mathbf{v}$ is extreme, from the identity

$$\mathbf{v} = \frac{1}{2}(\mathbf{v} + \mathbf{u}) + \frac{1}{2}(\mathbf{v} - \mathbf{u})$$

we see immediately that $\mathbf{u} = 0$.

For the converse it suffices to show the contrapositive: $\mathbf{v}$ *not extreme* $\Rightarrow$ *there is a* $\mathbf{u} \neq 0$ *in* E *with* $\mathbf{v} \pm \mathbf{u} \in \mathcal{C}$. If $\mathbf{v}$ is not extreme, then we can find $\mathbf{v}_1, \mathbf{v}_2$ in $\mathcal{C}$ distinct from $\mathbf{v}$ so that $\mathbf{v} = \lambda \mathbf{v}_1 + (1-\lambda)\mathbf{v}_2$. We rearrange this as

$$\lambda(\mathbf{v} - \mathbf{v}_1) = (1-\lambda)(\mathbf{v}_2 - \mathbf{v}) =: \mathbf{u}.$$

Then

$$\begin{aligned} \mathbf{v} + \mathbf{u} &= \mathbf{v} + (1-\lambda)(\mathbf{v}_2 - \mathbf{v}) = \lambda \mathbf{v} + (1-\lambda)\mathbf{v}_2 \in \mathcal{C}, \\ \mathbf{v} - \mathbf{u} &= \mathbf{v} - \lambda(\mathbf{v} - \mathbf{v}_1) = (1-\lambda)\mathbf{v} + \lambda \mathbf{v}_1 \in \mathcal{C} \end{aligned}$$

from which we see that the vector $\mathbf{u}$ has the needed property. ∎

Given a collection of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_\kappa\}$ in a linear space E and given another vector $\mathbf{v}$ in E, we say that $\mathbf{v}$ *is in the convex hull of the set of vectors* $\mathcal{S} = \{\mathbf{v}_1, \dots, \mathbf{v}_\kappa\}$, written as

$$\mathbf{v} \in \operatorname{conv} \mathcal{S},$$

if $\mathbf{v}$ can be written as a convex linear combination $\mathbf{v} = \sum_{i=1}^{\kappa} \lambda_i \mathbf{u}_i$ of the $\mathbf{u}_i$s (so $\lambda_i \geqslant 0$ and $\sum_{i=1}^{\kappa} \lambda_i = 1$). In case the linear space E carries a locally convex topology and the convex subset $\mathcal{C}$ is compact in this topology, the well known theorem of Kreĭn–Milman (see e.g. page 75 of [31]) asserts that $\mathcal{C}$ is the closed convex hull of the set of its extreme points $\partial_e \mathcal{C}$. There is a refinement of the Kreĭn–Milman theorem known generically as *Choquet theory*. In general let us say that a vector $\mathbf{v}$ in the nonempty compact subset X of the linear topological vector space E is *represented* by the probability Borel measure ν on X if it is the case that

$$\ell(\mathbf{v}) = \int_X \ell(\mathbf{u})\, \mathrm{d}\nu(\mathbf{u})$$

for all continuous linear functionals $\ell \in E^*$. A consequence of the Hahn–Banach theorem then is that ν uniquely determines the element $\mathbf{v} \in E$. The following theorem summarizes what we need from Choquet theory and is due mainly to Choquet [10] and Bishop–de Leeuw [9].

THEOREM 1.2 (See [27] and Section IV.6 of [32]). *Suppose that* $\mathcal{C}$ *is a compact convex subset of the linear topological vector space* E *and* $\mathbf{v} \in \mathcal{C}$. *Then there is a probability measure* ν *supported on the closure of the set of extreme points* $(\partial_e \mathcal{C})^-$ *which represents* $\mathbf{v}$. *In case* $\mathcal{C}$ *is metrizable, then* $\partial_e \mathcal{C}$ *is a Borel set and one can arrange that* ν *is supported exactly on the set of extreme points* $\partial_e \mathcal{C}$.

We shall have need of a refinement of the notion of convex hull, namely *interior convex hull* defined as follows. Given a collection of vectors $\mathcal{S} = \{\mathbf{u}_1, \dots, \mathbf{u}_\kappa\}$ in a linear space E and another vector $\mathbf{v}$ in E as above, we say that $\mathbf{v}$ *is in the interior of the convex hull of the set of vectors* $\mathcal{S} = \{\mathbf{u}_1, \dots, \mathbf{u}_\kappa\}$, written as

$$\mathbf{v} \in \operatorname{conv}^0 \mathcal{S},$$

if $\mathbf{v}$ can be written as a convex combination $\mathbf{v} = \sum_{i=1}^{\kappa} \lambda_i \mathbf{u}_i$ of the elements of $\mathcal{S}$ with the coefficients λ_i satisfying the strict inequalities $\lambda_i > 0$ together with $\sum_{i=1}^{\kappa} \lambda_i = 1$ uniquely determined. We will be particularly interested in the case when the zero vector $\mathbf{0}$ in E is in the interior convex hull of $\mathcal{S}$. In general there are several equivalent formulations of the condition that $\mathbf{0} \in \operatorname{conv}^0 \mathcal{S}$.

PROPOSITION 1.3. *Given a finite subset $\mathcal{S} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ of a linear space E, suppose that the zero vector $\mathbf{0}$ is a proper convex combination of $\mathbf{u}_1, \dots, \mathbf{u}_n$ in the linear space E:*

$$\mathbf{0} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \quad \textit{with } \lambda_i > 0 \textit{ for all } i \textit{ and } \sum_{i=1}^{n} \lambda_i = 1.$$

Then the following conditions are equivalent:

(i) $\mathbf{0} \in \operatorname{conv}^0 \mathcal{S}$, *i.e., the real numbers $\lambda_1, \dots, \lambda_n$ are uniquely determined by the conditions:*

$$\lambda_i > 0 \quad \textit{for all } i; \quad \sum_{i=1}^{n} \lambda_i = 1; \quad \sum_{i=1}^{n} \lambda_i \mathbf{u}_i = \mathbf{0}. \tag{1.1}$$

(ii) *The linear subspace of $\mathbb{R}^n$ consisting of vectors $\mathbf{c} = (c_1, \dots, c_n)$ such that $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = 0 \in E$ is one-dimensional (and hence is spanned by $\lambda = (\lambda_1, \dots, \lambda_n)$).*

(iii) *The only solution $\mathbf{c} = (c_1, \dots, c_n)$ of the system of equations*

$$\sum_{i=1}^{n} c_n = 0, \quad \sum_{i=1}^{n} c_i \mathbf{u}_i = \mathbf{0}, \tag{1.2}$$

is $\mathbf{c} = (0, \dots, 0)$.

Proof. We show (not (i)) $\Rightarrow$ (not (ii)) $\Rightarrow$ (not (iii)) $\Rightarrow$ (not (i)).

Step (not (i)) $\Rightarrow$ (not (ii)). Suppose that $\lambda = (\lambda_1, \dots, \lambda_n)$ and $\lambda' = (\lambda'_1, \dots, \lambda'_n)$ are two distinct elements of $\mathbb{R}^n$ satisfying the conditions in (i). Set $c_i = \lambda_i - \lambda'_i$ and $\mathbf{c} = (c_1, \dots, c_n)$. Then $\mathbf{c}$ is a second nonzero solution of $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$. Since $\sum_{i=1}^{n} c_i = \sum_{i-1} (\lambda_i - \lambda'_i) = 0$ while $\lambda_i > 0$ for all i, we see that $\mathbf{c}$ is linearly independent of λ. Hence the set of solutions $\mathbf{c} = (c_1, \dots, c_n)$ of $\sum_{i=1}^{n} c_i \mathbf{u}_i = \mathbf{0}$ has dimension at least 2, in contradiction to (ii).

Step (not (ii)) $\Rightarrow$ (not (iii)). If the space of solutions $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{R}^n$ of the linear system $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$ has dimension at least 2, then by the null-kernel theorem from linear algebra we can find a nonzero such vector which satisfies the single additional linear constraint $c_1 + \cdots + c_n = 0$, in contradiction with (iii).

Step (not (iii)) $\Rightarrow$ (not (i)). Assume that $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{R}^n$ is a nonzero solution of (1.2). Set $\lambda' = (\lambda_1 + \varepsilon c_1, \dots, \lambda_n + \varepsilon c_n)$ for some $\varepsilon > 0$. Then as long as ε is chosen sufficiently small, λ' is a second solution of (1.1), in contradiction with (i). ∎

2. AN EXTREME-POINT PROBLEM FOR A CONSTRAINED NORMALIZED BALL OF MATRIX MEASURES

In this section we consider the following general extreme-point problem which is central for our analysis of the matrix-valued Herglotz and Schur classes of holomorphic functions over a finitely connected planar domain discussed in the next section. We suppose that we are given a compact Hausdorff space X. We let $M(X)$ denote the space of complex Borel measures on X and $C_{\mathbb{R}}(X)$ denote the space of real-valued continuous functions on X. For N a positive integer, $M(X)^{N\times N}$ then denotes the space of complex $N \times N$ matrix-valued Borel measures on X. We will also have occasion to use $[M(X)^{N\times N}]_{\mathrm{h}}$ to denote complex Hermitian $N \times N$ matrix-valued measures and $[M(X)^{N\times N}]_{+}$ the subset of $[M(X)^{N\times N}]_{\mathrm{h}}$ consisting of positive matrix measures. We suppose that we are also given a collection $\boldsymbol{\phi} = \{\phi_1, \dots, \phi_m\}$ of m real-valued continuous functions on X (i.e., $\phi_1, \dots, \phi_m \in C_{\mathbb{R}}(X)$). We then let $\mathcal{C}(X, N, \boldsymbol{\phi})$ be the subset of $[M(X)^{N\times N}]_{\mathrm{h}}$ given by

$$\begin{aligned} \mathcal{C}(X,N,\boldsymbol{\phi}) = \Big\{ \mu \in [M(X)^{N\times N}]_{+} : \mu(X) = I, \text{ and} \\ \mu(\phi_r) := \int_X \phi_r(x)\,\mathrm{d}\mu(x) = 0 \text{ for } r = 1, \dots, m \Big\}. \end{aligned} \tag{2.1}$$

Note that $\mathcal{C}(X, N, \boldsymbol{\phi})$ is a convex subset of the real Banach space of complex Hermitian matrices $[M(X)^{N\times N}]_{\mathrm{h}}$ which is compact in the weak* topology on $[M(X)^{N\times N}]_{\mathrm{h}}$ induced by its duality with respect to the real Banach space of complex-Hermitian matrix-valued continuous functions $[C(X)^{N\times N}]_{\mathrm{h}}$. In view of the Kreĭn–Milman theorem and the results discussed in Section 1, it is then natural to pose the problem:

PROBLEM 2.1. Given a data set $(X, N, \boldsymbol{\phi})$ as above, characterize the set of extreme points of the associated compact, convex set $\mathcal{C}(X, N, \boldsymbol{\phi})$ given by (2.1).

Simple examples show that it is possible that $\mathcal{C}(X, N, \phi)$ is empty: for example, take X equal to the unit interval $[0, 1]$, $N = 1$, $m = 1$ with $\phi_1(x) = 1$. Then the condition that $1 = \mu(X) = \int_X \mathrm{d}\mu(x)$ and that $0 = \int_X \phi_1(x)\mathrm{d}\mu(x) = \int_X \mathrm{d}\mu(x)$ are contradictory. In the discussion to follow we will implicitly assume that

$\mathcal{C}(X, N, \phi) \neq \emptyset$; in all examples arising from some natural context, it is the case that $\mathcal{C}(X, N, \phi) \neq \emptyset$.

The following result is a first step toward obtaining more definitive solutions for various special cases of interest.

THEOREM 2.2. *Suppose that $\mu \in [M(X)^{N\times N}]_+$ is an extreme point of the set $\mathcal{C}(X, N, \boldsymbol{\phi})$ (2.1). Then there is a natural number n with $1 \leqslant n \leqslant (m+1)N^2$, n distinct points $x_1, \dots, x_n$ in X, and n positive semidefinite $N \times N$ matrices $W_1, \dots, W_n$ subject to the system of linear equations*

$$\sum_{r=1}^{n} W_i = I, \quad \sum_{r=1}^{n} \phi_i(x_r) W_r = 0 \quad \textit{for } i = 1, \dots, m \tag{2.2}$$

so that μ has the form

$$\mu = \sum_{j=1}^{n} W_j \delta_{x_j} \tag{2.3}$$

where δ_{x_j} is the scalar-valued measure equal to the unit point-mass at the point x_j.

Proof. It suffices to show that any measure $\mu = [\mu_{ij}]_{i,j=1,\dots,N}$ which is an extreme point has the form (2.3); conditions (2.2) then follow just by the condition that μ is an element of $\mathcal{C}(X, N, \boldsymbol{\phi})$. By way of contradiction, suppose that $\mu \in [M(X)^{N\times N}]_+$ is a positive matrix measure which is not of the form (2.3). We then must show that μ is not extreme.

If μ is not of the form (2.3) with $1 \leqslant n \leqslant (m+1)N^2$, then there are κ (with $\kappa > (m+1)N^2$) disjoint Borel sets $\Delta_1, \dots, \Delta_\kappa$ with $\mu(\Delta_j) \neq 0$. Define new measures $\mu_1, \dots, \mu_\kappa$ by

$$\mu_j(\Delta) = \mu(\Delta \cap \Delta_j) \quad \text{for } j = 1, \dots, \kappa.$$

Then the collection $\{\mu_1, \dots, \mu_\kappa\}$ is linearly independent in the real vector space $[M(X)^{N\times N}]_{\mathrm{h}}$ of complex-Hermitian matrix-valued Borel measures on X. Now define real linear functionals on $[M(X)^{N\times N}]_{\mathrm{h}}$ by

$$\begin{aligned}
&L_i : \mu \mapsto \mu_{ii}(X), \quad 1 \leqslant i \leqslant N,\\
&L_{\mathrm{Re},ij} : \mu \mapsto \ \mathrm{Re}\, \mu_{ij}(X), \quad 1 \leqslant i < j \leqslant N,\\
&L_{\mathrm{Im},ij} : \mu \mapsto \ \mathrm{Im}\, \mu_{ij}(X), \quad 1 \leqslant i < j \leqslant N,\\
&L_{i,r} : \mu \mapsto \mu_{ii}(\phi_r), \quad 1 \leqslant i \leqslant N, 1 \leqslant r \leqslant m,\\
&L_{\mathrm{Re},ij,r} : \mu \mapsto \ \mathrm{Re}\, \mu_{ij}(\phi_r), \quad 1 \leqslant i < j \leqslant N, 1 \leqslant r \leqslant m,\\
&L_{\mathrm{Im},ij,r} : \mu \mapsto \ \mathrm{Im}\, \mu_{ij}(\phi_r), \quad 1 \leqslant i < j \leqslant N, 1 \leqslant r \leqslant m.
\end{aligned}$$

Note that in total there are at most (exactly in case the linear functionals on the above list are all distinct)

$$N + \frac{N(N-1)}{2} + \frac{N(N-1)}{2} + Nm + \frac{N(N-1)}{2}m + \frac{N(N-1)}{2}m = N^2(m+1)$$

such real linear functionals. Note that for $1 \leqslant i \leqslant j \leqslant N$ and $1 \leqslant r \leqslant m$ we have

$$\mu_{ji}(X) = \mu_{ij}(X)^* \quad \text{and} \quad \mu_{ji}(\phi_r) = \mu_{ij}(\phi_r)^*. \tag{2.4}$$

We now define a real linear map L from $[M(X)^{N\times N}]_{\mathrm{h}}$ to $\mathbb{R}^{(m+1)N^2}$ by

$$L(\mu) = \begin{bmatrix} \mathrm{col}_j\{L_j(\mu) : 1 \leqslant j \leqslant N\} \\ \mathrm{col}_{i,j}\{L_{\mathrm{Re},ij}(\mu) : 1 \leqslant i < j \leqslant N\} \\ \mathrm{col}_{i,j}\{L_{\mathrm{Im},ij}(\mu) : 1 \leqslant i < j \leqslant N\} \\ \mathrm{col}_{i,r}\{L_{i,r}(\mu) : 1 \leqslant i \leqslant N,\, 1 \leqslant r \leqslant m\} \\ \mathrm{col}_{i,j,r}\{L_{\mathrm{Re},ij,r}(\mu) : 1 \leqslant i < j \leqslant N,\, 1 \leqslant r \leqslant m\} \\ \mathrm{col}_{i,j,r}\{L_{\mathrm{Im},ij,r}(\mu) : 1 \leqslant i < j \leqslant N,\, 1 \leqslant r \leqslant m\} \end{bmatrix}$$

where we use the notation $\mathrm{col}\{X_j : 1 \leqslant j \leqslant N\}$ to denote the column matrix $\mathrm{col}\{X_j : 1 \leqslant j \leqslant N\} = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}$. Consider the restriction of L to the κ-dimensional subspace $\mathcal{M} := \mathrm{span}\{\mu_1, \dots, \mu_\kappa\}$. Since $\kappa > (m+1)N^2$, as a consequence of the null-kernel theorem from linear algebra we see that there exists a nonzero measure ν of the form $\nu = \sum_{\ell=1}^{\kappa} c_\ell \mu_\ell \in \mathcal{M}$ with $c_\ell \in \mathbb{R}$ for all ℓ so that $L(\nu) = 0$. Consequently the matrix measure $\nu = [\nu_{ij}]_{i,j=1,\dots,N}$ satisfies $\nu_{ii}(X) = 0$ for all $i = 1, \dots, N$ and $\nu_{ij}(X) = 0$ for $1 \leqslant i < j \leqslant N$. From (2.4) we see that $\nu_{ij}(X) = 0$ for all $1 \leqslant i \leqslant j \leqslant N$ as well and we conclude that

$$\nu(X) = 0. \tag{2.5}$$

In a similar way we see in addition that

$$\nu(\phi_r) = 0 \quad \text{for } r = 1, \dots, m. \tag{2.6}$$

We next choose $\varepsilon > 0$ so that $\varepsilon < \min\{\frac{1}{c_j} : j \text{ with } c_j \neq 0\}$ where $c_1, \dots, c_\kappa$ are the coefficients in the representation $\nu = c_1\mu_1 + \cdots + c_\kappa\mu_\kappa$ for ν as an element of the space $\mathcal{M} = \mathrm{span}\{\mu_j : j = 1, \dots, \kappa\}$. Then by construction

$$1 \pm \varepsilon c_j \geqslant 0 \quad \text{for } j = 1, \dots, \kappa.$$

It follows that $(\mu \pm \varepsilon\nu)(X) = I$, $(\mu \pm \varepsilon\nu)(\Delta) \geqslant 0$ for all Borel Δ and $(\mu \pm \varepsilon\nu)(\phi_r) = \mu(\phi_r) = 0$ for $1 \leqslant r \leqslant m$, i.e., $\mu \pm \varepsilon\nu \in \mathcal{C}(X, N, \boldsymbol{\phi})$. Since it is also the case that $\varepsilon\nu$ is not the zero element of $[M(X)^{N\times N}]_{\mathrm{h}}$, it follows as a consequence of Lemma 1.1 that $\mu \notin \partial_{\mathrm{e}}\mathcal{C}(X, N, \boldsymbol{\phi})$, as needed to be shown. ∎

2.1. THE SCALAR-VALUED CASE: $N = 1$. We now analyze Problem 2.1 for the scalar-valued case ($N = 1$). The following result gives a complete characterization of $\partial_{\mathrm{e}}\mathcal{C}(X, N, \phi)$ ($\mathcal{C}(X, N, \phi)$ as in (2.1)) for the scalar case ($N = 1$).

THEOREM 2.3. *Suppose that we are given a compact Hausdorff space X along with m real-valued continuous functions $\boldsymbol{\phi} = \{\phi_1, \dots, \phi_m\}$ and we let $\mathcal{C}(X, 1, \boldsymbol{\phi})$ be the* *associated compact convex set of scalar measures given by* (2.1) *(with $N = 1$). Suppose that the positive scalar measure μ has the form* (2.3) *(tailored to the scalar case):*

$$\mu = \sum_{j=1}^{n} w_j \delta_{x_j} \tag{2.7}$$

where $x_1, \dots, x_n$ are distinct points in X ($1 \leqslant n \leqslant m+1$) and $w_j \in \mathbb{R}$ are subject to

$$w_j > 0 \quad \textit{for } 1 \leqslant j \leqslant n; \quad \sum_{j=1}^{n} w_j = 1; \quad \sum_{j=1}^{n} \phi_i(x_j) w_j = 0 \quad \textit{for } i = 1, \dots, m. \tag{2.8}$$

Denote by $\boldsymbol{\phi}(x_j)$ the vector $\boldsymbol{\phi}(x_j) = \begin{bmatrix} \phi_1(x_j) \\ \vdots \\ \phi_m(x_j) \end{bmatrix}$ in $\mathbb{R}^m$ for $j = 1, \dots, n$. Then $\mu \in \partial_{\mathrm{e}} \mathcal{C}(X, 1, \boldsymbol{\phi})$ if and only if $0 = \sum_{j=1}^{n} w_j \boldsymbol{\phi}(x_j)$ is an interior point of the convex hull of $\{\boldsymbol{\phi}(x_1), \dots, \boldsymbol{\phi}(x_n)\}$ in $\mathbb{R}^m$.

Proof. By Theorem 2.2 tailored to the scalar-valued case, we know that any $\mu \in \partial_{\mathrm{e}} \mathcal{C}$ has the form (2.7) with base points $x_1, \dots, x_n$ and weights $w_1, \dots, w_n$ subject to (2.8); the question is: which such μs are actually extreme points?

Note that conditions (2.8) can be interpreted as exhibiting $0 \in \mathcal{R}^m$ as lying in the convex hull of $\{\boldsymbol{\phi}(x_1), \dots, \boldsymbol{\phi}(x_n)\}$. It remains to show that the measure $\mu = \sum_{j=1}^{n} w_j \delta_{x_j}$ is an extreme point of $\mathcal{C}(X, 1, \boldsymbol{\phi})$ if and only if in fact 0 is in the interior of the convex hull of $\{\boldsymbol{\phi}(x_1), \dots, \boldsymbol{\phi}(x_n)\}$.

Let us suppose the $0 = \sum_{j=1}^{n} w_j \boldsymbol{\phi}(x_j)$ is *not* in the interior of the convex hull. By statement (iii) in Proposition 1.3, this is the same as the existence of real numbers $c_1, \dots, c_n$ not all zero with

$$\sum_{j=1}^{n} c_j = 0, \quad \sum_{j=1}^{n} c_j \boldsymbol{\phi}(x_j) = 0.$$

Define a measure $\nu = \sum_{j=1}^{n} c_j \delta_{x_j}$. Then $\nu \neq 0$, $\nu(X) = 0$ and $\nu(\phi_j) = \int_X \phi_j \mathrm{d}\nu = \sum_{j=1}^{n} c_j \phi(x_j) = 0$. If we choose $\varepsilon > 0$ sufficiently small, then $\mu \pm \varepsilon \nu \in \mathcal{C}(1, X, \boldsymbol{\phi})$. We conclude by Lemma 1.1 that μ is not extremal in $\mathcal{C}(X, 1, \boldsymbol{\phi})$.

Suppose next that $0 = \sum_{j=1}^{n} w_j \boldsymbol{\phi}(x_j)$ is an interior point of the convex hull of $\{\boldsymbol{\phi}(x_1), \dots, \boldsymbol{\phi}(x_n)\}$ in $\mathbb{R}^m$. We wish to show that then $\mu \in \partial_{\mathrm{e}} \mathcal{C}(X, 1, \boldsymbol{\phi})$. We therefore suppose that $\mu = t_1 \mu_1 + t_2 \mu_2$ with $\mu_k \in \mathcal{C}(X, 1, \boldsymbol{\phi})$ and $t_k > 0$ for each $k = 1, 2$ and $t_1 + t_2 = 1$. Since μ_k is a positive measure for each k, we read off from (2.7) that $\operatorname{supp} \mu_k \subset \{x_1, \dots, x_n\}$, so each μ_k has the form $\mu_k = \sum_{j=1}^{n} w_j^{(k)} \delta_{x_j}$

for some weights $w_j^{(k)} \geqslant 0$ with $\sum_{j=1}^{n} w_j^{(k)} = 1$. From the fact that $\mu_k \in \mathcal{C}(X, 1, \boldsymbol{\phi})$ we also have that $\mu_k(\phi_i) = \sum_{j=1}^{n} w_j^{(k)} \phi_i(x_j) = 0$ for each $i = 1, \dots, m$, or, in vectorial form, $\sum_{j=1}^{n} w_j^{(k)} \boldsymbol{\phi}(x_j) = 0 \in \mathbb{R}^m$. By the assumption that $0 = \sum_{j=1}^{n} w_j \boldsymbol{\phi}(x_j)$ is in interior point for the convex hull of $\{\boldsymbol{\phi}(x_1), \dots, \boldsymbol{\phi}(x_n)\}$ in $\mathbb{R}^m$, statement (i) in Proposition 1.3 gives us that $w_j^{(k)} = w_j$ for $j = 1, \dots, n$ for each $k = 1, 2$, i.e., $\mu_k = \mu$. We conclude that μ is indeed an extreme point of $\mathcal{C}(X, 1, \boldsymbol{\phi})$ as wanted. ∎

COROLLARY 2.4. *Suppose that we are given a data set*

$$X = \text{ a compact Hausdorff space}, \quad \boldsymbol{\phi} = \{\phi_1, \dots, \phi_m\} \subset C_{\mathbb{R}}(X)$$

and we let $\mathcal{C}(X, 1, \boldsymbol{\phi})$ be as in (2.1) *(with $N = 1$). Then $\mathcal{C}(X, 1, \boldsymbol{\phi}) \neq \emptyset$ if and only if, for some natural number n with $1 \leqslant n \leqslant m + 1$, there exists a collection of n distinct points $x_1, \dots, x_n$ in X such that $0 = \sum_{j=1}^{n} w_j \boldsymbol{\phi}(x_j)$ is an interior point of the convex hull of the set $\{\boldsymbol{\phi}(x_1), \dots, \boldsymbol{\phi}(x_n)\}$ in $\mathbb{R}^m$, where $\boldsymbol{\phi}(x_j) = \begin{bmatrix} \phi_1(x_j) \\ \vdots \\ \phi_m(x_j) \end{bmatrix}$.*

Proof. By the Kreĭn–Milman theorem, $\mathcal{C}(X, 1, \boldsymbol{\phi})$ has extreme points if and only if $\mathcal{C}(X, 1, \boldsymbol{\phi})$ is not empty. The conclusion is now immediate from Theorem 2.3. ∎

REMARK 2.5. One can interpret Theorem 2.3 even for the case $m = 0$. In this case $1 \leqslant n \leqslant m + 1 = 1$ forces $n = 1$. The constraints (2.8) force μ to have the form $\mu = \delta_x$ for some $x \in X$. As there are no ϕs, the condition that 0 be an interior point of the convex hull of $\{\boldsymbol{\phi}(x_1), \dots, \boldsymbol{\phi}(x_n)\}$ can be interpreted to hold vacuously. This recovers the correct result that the set of extreme points of the normalized ball $\{\mu \in M(X)_+ : \mu(X) = 1\}$ consists of the unit point masses $\{\delta_x : x \in X\}$.

REMARK 2.6. Another special case of Theorem 2.3 of interest is the case where $X = \mathbb{T}$ is the unit circle in the complex plane, $m = 2$ with $\phi_1(z) = \operatorname{Re} z$ and $\phi_2(z) = \operatorname{Im} z$. In this case one can give an explicit geometric characterization of $\partial_e \mathcal{C}$. Indeed, pairs of points with $\mathbf{0} \in \mathbb{R}^2 \cong \mathbb{C}$ correspond to antipodal points on the unit circle, and triples of points x_1, x_2, x_3 on the unit circle with $\mathbf{0} \in \operatorname{conv}^0\{x_1, x_2, x_3\}$ amount to non-collinear points on the unit circle having $\mathbf{0} \in \mathbb{C}$ in the interior of the simplex spanned by x_1, x_2, x_3. This analysis has been worked out by Dritschel and Pickering in [15]. Motivation for this example comes from the search for a collection of test functions for the Schur class associated with the constrained H^∞ class $H_1^\infty = \{f \in H^\infty : f'(0) = 0\}$ (see [11] and Remark 4.7 below). In this context there is an additional equivalence relation imposed on

$\partial_{\mathrm{e}}\mathcal{C}(\mathbb{T}, 1, \boldsymbol{\phi})$ and the set of equivalence classes of $\partial_{\mathrm{e}}\mathcal{C}(\mathbb{T}, 1, \boldsymbol{\phi})$ can be identified topologically with the unit sphere.

A similar albeit less explicit analysis can be worked out for the more general subalgebras $H^\infty_B = \mathbb{C} + BH^\infty$ (where B is an inner function) studied in [30], at least for the case where B is a finite Blaschke product. For example, if $B(z) = \prod\limits_{i=1}^{\kappa} (\frac{z-\alpha_i}{1-\overline{\alpha}_i z})^{m_i}$ for distinct points $\alpha_1, \ldots, \alpha_\kappa \in \mathbb{D}$ with $\sum\limits_{i=1}^{\kappa} m_i \geqslant 2$, then the algebra H^∞_B has the explicit function-theoretic description involving $\sum\limits_{i=1}^{\kappa} m_i - 1$ complex linear constraints:

$$\begin{aligned} H^\infty_B = \{f \in H^\infty : f(\alpha_i) = f(\alpha_1) \text{ for } 2 \leqslant i \leqslant \kappa; \\ \frac{\mathrm{d}^k f}{\mathrm{d}z^k}(\alpha_i) = 0 \text{ for } 2 \leqslant k \leqslant m_i, i = 1, \ldots, \kappa\}. \end{aligned}$$

We let the collection $\boldsymbol{\phi}$ consist of the $2\Big(\sum\limits_{i=1}^{\kappa} m_i - 1\Big)$ real-valued continuous functions

$$\begin{aligned} \phi_{i,1,\mathrm{Re}}(\zeta) &= \mathrm{Re}\,\Big(\frac{\alpha_i + \zeta}{\alpha_i - \zeta} - \frac{\alpha_1 + \zeta}{\alpha_1 - \zeta}\Big) \quad \text{for } 2 \leqslant i \leqslant \kappa, \\ \phi_{i,1,\mathrm{Im}}(\zeta) &= \mathrm{Im}\,\Big(\frac{\alpha_i + \zeta}{\alpha_i - \zeta} - \frac{\alpha_1 + \zeta}{\alpha_1 - \zeta}\Big) \quad \text{for } 2 \leqslant i \leqslant \kappa, \\ \phi_{i,k,\mathrm{Re}}(\zeta) &= \mathrm{Re}\,\frac{\mathrm{d}^k}{\mathrm{d}z^k}\Big(\frac{z+\zeta}{z-\zeta}\Big)\Big|_{z=\alpha_i} \quad \text{for } 2 \leqslant k \leqslant m_i,\ i = 1, \ldots, \kappa, \\ \phi_{i,k,\mathrm{Im}}(\zeta) &= \mathrm{Im}\,\frac{\mathrm{d}^k}{\mathrm{d}z^k}\Big(\frac{z+\zeta}{z-\zeta}\Big)\Big|_{z=\alpha_i} \quad \text{for } 2 \leqslant k \leqslant m_i,\ i = 1, \ldots, \kappa. \end{aligned}$$

We may apply the result of Theorem 2.3 to obtain a reasonably explicit characterization of the extreme points μ of the associated compact convex set of measures $\mathcal{C}(\mathbb{T}, 1, \boldsymbol{\phi})$. For μ such an extreme point, the associated Herglotz function $F_\mu(z) = \int\limits_{\mathbb{T}} \frac{z+\zeta}{z-\zeta}\, \mathrm{d}\mu(\zeta)$ is an extreme point of the normalized Herglotz class

$$\mathcal{H}_B = \{F = (I - S)(I + S)^{-1} : S \in (H^\infty_B)^{N \times N} \text{ with } \|S\| \leqslant 1 \text{ and } S(0) = 0\}.$$

It then follows by the same argument as in [15] corresponding to the special case $B(z) = z^2$ that the functions

$$\{S_\mu = (F_\mu + I)^{-1}(F_\mu - I) : \mu \text{ extreme for } \mathcal{C}(\mathbb{T}, 1, \boldsymbol{\phi})\}$$

form a collection of test functions for the algebra H^∞_B. This complements the other results obtained in [30] for the algebras H^∞_B.

2.2. RETURN TO THE GENERAL MATRIX-VALUED CASE. We now indicate how one can analyze the general case of Problem 2.1 by using the language of noncommutative convexity (see [17], [19], [22], [23]).

Rather than delve into the general setting of C^*-convex combinations of elements of a C^*-algebra or of the generalized state space of a C^*-algebra and associated C^*-convex subsets and C^*-extreme points, we discuss only the concrete special case which we need for our application (but see Remark 2.11 below). We fix a positive integer N and consider a collection of n vectors in the space $\mathcal{X} := ([\mathbb{C}^{N\times N})]_{\mathrm{h}})^{m\times 1}$, i.e., column vectors of length m, each entry of which is an $N \times N$ complex Hermitian matrix. Given a collection of n elements $\Phi^{(1)}, \dots, \Phi^{(n)}$ in $\mathcal{X}$, we say that $\Phi \in \mathcal{X}$ is a C^**-convex combination* of $\Phi^{(1)}, \dots, \Phi^{(n)}$ if there are n matrices $A_1, \dots, A_n$ of size $N \times N$ with $\sum_{j=1}^{n} A_j^* A_j = I$ so that

$$\Phi = \sum_{j=1}^{n} A_j^* \Phi^{(j)} A_j \tag{2.9}$$

where we set

$$A_j^* \Phi^{(j)} A_j = \begin{bmatrix} A_j^* \Phi_1^{(j)} A_j \\ \vdots \\ A_j^* \Phi_m^{(j)} A_j \end{bmatrix} \quad \text{if } \Phi^{(j)} = \begin{bmatrix} \Phi_1^{(j)} \\ \vdots \\ \Phi_m^{(j)} \end{bmatrix} \in ([\mathbb{C}^{N\times N}]_{\mathrm{h}})^{m\times 1}. \tag{2.10}$$

For our application, we only deal with the special case where $\Phi_i^{(j)}$ is a scalar multiple of the identity: $\Phi_i^{(j)} = \phi_i^{(j)} I_N$ where $\phi_i^{(j)}$ is a real number; we denote the subspace of all such elements of $\mathcal{X}$ by $\mathcal{X}_s$. For $\Phi^{(1)}, \dots, \Phi^{(n)} \in \mathcal{X}_s$, the C^*-convex combination (2.9) and (2.10) simplifies to

$$\Phi = \sum_{j=1}^{n} W_j \Phi^{(j)} \tag{2.11}$$

where we set $W_j = A_j^* A_j$, so $\{W_j : j = 1, \dots, n\}$ is any collection of $N \times N$ matrices satisfying

$$W_j \geqslant 0 \quad \text{for } j = 1, \dots, n, \quad \sum_{j=1}^{n} W_j = I_N \tag{2.12}$$

and the meaning of the j-th term in (2.11) is

$$W_j \Phi^{(j)} = \begin{bmatrix} \phi_1^{(j)} W_j \\ \vdots \\ \phi_m^{(j)} W_j \end{bmatrix} \quad \text{if } \Phi^{(j)} = \begin{bmatrix} \phi_1^{(j)} I_N \\ \vdots \\ \phi_m^{(j)} I_N \end{bmatrix}. \tag{2.13}$$

We shall furthermore only be interested in the case where the C^*-convex combination of such $\Phi^{(1)}, \dots, \Phi^{(n)}$ in $\mathcal{X}_s$ is the zero element $\mathbf{0}$ in $\mathcal{X}$:

$$\mathbf{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in ([\mathbb{C}^{N\times N}]_{\mathrm{h}})^{m\times 1}.$$

In analogy with the notion of the interior point of the convex hull of a collection of vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ for the classical case presented in Section 1, we propose the following definition of the notion that $\mathbf{0}$ is an interior point of the C^*-convex hull of a collection of vectors in $\mathcal{X}_s$. The statement of the result requires some additional terminology, all of which we collect in the following definition.

DEFINITION 2.7. Given an operator T on a Hilbert space $\mathcal{H}$ (e.g., $\mathcal{H} = \mathbb{C}^N$ and T presented as a matrix in $\mathbb{C}^{N\times N}$) together with a closed subspace $\mathcal{M}$ of $\mathcal{H}$, we say that T *is supported on* $\mathcal{M}$ if $T = TP_{\mathcal{M}} = P_{\mathcal{M}}T$ (where $P_{\mathcal{M}}$ is the orthogonal projection from $\mathcal{H}$ to $\mathcal{M}$).

We say that a family of closed subspaces $\{\mathcal{M}_1, \dots, \mathcal{M}_n\}$ is *weakly independent* if, *whenever $T_1, \dots, T_n$ are linear operators on $\mathcal{H}$ with*

$$T_j \text{ is supported on } \mathcal{M}_j \text{ for each } j \text{ and } \sum_{j=1}^{n} T_j = 0, \tag{2.14}$$

it follows that $T_j = 0$ for each $j = 1, \dots, n$.

Suppose that in addition we are given a collection $\boldsymbol{\phi} = \{\phi^{(1)}, \dots, \phi^{(n)}\}$ of n vectors in $\mathbb{R}^m$ (so $\phi^{(j)} = \begin{bmatrix} \phi_1^{(j)} \\ \vdots \\ \phi_m^{(j)} \end{bmatrix}$ with $\phi_1^{(j)}, \dots, \phi_m^{(j)} \in \mathbb{R}$). Then we say that the family of closed subspaces $\{\mathcal{M}_1, \dots, \mathcal{M}_n\}$ is *$\boldsymbol{\phi}$-constrained weakly independent* if, *whenever* $T_1, \dots, T_n$ are linear operators on $\mathcal{H}$ satisfying the conditions

$$\begin{aligned} &\text{(i) } T_j \text{ is supported on } \mathcal{M}_j \text{ for each } j, \\ &\text{(ii) } \sum_{j=1}^{n} T_j = 0, \\ &\text{(iii) } \sum_{j=1}^{n} \phi_i^{(j)} T_j = 0 \text{ for each } i = 1, \dots, m, \end{aligned} \tag{2.15}$$

it follows that $T_j = 0$ for each $j = 1, \dots, n$.

Finally, suppose that we are given n-vectors $\boldsymbol{\phi} = \left\{ \begin{bmatrix} \phi_1^{(j)} \\ \vdots \\ \phi_m^{(j)} \end{bmatrix} : j = 1, \dots, n \right\}$ in $\mathbb{R}^m$ with associated set of n vectors $\boldsymbol{\Phi} = \left\{ \begin{bmatrix} \phi_1^{(j)} I_N \\ \vdots \\ \phi_m^{(j)} I_N \end{bmatrix} : j = 1, \dots, n \right\}$ in $\mathcal{X}_s$, and suppose that 0 is in the C^*-convex hull of $\boldsymbol{\Phi}$: there are matrices $W_1, \dots, W_n$ satisfying conditions (2.12) so that

$$0 = \sum_{j=1}^{n} W_j \Phi^{(j)}$$

with $W_j \Phi^{(j)}$ defined as in (2.13). Then we say that 0 *is an interior point of the* C^*-*convex hull* of $\{\Phi^{(j)} : j = 1, \dots, n\}$ if *the family of subspaces* $\{\operatorname{Ran} W_1, \dots, \operatorname{Ran} W_n\}$ *is* $\boldsymbol{\phi}$*-constrained weakly independent.*

An easy observation is that for the case $N = 1$, the notion of 0 being an interior point of the C^*-convex hull of $\boldsymbol{\Phi} = \{\phi^{(1)}, \dots, \phi^{(n)}\} \subset \mathbb{R}^m$ coincides with 0 being an interior point of the convex hull of $\{\phi^{(1)}, \dots, \phi^{(n)}\}$ as characterized in Proposition 1.3. Indeed, supposes that $0 = \sum\limits_{k=1}^{n} w_k \phi^{(k)}$ for positive numbers $w_1, \dots, w_n$ summing to 1, and $t_1, \dots, t_n$ is a collection of real numbers with

$$\sum_{k=1}^{n} t_k = 0, \quad \sum_{k=1}^{n} t_k \phi^{(k)} = 0 \in \mathbb{R}^m.$$

Since $\operatorname{Ran} w_k$ is the whole space $\mathbb{C}$ (when w_k is considered as an operator on $\mathbb{C}$), it is automatically the case that t_k is supported on $\operatorname{Ran} w_k$. Thus the condition for 0 being an interior point of the C^*-convex hull of $\{\phi^{(1)}, \dots, \phi^{(n)}\}$ reduces to condition (iii) in Proposition 1.3 (with $\phi^{(j)}$ in place of $\mathbf{u}_j$), i.e., to 0 being an interior point of the classical convex hull of $\{\phi^{(1)}, \dots, \phi^{(n)}\}$.

We are now ready to state the following general result concerning Problem 2.1.

THEOREM 2.8. *Let the convex set of measures* $\mathcal{C} = \mathcal{C}(X, N, \boldsymbol{\phi}\}$ *be given as in* (2.1). *Then a measure* μ *in* $\mathcal{C}$ *is extremal* ($\mu \in \partial_{\mathrm{e}} \mathcal{C}$) *if and only if there is a natural number* n *with* $1 \leqslant n \leqslant (m+1)N^2$ *and* n *distinct points* $\mathbf{x} = (x_1, \dots, x_n)$ *in* X *together with* $N \times N$ *matrix weights* $W_1, \dots, W_n$ *satisfying the conditions* (2.2) *so that* μ *has a representation as in Theorem* 2.2 *(see* (2.3)*)*

$$\mu = \sum_{j=1}^{n} W_j \delta_{x_j} \tag{2.16}$$

where, in addition, the family of subspaces $\{\operatorname{Ran} W_1, \dots, \operatorname{Ran} W_n\}$ *is* $\boldsymbol{\phi}(\mathbf{x})$*-constrained weakly independent, where we set*

$$\boldsymbol{\phi}(\mathbf{x}) = \{\phi(x_1), \dots, \phi(x_n)\}.$$

Proof. Suppose first that μ has the form (2.16) with $\{\operatorname{Ran} W_1, \dots, \operatorname{Ran} W_n\}$ a $\boldsymbol{\phi}$-constrained weakly independent family of subspaces, and also suppose that ν is a complex Hermitian $N \times N$-matrix measure on X such that

$$\nu(X) = 0; \quad \nu(\phi_i) = \int_X \phi_i \mathrm{d}\nu = 0 \quad \text{for } i = 1, \dots, m; \quad \mu \pm \nu \geqslant 0. \tag{2.17}$$

From the last of conditions (2.17) we see that $\operatorname{supp} \nu \subset \{x_1, \dots, x_m\}$ and hence there are complex Hermitian matrices $T_1, \dots, T_n$ so that

$$\nu = \sum_{j=1}^{n} T_j \delta_{x_j}.$$

By evaluating $\mu \pm \nu$ on the singleton Borel set $\{x_j\}$, we see that $W_j \pm T_j \geqslant 0$. This enables us to conclude that T_j is supported on $\operatorname{Ran} W_j$ for each j. From the first two conditions in (2.17) we deduce that

$$\sum_{j=1}^{n} T_j = 0; \quad \sum_{j=1}^{n} \phi_i(x_j) T_j = 0 \quad \text{for } i = 1, \dots, m.$$

From the hypothesis that $\{\operatorname{Ran} W_j : j = 1, \dots, n\}$ is $\boldsymbol{\phi}(\mathbf{x})$-constrained weakly independent, we conclude that $T_j = 0$ for each j, and hence $\nu = 0$. From the criterion in Lemma 1.1, we now conclude that μ is extremal as wanted.

Conversely, suppose that $\mu \in \partial_{\mathrm{e}} \mathcal{C}$ and suppose that $\{T_1, \dots, T_n\}$ is a collection of operators satisfying the conditions (2.15) (with $\operatorname{Ran} W_j$ in place of $\mathcal{M}_j$). Note that $\{\operatorname{Re} T_1, \dots, \operatorname{Re} T_n\}$ and $\{\operatorname{Im} T_1, \dots, \operatorname{Im} T_n\}$ satisfy the same hypotheses and, in order to show that $T_j = 0$, it suffices to show that $\operatorname{Re} T_j = 0$ and $\operatorname{Im} T_j = 0$. Thus without loss of generality we may assume that T_j is complex Hermitian. Define a measure ν by $\nu = \varepsilon \sum\limits_{j=1}^{n} T_j \delta_{x_j}$ where $\varepsilon > 0$. One can check that then ν meets all the conditions (2.17) as long as $\varepsilon > 0$ is chosen sufficiently small. If μ is extremal, then Lemma 1.1 forces $\nu = 0$. As we were careful to arrange that $\varepsilon \neq 0$, it follows that $T_j = 0$ for each $j = 1, \dots, n$. It now follows that indeed $\{\operatorname{Ran} W_j : j = 1, \dots, n\}$ is $\boldsymbol{\phi}(\mathbf{x})$-constrained weakly independent as was to be shown. ∎

It is of interest to specialize Theorem 2.8 to the case $m = 0$; in this way we recover a result of Arveson (see Theorem 1.4.10 of [5]).

COROLLARY 2.9. *Let* $\mathcal{C}(X, N, \varnothing)$ *be the compact convex set of positive* $N \times N$*-matrix measures* μ *on a compact Hausdorff space* X *normalized to have* $\mu(X) = I_N$*. Then* μ *is extremal in* $\mathcal{C}(X, N, \varnothing)$ *if and only if, for some natural number* n *with* $1 \leqslant n \leqslant N^2$*, there are* n *distinct points* $x_1, \dots, x_n$ *and* $N \times N$ *matrix weights* $W_1, \dots, W_n$ *satisfying:*

(i) $W_j \geqslant 0$ *for each* $j = 1, \dots, n$ *and* $\sum_{j=1}^{n} W_j = I_N$, *and*

(ii) *the family of subspaces* $\{\operatorname{Ran} W_1, \dots, \operatorname{Ran} W_n\}$ *is weakly independent so that* μ *is given by*

$$\mu = \sum_{k=1}^{n} W_k \delta_{x_k}. \tag{2.18}$$

Proof. Simply observe that this is just the $m = 0$ case of Theorem 2.8. We note that our proof (i.e., the proof of Theorem 2.8 specialized to the $m = 0$ case) is elementary and direct while the proof in [5] has a more sophisticated flavor bringing in ideas from C^*-representation and dilation theory. ∎

REMARK 2.10. In the paper of Arveson [5], it is not noted explicitly that the number of terms n in the decomposition (2.18) for an extremal measure of $\mathcal{C}(X, N, \emptyset)$ can be at most N^2 for the finite-dimensional case ($\mathcal{H} = \mathbb{C}^N$). However it is observed there (see page 165 of [5]) that, in the finite-dimensional case, weak independence of a family of subspaces $\{\mathcal{M}_1, \dots, \mathcal{M}_n\} \subset \mathbb{C}^N$ is equivalent to classical linear independence for the family of subspaces $\{\mathcal{N}_1, \dots, \mathcal{N}_n\} \subset \mathcal{C}^N \otimes \mathcal{C}^N$, where we have set $\mathcal{N}_j = \operatorname{span}\{\xi \otimes \eta : \xi, \eta \in \mathcal{M}_j\}$. Since $\dim(\mathbb{C}^N \otimes \mathbb{C}^N)$ is N^2, we have the bound N^2 on the number of subspaces in a weakly independent family of subspaces $\{\mathcal{M}_1, \dots, \mathcal{M}_n\}$ contained in $\mathbb{C}^N$.

There is also an example given in [5] of a weakly independent family of subspaces which is not linearly independent in the classical sense, e.g., $\mathcal{M}_1 = \operatorname{span}\{\xi\}$, $\mathcal{M}_2 = \operatorname{span}\{\eta\}$, $\mathcal{M}_3, = \operatorname{span}\{\xi + \eta\}$ where ξ and η are linearly independent vectors. This example can be enhanced as follows (see pages 32–35 of [21]). One can choose three vectors ξ_1, ξ_2, ξ_3 in $\mathbb{C}^2$ so that the family of subspaces $\mathcal{M}_j = \operatorname{span}\{\xi_j\}$ ($j = 1, 2, 3$) is weakly independent and in addition the associated matrix weights $W_j = \xi_j \xi_j^*$ ($j = 1, 2, 3$) satisfy the normalization $W_1 + W_2 + W_3 = I_2$. We conclude that the measure $\mu = W_1 \delta_{x_1} + W_2 \delta_{x_2} + W_3 \delta_{x_3}$ (where x_1, x_2, x_3 are any three distinct points in X) is extremal in $\mathcal{C}(X, 2, \emptyset)$ while not being a spectral measure, i.e., μ is not of the form $\mu = P_1 \delta_{x_1} + P_2 \delta_{x_2}$ with P_1, P_2 orthogonal projections with pairwise orthogonal ranges in $\mathbb{C}^2$ (compare with Theorem 2.14 below).

REMARK 2.11. The notion of C^*-convex combination and associated notions of C^*-convex set and C^*-extremal point are defined more broadly in the literature than what we have indicated so far here. The setting of [17], [19] is the generalized state space $S_{\mathcal{H}}(A)$ of unit-preserving completely positive maps from the C^*-algebra A into the C^*-algebra $\mathcal{L}(\mathcal{H})$ of bounded linear operators on the Hilbert space $\mathcal{H}$. A C^*-convex combination of n such maps $\phi_1, \dots, \phi_n$ is defined to be a ϕ given by

$$\phi(a) = \sum_{j=1}^{n} t_j^* \phi_j(a) t_j$$

where $t_j \in \mathcal{L}(\mathcal{H})$ satisfy $\sum_{j=1}^{n} t_j^* t_j = I_{\mathcal{H}}$. The main interest in [17], [19] (as well as in other papers) is the structure of C^*-extreme points of $S_{\mathcal{H}}(A)$. More broadly, one could consider real maps $\phi : A \to \mathcal{L}(\mathcal{H})$, i.e., maps which preserve selfadjoint elements, and in particular, examine when the zero map 0 is a C^*-convex combination of n given such maps $\phi_1, \dots, \phi_n$. This becomes exactly the setting introduced in Section 2.2 if we take A to be the C^*-algebra of continuous functions on the finite-point set $\{1, \dots, m\}$, i.e., $A = C(\{1, \dots, m\}) \cong \mathbb{C}^m$ (so selfadjoint elements are identified with $\mathbb{R}^m$), $\mathcal{H} = \mathbb{C}^N$, and identify $\mathcal{X} = ([\mathbb{C}^{N \times N}]_{\mathrm{h}})^{m \times 1}$ with maps from $C(\{1, \dots, m\})$ into $\mathcal{L}(\mathbb{C}^N) \cong \mathbb{C}^{N \times N}$:

$$\Phi = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_m \end{bmatrix} \in \mathcal{X} \mapsto \phi_\Phi : f \in C(\{1, \dots, m\}) \mapsto f(1)\Phi_1 + \cdots + f(m)\Phi_m.$$

We have not seen the notion of *interior point of the C^*-convex hull* elsewhere in the literature. Note that we define this notion here only for the special case where $\Phi_1, \dots, \Phi_n$ are in $\mathcal{X}_s$; we do not hazard a guess here as to what the appropriate notion should be for the more noncommutative situation where $\Phi_1, \dots, \Phi_n \in \mathcal{X}$, or for the still more general situation where $\Phi_1, \dots, \Phi_n$ are real elements of $\mathcal{L}(A, \mathcal{L}(\mathcal{H}))$.

As observed in [19], given a compact Hausdorff space X, the generalized state space $S_{\mathcal{H}}(C(X))$ of the commutative C^*-algebra $C(X)$ can be identified with positive $\mathcal{L}(\mathcal{H})$-valued measures μ on X having total mass $\mu(X)$ equal to $I_{\mathcal{H}}$. Thus, when $\mathcal{H} = \mathbb{C}^N$, $S_{\mathbb{C}^N}(C(X))$ is exactly the convex set $\mathcal{C}(X, N, \emptyset)$ whose classical extreme points are described in Corollary 2.9. One of the central goals in [17], [19] is to describe the C^*-extreme points of $S_{\mathcal{H}}(C(X))$. It is interesting that the problem of describing the classical extreme points of the linearly-constrained generalized state space $\mathcal{C}(X, N, \boldsymbol{\phi})$, a problem formulated completely in the confines of classical convexity theory, has a solution (see Theorem 2.8) which draws on ideas from noncommutative convexity theory.

We note that other papers (e.g. [16], [22], [23]) study C^*-convex sets (and associated extremal-point theory) in $\mathcal{L}(\mathcal{H})$ or, more generally, in a general C^*-algebra A. From our point of view this amounts to the special case $m = 1$.

It can be argued that the characterization of $\partial_{\mathrm{e}} \mathcal{C}(N, X, \boldsymbol{\phi})$ in Theorem 2.8 is not particularly explicit and is a little difficult to work with. To address this issue, we give a couple of illustrative examples where more detailed information is possible.

THEOREM 2.12. *Suppose that $\mathcal{C}(X, N, \boldsymbol{\phi})$ is as in* (2.1). *Suppose that the measure $\mu \in [M(X)^{N \times N}]_{\mathrm{h}}$ has the form*

$$\mu = \sum_{k=1}^{n} \mu_k L_k \tag{2.19}$$

where:

(i) *each μ_k is a scalar positive measure which is an extreme point for the associated convex compact subset $\mathcal{C}(X, 1, \boldsymbol{\phi})$ of positive scalar measures where in addition the support sets $\mathcal{S}_k := \{\operatorname{supp} \mu_k : k = 1, \dots, n\}$ are disjoint ($\mathcal{S}_k \cap \mathcal{S}_{k'} = \emptyset$ for $k \neq k'$);*

(ii) *the matrix weights L_k satisfy the conditions*

$$L_k \geqslant 0 \quad \textit{for each } k; \quad \sum_{k=1}^{n} L_k = I_N;$$

(iii) *the family of subspaces $\{\operatorname{Ran} L_k : k = 1, \dots, n\}$ is weakly independent (as defined in Definition 2.7).*
Then $\mu \in \partial_e \mathcal{C}(X, N, \boldsymbol{\phi})$.

Proof. We argue that if μ is as in the statement of the theorem, then it meets the conditions of Theorem 2.8 and therefore is extremal in $\mathcal{C}(X, N, \boldsymbol{\phi})$. Toward this end, we let $\{x_{k,1}, \dots, x_{k,n_k}\}$ be the support of the measure μ_k. By hypothesis, these points are all distinct. We then write μ_k as

$$\mu_k = \sum_{j=1}^{n_k} w_j^{(k)} \delta_{x_{k,j}}$$

for scalar weights $w_j^{(k)} > 0$ and then rewrite μ as

$$\mu = \sum_{k=1}^{n} \sum_{j=1}^{n_k} w_j^{(k)} L_k \delta_{x_{k,j}} = \sum_{k,j: 1 \leqslant k \leqslant n; 1 \leqslant j \leqslant n_k} W_{k,j} \delta_{x_{k,j}} \tag{2.20}$$

where we have set $W_{k,j} = w_j^{(k)} L_k$. Then (2.20) represents μ in the form (2.16), but with index set $\{(i,j) : 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n_k\}$ rather than $\{j : 1 \leqslant j \leqslant n\}$. It remains only to show that the collection of subspaces $\{\operatorname{Ran} W_{k,j} : 1 \leqslant k \leqslant n, 1 \leqslant j \leqslant n_k\}$ is $\boldsymbol{\phi}$-constrained weakly independent.

We therefore suppose that we are given a collections of operators $T_{k,j}$ on $\mathbb{C}^N$ satisfying:

$$\begin{aligned} & T_{k,j} \text{ is supported on } \operatorname{Ran} W_{k,j} = \operatorname{Ran} L_k \text{ for each } (k,j); \\ & \sum_{k,j: 1 \leqslant k \leqslant n; 1 \leqslant j \leqslant n_k} T_{k,j} = 0; \\ & \sum_{k,j: 1 \leqslant k \leqslant n; 1 \leqslant j \leqslant n_k} \phi_i(x_{k,j}) T_{k,j} = 0 \text{ for } i = 1, \dots, m; \end{aligned}$$

with the goal to show that each $T_{k,j} = 0$. From the hypothesis that $\{\operatorname{Ran} L_k : k = 1, \dots, n\}$ is weakly independent and the observation that both $\sum_{j=1}^{n_k} T_{k,j}$ and

$\sum_{j=1}^{n_k} \phi_i(x_{k,j}) T_{k,j}$ are supported on $\operatorname{Ran} L_k$, it follows that

$$\sum_{j=1}^{n_k} T_{k,j} = 0 \quad \text{and} \quad \sum_{j=1}^{n_k} \phi_i(x_{k,j}) T_{k,j} = 0 \quad \text{for each } i \text{ and } k. \tag{2.21}$$

We next use that $\mu_k = \sum_{j=1}^{n_k} w_j^{(k)} \delta_{x_{k,j}}$ is a scalar extreme point: it follows from Theorem 2.3 that $0 = \sum_{j=1}^{n_k} w_j^{(k)} \boldsymbol{\phi}(x_{k,j})$ is an interior point of the convex hull of the vectors $\boldsymbol{\phi}(x_{k,1}), \dots, \boldsymbol{\phi}(x_{k,n_k})$ in $\mathbb{R}^m$. By criterion (iii) in Proposition 1.3, the conditions (2.21) applied entrywise now force that $T_{k,j} = 0$ for each $j = 1, \dots, n_k$. As k is arbitrary, we have shown that $T_{k,j} = 0$ as required. ∎

The next result gives a partial converse to Theorem 2.12.

THEOREM 2.13. *Suppose that $\mu \in \mathcal{C}(X, N, \boldsymbol{\phi})$ has the form* (2.19) *such that:*

(i) *each μ_k is a scalar positive measure which is an extreme point for $\mathcal{C}(X, 1, \boldsymbol{\phi})$ (with supports not necessarily disjoint),*

(ii) *each W_k is positive semidefinite, and*

(iii) *the family of subspaces $\{\operatorname{Ran} W_k : k = 1, \dots, n\}$ is not weakly independent.*

Then μ is not an extreme point of $\mathcal{C}$.

Proof. The assumption that $\{\operatorname{Ran} W_k : k = 1, \dots, n\}$ is *not* weakly independent means that we can find a family of operators $\{T_k : k = 1, \dots, n\}$ on $\mathbb{C}^N$ such that $T_k \neq 0$ for some k, $\sum_{k=1}^{N} T_k = 0$ and T_k is supported on $\operatorname{Ran} W_k$ for each k. By considering either $\{\operatorname{Re} T_k : k = 1, \dots, N\}$ or $\{\operatorname{Im} T_k : k = 1, \dots, N\}$, we may suppose without loss of generality that each T_k is complex Hermitian. By choosing $\varepsilon > 0$ but sufficiently small, we can then arrange that $W_k \pm \varepsilon T_k \geqslant 0$ for each $k = 1, \dots, n$. We then define a measure ν by

$$\nu = \sum_{k=1}^{n} \varepsilon T_k \mu_k.$$

Then it is easily checked that $\nu \neq 0$, $\nu(\phi_i) = 0$ for each $i = 1, \dots, m$, $\nu(X) = 0$, and $\mu \pm \nu \geqslant 0$. As a consequence of Lemma 1.1 it follows that μ is not an extreme point of $\mathcal{C}(X, N, \boldsymbol{\phi})$. ∎

We now present another concrete class of extreme points for a general convex compact set of the form $\mathcal{C}(N, X, \boldsymbol{\phi})$ as in (2.1).

THEOREM 2.14. *Suppose that $\mu \in [M(X)^{N \times N}]_+$ has the form*

$$\mu = \sum_{k-1}^{N} \mu_k P_k \tag{2.22}$$

where $\{P_1, \dots, P_N\}$ is a pairwise-orthogonal family of rank-1 orthogonal projections summing to the identity operator I on $\mathbb{C}^N$ and where $\mu_1, \dots, \mu_N$ are scalar measures (not necessarily having disjoint supports and perhaps not even distinct) which are extremal for the compact convex set of scalar measures $\mathcal{C}(X, 1, \boldsymbol{\phi})$. Then μ is extremal in the compact convex set of matrix measures $\mathcal{C}(X, N, \boldsymbol{\phi})$.

Proof. Let μ be as in the statement of the theorem and suppose that ν is a complex Hermitian $N \times N$ matrix measure with

$$\nu(X) = 0; \quad \nu(\phi_i) = 0 \quad \text{for } i = 1, \dots, m; \quad \mu \pm \nu \geqslant 0. \tag{2.23}$$

Factor the rank-1 orthogonal projection P_k as $P_k = e_k e_k^*$ for a unit column vector $e_k \in \mathbb{C}^{N \times N}$. Then we have

$$0 \leqslant P_k(\mu \pm \nu) P_k = (\mu_k \pm \nu_{kk}^{(1)}) P_k$$

where $\nu_{kk}^{(1)}$ is the scalar measure given by $\nu_{kk}^{(1)} = e_k^* \nu e_k$. The conditions (2.23) satisfied by ν imply that each $\nu_{kk}^{(1)}$ satisfies the conditions

$$\nu_{kk}^{(1)}(X) = 0; \quad \nu_{kk}^{(1)}(\phi_i) = 0 \quad \text{for } i = 1, \dots, m; \quad \mu_k \pm \nu_{kk}^{(1)} \geqslant 0.$$

Since μ_k is extremal for $\mathcal{C}(X, 1, \boldsymbol{\phi})$, a consequence of Lemma 1.1 is that $\nu_{kk}^{(1)} = 0$ for each $k = 1, \dots, n$. It remains to show that the off-diagonal components of ν with respect to the orthonormal basis $\{e_1, \dots, e_N\}$ are also all zero, i.e., $\nu_{kk'}^{(1)} = 0$ for $1 \leqslant k < k' \leqslant N$, where $\nu_{kk'}^{(1)} = e_i^* \nu e_{k'}$.

Toward this goal, we consider the 2×2 matrix measure $\left[\begin{smallmatrix} e_k^* \\ e_{k'}^* \end{smallmatrix}\right] (\mu \pm \nu) \left[\begin{smallmatrix} e_k & e_{k'} \end{smallmatrix}\right]$ which we can identify with the 2×2 matrix measure

$$\begin{bmatrix} \underline{\mu_k} & \pm\nu_{kk'}^{(1)} \\ \pm\nu_{kk'}^{(1)} & \mu_{k'} \end{bmatrix}.$$

From the fact that $\mu \pm \nu \geqslant 0$, we see that this 2×2 matrix measure is positive for both choices of signs $\pm$. It follows that necessarily $\operatorname{supp} \nu_{kk'}^{(1)} \subset \operatorname{supp} \mu_k \cap \operatorname{supp} \mu_{k'}$. If $\{x_{k,1}, \dots, x_{k,n_k}\}$ is the support of μ_k, then necessarily $\nu_{kk'}$ has the form

$$\nu_{kk'}^{(1)} = \sum_{j=1}^{n_k} v_j^{kk'} \delta_{x_{k,j}}$$

for some weights $v_j^{kk'}$ (where $v_j^{kk'} = 0$ whenever $x_{k,j} \in \operatorname{supp} \mu_k$ is not in $\operatorname{supp} \mu_{k'}$). A consequence of the conditions (2.23) is the set of conditions on the weights $\{v_j^{kk'} : j = 1, \dots, n_k\}$:

$$\sum_{j=1}^{n_k} v_j^{kk'} = 0; \quad \sum_{j-1}^{n_k} \phi_i(x_{k,j}) v_j^{kk'} = 0 \quad \text{for } i = 1, \dots, m. \tag{2.24}$$

Since $\mu_k = \sum\limits_{j=1}^{n_k} w_j^{(k)} \delta_{x_{k,j}}$ is extremal for $\mathcal{C}(X, 1, \boldsymbol{\phi})$, we know that

$$0 = \sum_{j=1}^{n_k} w_j^{(k)} \boldsymbol{\phi}(x_{k,j})$$

is an interior point of the convex hull of $\{\boldsymbol{\phi}(x_{k,1}), \dots, \boldsymbol{\phi}(x_{k,n_k}\}$. By criterion (iii) in Proposition 1.3, conditions (2.24) then lead to the conclusion that $v_j^{kk'} = 0$ for $j = 1, \dots, n_k$. We conclude that $\nu_{kk'}^{(1)}$ is the zero measure. Since the pair of indices (k, k') is arbitrary, we now have that $\nu = 0$. An application of Lemma 1.1 then tells us that μ is extremal for $\mathcal{C}(X, N, \boldsymbol{\phi})$ as wanted. ∎

REMARK 2.15. By combining Theorems 2.14 and 2.8, we see that any measure of the form (2.22) must satisfy the conditions of Theorem 2.8 when expressed in the form (2.16). There does not appear to be any obvious direct proof of this implication.

REMARK 2.16. It is easily seen that the family of subspaces $\{\operatorname{Ran} P_k : 1 \leqslant k \leqslant N\}$ is weakly independent whenever $\{P_1, \dots, P_N\}$ is a pairwise-orthogonal family of orthogonal projections on $\mathbb{C}^N$. Let us say that a $\mu \in \mathcal{C}(X, N, \boldsymbol{\phi})$ (as in (2.1)) is *special* if μ has a presentation of the form

$$\mu = \sum_{k=1}^{n} \mu_k W_k \tag{2.25}$$

where each μ_k is a scalar positive measure extremal in the compact convex set of scalar measures $\mathcal{C}(X, 1, \boldsymbol{\phi})$ and where the family of subspaces $\{\operatorname{Ran} W_k : k = 1, \dots, n\}$ is weakly independent. Thus the extremal measures identified in Theorem 2.12 and those identified in Theorem 2.14 are all special, but the class of special measures is more general than either of these particular cases. For the special cases $N = 1$ or $m = 0$, we see that the set of extreme points $\partial_{\mathrm{e}}\mathcal{C}$ consists exactly of the special measures. More generally, in the examples which we have computed, it turns out that special measures are extremal, but we do not know if this always holds. On the other hand there are examples of extremal measures which are not necessarily special (see Corollary 5.1 below).

3. THE HERGLOTZ CLASS OVER A FINITELY CONNECTED PLANAR DOMAIN

In this section we let $\mathcal{R}$ denote a bounded domain (connected, open set) in the complex plane whose boundary consists of $m + 1$ smooth Jordan curves; we refer to [18], [20] as general references for the function theory on such domains. We let $\partial_0, \partial_1, \dots, \partial_m$ denote the $m + 1$ boundary components with ∂_0 denoting the boundary of the unbounded component of the complement $\mathbb{C} \setminus \mathcal{R}$ of $\mathcal{R}$ in the complex plane. For a fixed natural number N, we let $\mathcal{H}^N(\mathcal{R})$ denote the set of (single-valued) holomorphic $N \times N$ matrix-valued functions $F(z)$ on $\mathcal{R}$ with positive real part: Re $F(z) \geqslant 0$ for $z \in \mathcal{R}$. We often fix a point $t_0 \in \mathcal{R}$ and consider $\mathcal{H}^N(\mathcal{R})$ subject to the normalization $F(t_0) = I_N$; denote this normalized Herglotz class by $\mathcal{H}^N(\mathcal{R})_I$. In case $N = 1$ we write simply $\mathcal{H}(\mathcal{R})_1$.

A standard normal families argument combined with the classical Harnack inequality shows that $\mathcal{H}^N(\mathcal{R})_I$ is a compact convex subset of the locally convex topological space $\mathrm{Hol}(\mathcal{R})$ consisting of all holomorphic $N \times N$ matrix-valued functions on $\mathcal{R}$ with the topology of uniform pointwise convergence on compact subsets of $\mathcal{R}$. Our goal in this section is to apply the results of Section 2 to characterize the extreme points of $\mathcal{H}^N(\mathcal{R})_I$.

3.1. THE SCALAR-VALUED HERGLOTZ CLASS OVER $\mathcal{R}$. For the scalar-valued case $N = 1$, characterization of the extreme points of $\mathcal{H}(\mathcal{R})_1$ is worked out in various places (see [4], [13], [29]). The first step is to transform the problem to one of the form in Section 2.1 as follows. One can solve the Dirichlet problem on such domains: thus for given $u \in C_{\mathbb{R}}(\partial\mathcal{R})$, there is a unique function $u^\wedge \in C_{\mathbb{R}}(\mathcal{R}^-)$ (where $\mathcal{R}^-$ denotes the *closure* of $\mathcal{R}$) so that $u^\wedge|_{\mathcal{R}}$ is harmonic on $\mathcal{R}$ and

$$u^\wedge|_{\partial R} = u. \tag{3.1}$$

By the Riesz representation theorem, there is a Borel measure ω_{t_0} on $\partial\mathcal{R}$ (the *harmonic measure for the fixed point* t_0) so that

$$u^\wedge(t_0) = \int_{\partial\mathcal{R}} u(\zeta)\, \mathrm{d}\omega_{t_0}(\zeta).$$

It is known that the harmonic measure $\mathrm{d}\omega_z$ for any other point $z \in \mathcal{R}$ is mutually boundedly absolutely continuous with respect to $\mathrm{d}\omega_{t_0}$; hence there is a function $\mathcal{P}_z(\cdot)$ on $\partial\mathcal{R}$ (the *Poisson kernel* for the region $\mathcal{R}$ with the normalization that $\mathcal{P}_{t_0}(\zeta) \equiv 1$ on $\partial\mathcal{R}$) so that (3.1) becomes

$$u^\wedge(z) = \int_{\partial\mathcal{R}} u(\zeta)\mathcal{P}_z(\zeta)\, \mathrm{d}\omega_{t_0}(\zeta) \tag{3.2}$$

This formula can be extended to measures in a natural way: given a Borel measure μ on $\partial\mathcal{R}$, define a function $\mu^\wedge$ on $\mathcal{R}$ by

$$\mu^\wedge(z) = \int_{\partial\mathcal{R}} \mathcal{P}_z(\zeta)\, \mathrm{d}\mu(\zeta). \tag{3.3}$$

Note that if a continuous function u on $\partial\mathcal{R}$ is identified with the measure $\mathrm{d}\mu_u(\zeta)$ $= u(\zeta)\mathrm{d}\omega_{t_0}(\zeta)$, then formula (3.3) agrees with (3.2). Moreover, there is a converse for the case of positive harmonic functions: *any* positive harmonic function on $\mathcal{R}$ is of the form $u^\wedge$ as in (3.3) for a uniquely determined positive Borel measure μ on $\mathcal{R}$.

One of the difficulties with function theory on multiply-connected domains (in contrast with function theory on the unit disk) is that harmonic functions need not have single-valued harmonic conjugates; consequently, a given harmonic function $u^\wedge$ on $\mathcal{R}$ can fail to be the real part of any (single-valued) holomorphic function on $\mathcal{R}$. A natural task then is: *given a harmonic function h on $\mathcal{R}$ having the form $u^\wedge$ as in* (3.2) *or more generally $\mu^\wedge$ as in* (3.3), *describe in terms of the function $u \in C_{\mathbb{R}}(\partial\mathcal{R})$ (or in terms of the measure μ on $\partial\mathcal{R}$) when it is the case that h has a single-valued harmonic conjugate*. Note that the second case covers the first case by putting $\mathrm{d}\mu = u\,\mathrm{d}\omega_{t_0}$ so it suffices to solely consider the second case. The solution is quite elegant (see [4], [13], [29]) and can be described as follows. One can show that there exists a set $\boldsymbol{\phi} = \{\phi_1, \dots, \phi_m\}$ of m continuous real-valued functions on $\partial\mathcal{R}$ such that

$$\boldsymbol{\phi} = \text{ real basis for } L^2(\omega_{t_0}) \ominus [H^2(\omega_{t_0}) + \overline{H^2(\omega_{t_0})}], \text{ or equivalently}$$

$$\{\phi_1 \mathrm{d}\omega_{t_0}, \dots, \phi_m \mathrm{d}\omega_{t_0}\} = \text{ real basis for } (A(\mathcal{R}) + \overline{A(\mathcal{R})})^{\perp}. \tag{3.4}$$

Here $H^2(\omega_{t_0})$ is the Hardy space of analytic functions over $\mathcal{R}$ based on the measure ω_{t_0} on $\partial\mathcal{R}$ (see e.g. [18]), the overline denotes complex conjugation, $A(R)$ is the algebra of continuous functions on $\mathcal{R}^-$ which are holomorphic on $\mathcal{R}$, and the notation $\perp$ denotes the annihilator computed in the space of Borel measures $M(\partial\mathcal{R})$ on $\partial\mathcal{R}$ dual to the Banach space $C(\partial\mathcal{R})$ of continuous functions on $\partial\mathcal{R}$. Then the result is: *$\mu^\wedge$ given by* (3.3) *has a single-valued harmonic conjugate $\widetilde{\mu}^\wedge$ if and only if the orthogonality conditions*

$$\int_{\partial\mathcal{R}} \phi_i(\zeta)\,\mathrm{d}\mu(\zeta) = 0 \quad \textit{for } i = 1, \dots, m \tag{3.5}$$

are satisfied. Moreover, the condition that $u^\wedge(t_0) = 1$ corresponds to the condition that $\mu(\partial\mathcal{R}) = 1$ (so μ is a probability measure on $\partial\mathcal{R}$). *Throughout this section, the notation $\boldsymbol{\phi} = \{\phi_1, \dots, \phi_m\}$ refers to a fixed r-tuple of real-valued continuous functions on $\partial\mathcal{R}$ constructed as in* (3.4).

All these observations lead to a parametrization of scalar-valued normalized Herglotz class $\mathcal{H}(\mathcal{R})_1$ as follows. Given a measure $\mu \in \mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$, define a positive harmonic function $\mu^\wedge$ on $\partial\mathcal{R}$ by (3.3). Since $\mu \in \mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$, μ satisfies the orthogonality conditions (3.5) and hence any harmonic conjugate of $\mu^\wedge$ is single-valued. Then there is a unique such harmonic conjugate $\widetilde{\mu}^\wedge$ so that $\widetilde{\mu}^\wedge(t_0) = 0$. If we set $f_\mu(z) = \mu^\wedge(z) + \mathrm{i}\widetilde{\mu}^\wedge(z)$, then $f_\mu \in \mathcal{H}(\mathcal{R})_1$. Furthermore, any $f \in \mathcal{H}(\mathcal{R})_1$ has the form f_μ for a uniquely determined $\mu \in \mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$. Thus *there is a one-to-one correspondence between the normalized Herglotz class $\mathcal{H}(\partial\mathcal{R})_1$ and the compact convex set of probability measures $\mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$.* As the correspondence is affine, we can also say: *the function f is extremal for the compact convex set $\mathcal{H}_1(\mathcal{R})$ if and only if $f = f_\mu$ where μ is an extremal measure for the compact convex set $\mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$.*

Thus to describe the set of extreme points for $\mathcal{H}(\mathcal{R})_1$, it suffices to describe the extreme points of $\mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$, exactly a problem analyzed in Theorem 2.3 above. However, for this function-theory context, much more definitive detailed information is available.

THEOREM 3.1 (See Theorem 1.3.17 of [4], Lemma 2.10 of [13], Lemma 3.7 of [29]). *For any* $m+1$*-tuple* $\mathbf{x} = (x_0, \dots, x_m)$ *of points on* $\partial\mathcal{R}$ *such that* $x_j \in \partial_j$ *for each* $j = 0, 1, \dots, m$*, there is a unique set of positive weights* $w_0^{\mathbf{x}}, \dots, w_m^{\mathbf{x}}$ *summing up to* 1 *such that the measure* μ *given by*

$$\mu = w_0^{\mathbf{x}}\delta_{x_0} + \cdots + w_m^{\mathbf{x}}\delta_{x_m} \tag{3.6}$$

is extremal for $\mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$*, and, conversely, any extremal measure* μ *for* $\mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$ *arises in this way.*

From the point of view of Theorem 2.3, the added content of Theorem 3.1 is as follows. For the case where $X = \partial\mathcal{R}$ and $\boldsymbol{\phi} = \{\phi_1, \dots, \phi_m\}$ is constructed as in (3.4), then *the* n*-tuple of points* $x_1, \dots, x_n$ *in* $\partial\mathcal{R}$ *is such that the zero vector* $0 \in \mathbb{R}^m$ *is an interior point of the convex hull of the set of vectors*

$$\left\{ \begin{bmatrix} \phi_1(x_1) \\ \vdots \\ \phi_m(x_1) \end{bmatrix}, \dots, \begin{bmatrix} \phi_1(x_n) \\ \vdots \\ \phi_m(x_n) \end{bmatrix} \right\} \subset \mathbb{R}^m$$

if and only if $n = m+1$ *and the* $m+1$*-tuple now indexed as* $x_0, x_1, \dots, x_m$ *consists of exactly one point from each boundary component* $\partial_j \subset \partial\mathcal{R}$ *of* $\mathcal{R}$*.*

Following [4], let us introduce the notation

$$\mathbb{T}_{\mathcal{R}} = \partial_0 \times \cdots \times \partial_m \tag{3.7}$$

for the Cartesian product of the boundary components of $\mathcal{R}$; we think of this as the "$\mathcal{R}$-torus" since it plays the same role in integral representation formulas for Herglotz functions over $\mathcal{R}$ as does the usual torus $\mathbb{T} = \partial\mathbb{D}$ for integral representation formulas for Herglotz functions over the unit disk $\mathbb{D}$. Theorem 3.1 provides a $\mathbb{T}_{\mathcal{R}}$-parametrization of the extreme points of $\mathcal{H}(\mathcal{R})_1$ as follows. For a given $\mathbf{x} = (x_0, x_1, \dots, x_m) \in \mathbb{T}_{\mathcal{R}}$, we let $\mu_{\mathbf{x}}$ be the extremal measure of $\mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$ given by (3.6). Given any $\mathbf{x} \in \mathbb{T}_{\mathcal{R}}$, we let $f_{\mathbf{x}}$ be the unique holomorphic function on $\mathcal{R}$ determined by

$$\operatorname{Re} f_{\mathbf{x}}(z) = \int_{\partial\mathcal{R}} \mathcal{P}_z(\zeta) \mathrm{d}\mu_{\mathbf{x}}(\zeta), \quad \operatorname{Im} f_{\mathbf{x}}(t_0) = 0. \tag{3.8}$$

Then *the extreme points of normalized Herglotz functions* $\mathcal{H}(\mathcal{R})_1$ *consist exactly of the functions* $f_{\mathbf{x}}$ *with* $\mathbf{x} \in \mathbb{T}_{\mathcal{R}}$*.*

We next observe that the convex set $\mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$ is compact and convex in the space of real Borel measures $M(\partial\mathcal{R})$ over $\mathcal{R}$ where the latter space carries the weak* topology. As $M(\partial\mathcal{R})$ is the dual of $C_{\mathbb{R}}(\partial R)$ which is a separable Banach space, it follows that the unit ball in $M(\partial\mathcal{R})$ is metrizable. Hence the second statement in Theorem 1.2 applies. Furthermore, when we use the correspondence between $\partial_{\mathrm{e}}\mathcal{C}(\partial R, 1, \phi)$ and $\mathbb{T}_{\mathcal{R}}$ to transport the weak* topology on $\partial_{\mathrm{e}}\mathcal{C}(\partial R, 1, \boldsymbol{\phi})$ to a topology on $\mathbb{T}_{\mathcal{R}}$, one can check that the topology so obtained is just the Euclidean topology on $\mathbb{T}_{\mathcal{R}}$ inherited as a subset of $\mathbb{C}^{m+1}$. Thus, by Theorem 1.2

above, any measure $\mu \in \mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$ has an integral representation

$$\mu = \int_{\mathbb{T}_{\mathcal{R}}} \mu_{\mathbf{x}} \mathrm{d}\nu(\mathbf{x})$$

where the integral is defined in the weak sense:

$$\int_{\partial\mathcal{R}} \phi(\zeta)\mathrm{d}\mu(\zeta) = \int_{\mathbb{T}_{\mathcal{R}}} \Big[\int_{\partial\mathcal{R}} \phi(\zeta)\mathrm{d}\mu_{\mathbf{x}}(\zeta) \Big] \mathrm{d}\nu(\mathbf{x}) \quad \text{for each } \phi \in C_{\mathbb{R}}(\partial\mathcal{R}). \tag{3.9}$$

This leads to the following integral representation formula for functions f in the normalized Herglotz class $\mathcal{H}^1_1(\mathcal{R})$.

THEOREM 3.2 (See Theorem 1.3.26 of [4]). *Given $f \in \mathcal{H}(\mathcal{R})_1$, there is a probability measure ν on $\mathbb{T}_{\mathcal{R}}$ so that*

$$f(z) = \int_{\mathbb{T}_{\mathcal{R}}} f_{\mathbf{x}}(z)\, \mathrm{d}\nu(\mathbf{x}). \tag{3.10}$$

Proof. We have already seen that any $f \in \mathcal{H}(\mathcal{R})_1$ is associated with a uniquely determined measure $\mu \in \mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$ so that $\mathrm{Re}\, f = \mu^{\wedge}$ as in (3.3). Plugging this μ into (3.9) (and setting $f(z) = \mathcal{P}_z(\zeta)$) tells us that there is a probability measure ν on $\mathbb{T}_{\mathcal{R}}$ so that

$$\mathrm{Re}\, f(z) = \int_{\partial\mathcal{R}} \mathcal{P}_z(\zeta)\mathrm{d}\mu(\zeta) = \int_{\mathbb{T}_{\mathcal{R}}} \Big[\int_{\partial\mathcal{R}} \mathcal{P}_z(\zeta)\mathrm{d}\mu_{\mathbf{x}}(\zeta) \Big]\, \mathrm{d}\nu(\mathbf{x}) = \int_{\mathbb{T}_{\mathcal{R}}} \mathrm{Re}\, f_{\mathbf{x}}(z)\, \mathrm{d}\nu(\mathbf{x}).$$

By the uniqueness of the harmonic conjugate normalized to have value 0 at t_0, the formula (3.10) now follows. ∎

3.2. THE MATRIX-VALUED HERGLOTZ CLASS OF $\mathcal{R}$. We now wish to obtain results parallel to Theorem 3.1 and Theorem 3.2 for the normalized matrix-valued Herglotz class $\mathcal{H}^N(\mathcal{R})_I$. For the matrix-valued case the function theory is not as highly developed. The implication is that the results which we do obtain are not as explicit as for the scalar-valued case.

By applying Theorem 2.8 to the convex set $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ (where the collection $\boldsymbol{\phi}$ as usual is equal to $\{\phi_1, \dots, \phi_m\}$ as in (3.4)), we get most of the following somewhat less explicit analogue of Theorem 3.1.

THEOREM 3.3. *The $N \times N$ matrix-valued Borel measure μ on $\partial\mathcal{R}$ is extremal for the compact convex set $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ (with $\boldsymbol{\phi}$ as in* (3.4)*) if and only if μ has a representation of the form*

$$\mu = \sum_{j=1}^{n} W_j \delta_{x_j}, \tag{3.11}$$

where:

(i) *n is a natural number such that $1 \leqslant n \leqslant (m+1)N^2$;*

(ii) *$x_1, \dots, x_n$ are n distinct points in $\partial\mathcal{R}$;*

(iii) $W_1, \dots, W_n$ *are* n $N \times N$ *matrices such that* $W_j \geqslant 0$ *for each* j, $\sum_{j=1}^{n} W_j = I_N$, *and* $\{\operatorname{Ran} W_j : j = 1, \dots, n\}$ *is* $\boldsymbol{\phi}(\mathbf{x})$*-constrained weakly independent (see Definition 2.7), where here we set*

$$\boldsymbol{\phi}(\mathbf{x}) = \left\{ \begin{bmatrix} \phi_1(x_1) \\ \vdots \\ \phi_m(x_1) \end{bmatrix}, \dots, \begin{bmatrix} \phi_1(x_n) \\ \vdots \\ \phi_m(x_n) \end{bmatrix} \right\} \subset \mathbb{R}^m.$$

Furthermore, if μ *of the form* (3.11) *is extremal for* $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$, *then the additional property*

(iv) $\sum_{j: x_j \in \partial_r} W_j$ *is invertible for each* $r = 0, 1, \dots, m$

holds; consequently the number of points n *in the support* $\operatorname{supp}\mu$ *of* μ *in fact satisfies* $m+1 \leqslant n$ *with at least one point* x_j *from the support of* μ *in each connected component* ∂_r *of* $\partial\mathcal{R}$.

Proof. What is added here going beyond the structure given by Theorem 2.8 for the general (non function-theoretic) case is the additional property satisfied by any extremal measure given by condition (iv). To see this, we note that for any unit vector $\mathbf{u} \in \mathbb{C}^N$, the scalar-valued measure $\mu_{\mathbf{u}}(\Delta) := \mathbf{u}^* \mu(\Delta) \mathbf{u}$ is in the convex set of scalar measures $\mathcal{C}(\partial\mathcal{R}, 1, \phi)$. We now quote the result of Lemma 1.3.1 in [4]: *if* $\mu^{(1)}$ *is a nonzero scalar positive measure on* $\partial\mathcal{R}$ *such that* $\int_{\partial\mathcal{R}} \phi_i \mathrm{d}\mu^{(1)} = 0$ *for* $i = 1, \dots, m$ (so $\mu^{(1)\wedge}$ has a single-valued harmonic conjugate), *then* $\mu^{(1)}(\partial_r) > 0$ *for each* $r = 0, 1, \dots, m$. Consequently, if $\sum_{j: x_j \in \partial_r} W_j = \mu(\partial_r)$ is singular for some r, then there is a unit vector $\mathbf{u}$ so that $\mu_{\mathbf{u}}(\partial_r) = \mathbf{u}^* \mu(\partial_r) \mathbf{u} = 0$. From Lemma 1.3.1 of [4] we conclude that $\mathbf{u}^* \mu \mathbf{u}$ is the zero measure, in contradiction with $\mu(\partial\mathcal{R}) = I_N$. ∎

REMARK 3.4. When Theorem 3.3 is specialized to the scalar case ($N = 1$), the bounds $m+1 \leqslant n \leqslant (m+1)N^2 = m+1$ uniquely determine the value of n as $n = m+1$ for the total number of points of $\operatorname{supp}\mu$ for $\mu \in \partial_{\mathrm{e}}\mathcal{C}(\partial\mathcal{R}, 1, \boldsymbol{\phi})$; as each boundary component must contain at least one point of support of μ, we recover the result in Theorem 3.1 that any such μ contains exactly one point of support from each boundary component ∂_r ($r = 0, 1, \dots, m$). For the case $N > 1$, Theorem 3.3 (as compared to Theorem 3.1) lacks an explicit characterization as to which natural numbers n between $m+1$ and $(m+1)N^2$ and which associated n-tuples of points $\mathbf{x} = (x_1, \dots, x_n)$ actually arise in a representation (3.11) for an extreme point of $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$, beyond the information that $\mathbf{x}$ must include at least one point from each boundary component $\partial_1, \dots, \partial_m$. In particular, the only apparent upper bound on the number of points of $\operatorname{supp}\mu$ for an extremal $\mu \in \mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ in a particular boundary component ∂_r is the crude bound $(m+1)N^2 - m$. Also it is not clear to what extent the n-tuple of points $\mathbf{x}$ determines the associated n-tuple of matrix weights $W_1, \dots, W_n$; note that the classes of examples from Theorems 2.12 and 2.14 show that it is certainly not the case that the support $\mathbf{x} = \{x_1, \dots, x_n\}$ uniquely determines the associated set of matrix weights $\mathbf{w} = \{W_1, \dots, W_n\}$. However some additional detail is available for the case of an annulus ($m = 1$); see Remark 5.3 below.

Despite the lack of explicitness in the characterization of the extreme points of $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ as explained in Remark 3.4, we can still pursue the matrix analogue of much of the analysis done for the scalar-valued normalized Herglotz class as follows.

It is straightforward to see that positive $N \times N$ matrix-valued harmonic functions H are given via a Poisson representation

$$H(z) = \mu^\wedge(z) := \int_{\partial\mathcal{R}} \mathcal{P}_z(\zeta) \mathrm{d}\mu(\zeta)$$

where now μ is a complex Hermitian positive $N \times N$ matrix-valued measure on $\partial\mathcal{R}$. Moreover, the harmonic matrix-valued function $\mu^\wedge(z)$ has a single-valued matrix-valued harmonic conjugate if and only if the matrix measure μ satisfies the orthogonality conditions (3.5) (with respect to the *scalar-valued* functions $\boldsymbol{\phi} := \{\phi_1, \dots, \phi_m\}$), and the normalization condition that $\mu^\wedge(t_0) = I_N$ translates to the condition on μ that $\mu(X) = I_N$. By continuing an analysis parallel to what was done above for the scalar-valued case, we arrive at the following: *Given $\mu \in \mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$, there is a unique $F_\mu \in \mathcal{H}^N(\mathcal{R})_I$ so that*

$$\text{(3.12)} \qquad \operatorname{Re} F_\mu(z) = \int_{\partial\mathcal{R}} \mathcal{P}_z(\zeta) \mathrm{d}\mu(\zeta).$$

Conversely, any $F \in \mathcal{H}^N(\mathcal{R})_I$ arises in this way from a $\mu \in \mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$. Moreover, the function F_μ is extremal in $\mathcal{H}^N(\mathcal{R})_I$ if and only if μ is extremal in $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$.

The set of complex Hermitian $N \times N$ matrix-valued measures $[M(X)^{N\times N}]_{\mathrm{h}}$ is the dual space of the separable real Banach space $[C(\partial\mathcal{R})^{N\times N}]_{\mathrm{h}}$ of complex Hermitian $N \times N$ matrix-valued continuous functions on $\partial\mathcal{R}$, and hence the unit ball is metrizable and the second statement in Theorem 1.2 applies. The set of extreme points $\partial_{\mathrm{e}}\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ is a Borel subset of $[M(X)^{N\times N}]_{\mathrm{h}}$ and we have a (admittedly somewhat implicit) parametrization from Theorem 3.3. In detail, let us denote by $\mathbb{T}^N_{\mathcal{R}}$ (the matrix $\mathcal{R}$-torus) the set

$$\mathbb{T}^N_{\mathcal{R}} = \{(\mathbf{x}, \mathbf{w}) : \mathbf{x} = (x_1, \dots, x_n) \subset \partial\mathcal{R} \text{ and } \mathbf{w} = (W_1, \dots, W_n) \subset [\mathbb{C}^{N\times N}]_+ \\ \text{are as in Theorem 3.3}\}.$$

Given $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}^N_{\mathcal{R}}$, there is an associated extremal measure $\mu_{\mathbf{x},\mathbf{w}} \in \partial_{\mathrm{e}}\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ given by (3.11) in Theorem 3.3, and all extremal measures $\mu \in \partial_{\mathrm{e}}\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ are of the form $\mu_{\mathbf{x},\mathbf{w}}$ for some $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}^N_{\mathcal{R}}$. We topologize $\mathbb{T}^N_{\mathcal{R}}$ by transporting the weak* topology on the the associated set of measures $\mu_{\mathbf{x},\mathbf{w}}$ for $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}^N_{\mathcal{R}}$. Then by Theorem 1.2, given any $\mu \in \mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$, there is a Borel probability measure

ν on $\mathbb{T}^N_{\mathcal{R}}$ so that

$$\mu = \int\limits_{\mathbb{T}^N_{\mathcal{R}}} \mu_{\mathbf{x},\mathbf{w}} \, \mathrm{d}\nu(\mathbf{x}, \mathbf{w})$$

with the integral interpreted in the weak sense:

$$\operatorname{tr}\Big(\int\limits_{\partial\mathcal{R}} \Phi(\zeta) \, \mathrm{d}\mu(\zeta) \Big) = \int\limits_{\mathbb{T}^N_{\mathcal{R}}} \operatorname{tr}\Big(\int\limits_{\partial\mathcal{R}} \Phi(\zeta) \, \mathrm{d}\mu_{\mathbf{x},\mathbf{w}}(\zeta) \Big) \, \mathrm{d}\nu(\mathbf{x}, \mathbf{w}) \tag{3.13}$$

for any $\Phi \in [C(\partial\mathcal{R})^{N\times N}]_{\mathrm{h}}$.

Since extreme points of the set $\mathcal{H}^N(\mathcal{R})_I$ correspond to extreme points of the set $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ in accordance with the formula (3.12), we see that the extreme points of the normalized Herglotz class $\mathcal{H}^N(\mathcal{R})_I$ are exactly the functions $F_{\mathbf{x},\mathbf{w}}$ determined by

$$\operatorname{Re} F_{\mathbf{x},\mathbf{w}}(z) = \int\limits_{\partial\mathcal{R}} \mathcal{P}_z(\zeta) \, \mathrm{d}\mu_{\mathbf{x},\mathbf{w}}(\zeta), \quad \operatorname{Im} F_{\mathbf{x},\mathbf{w}}(t_0) = 0. \tag{3.14}$$

We are now led to the matrix-valued analogue of Theorem 3.2.

THEOREM 3.5. *Given $F \in \mathcal{H}^N(\mathcal{R})_I$, there is a quantum probability measure ν on $\mathbb{T}^N_{\mathcal{R}}$ so that*

$$F(z) = \int\limits_{\mathbb{T}^N_{\mathcal{R}}} F_{\mathbf{x},\mathbf{w}}(z) \, \mathrm{d}\nu(\mathbf{x}, \mathbf{w}). \tag{3.15}$$

Proof. We have noted that any $F \in \mathcal{H}^N(\mathcal{R})_I$ is associated with a measure $\mu \in \mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ as in (3.12). For X an arbitrary complex Hermitian $N \times N$ matrix, we use the representation (3.13) with $\Phi(\zeta) = X\mathcal{P}_z(\zeta)$ to get

$$\begin{aligned}\operatorname{tr}(X\operatorname{Re} F(z)) &= \operatorname{tr}\Big(\int\limits_{\partial\mathcal{R}} X\mathcal{P}_z(\zeta)\mathrm{d}\mu(\zeta) \Big) = \int\limits_{\mathbb{T}^N_{\mathcal{R}}} \operatorname{tr}\Big(\int\limits_{\partial\mathcal{R}} X\mathcal{P}_z(\zeta) \, \mathrm{d}\mu_{\mathbf{x},\mathbf{w}}(\zeta) \Big)\mathrm{d}\nu(\mathbf{x}, \mathbf{w}) \\ &= \int\limits_{\mathbb{T}^N_{\mathcal{R}}} \operatorname{tr}(X\operatorname{Re} F_{\mathbf{x},\mathbf{w}}(z)) \, \mathrm{d}\nu(\mathbf{x}, \mathbf{w}) = \operatorname{tr}\Big(X \int\limits_{\mathbb{T}^N_{\mathcal{R}}} \operatorname{Re} F_{\mathbf{x},\mathbf{w}}(z) \, \mathrm{d}\nu(\mathbf{x}, \mathbf{w}) \Big).\end{aligned}$$

Since $X \in [\mathbb{C}^{N\times N}]_{\mathrm{h}}$ is arbitrary, we conclude that

$$\operatorname{Re} F(z) = \int\limits_{\mathbb{T}^N_{\mathcal{R}}} \operatorname{Re} F_{\mathbf{x},\mathbf{w}}(z) \, \mathrm{d}\nu(\mathbf{x}, \mathbf{w}).$$

By uniqueness of harmonic conjugate with value 0 at t_0, the representation (3.15) now follows. ∎

4. THE SCHUR CLASS OVER A FINITELY CONNECTED PLANAR DOMAIN

We define the (strict) $N \times N$*-matrix Schur class* over $\mathcal{R}$, denoted by $\mathcal{S}^N(\mathcal{R})$, to be the class of all holomorphic functions on $\mathcal{R}$ with values equal to $N \times N$ matrices such that $\|S(z)\| < 1$ for $z \in \mathcal{R}$. The normalized strict Schur class $\mathcal{S}^N(\mathcal{R})_0$ consists of such functions S such that in addition $S(t_0) = 0$. A consequence of the Schwarz lemma is that a holomorphic function S with (not necessarily strict) contraction values and with $S(t_0) = 0$ necessarily has strictly contractive values on all of $\mathcal{R}$.

The classes $\mathcal{H}^N(\mathcal{R})_I$ and $\mathcal{S}^N(\mathcal{R})_0$ correspond via a linear-fractional change of variable, as summarized in the following proposition. We include the elementary proof since we shall make use of the formulas in subsequent proofs.

PROPOSITION 4.1. *The normalized Schur class $\mathcal{S}^N(\mathcal{R})_0$ and the normalized Herglotz class $\mathcal{H}^N(\mathcal{R})_I$ are related according to the following linear-fractional change-of-variable formulas:*

$$S \in \mathcal{S}^N(\mathcal{R})_0 \Leftrightarrow F := (I-S)^{-1}(I+S) \in \mathcal{H}^N(\mathcal{R})_I, \tag{4.1}$$

$$F \in \mathcal{H}^N(\mathcal{R})_I \Leftrightarrow S := (F+I)^{-1}(F-I) \in \mathcal{S}^N(\mathcal{R})_0. \tag{4.2}$$

Moreover, the transformations in (4.1) *and* (4.2) *are inverse to each other.*

Proof. If $S \in \mathcal{S}^N(\mathcal{R})_0$ and F is defined as in (4.1), then clearly $F(t_0) = I_N$ and

$$\begin{aligned} F(z)+F(w)^* &= (I-S(z))^{-1}(I+S(z)) + (I+S(w)^*)(I-S(w)^*)^{-1} \\ &= (I-S(z))^{-1}[(I+S(z))(I-S(w)^*)+(I-S(z))(I+S(w)^*)](I-S(w))^{-1} \\ &= 2(I-S(z))^{-1}(I-S(z)S(w)^*)(I-S(w)^*)^{-1} \end{aligned} \tag{4.3}$$

from which we see that Re $F(z)$ is positive for $z \in \mathcal{R}$ so $F \in \mathcal{H}^N(\mathcal{R})_I$.

Similarly, if $F \in \mathcal{H}^N(\mathcal{R})_I$ and S is defined as in (4.2), then clearly $S(t_0) = 0$ and

$$\begin{aligned} & I - S(z)S(w)^* \\ &= I - (F(z)+I)^{-1}(F(z)-I)(F(w)^*-I)(F(w)^*+I)^{-1} \\ &= (F(z)+I)^{-1}[(F(z)+I)(F(w)^*+I) - (F(z)-I)(F(w)^*-I)](F(w)^*+I)^{-1}. \end{aligned}$$

We conclude that

$$I - S(z)S(w)^* = 2(F(z)+I)^{-1}[F(z)+F(w)^*](F(w)^*+I)^{-1} \tag{4.4}$$

from which we see that $S(z)$ is contractive for $z \in \mathcal{R}$ and hence $S \in \mathcal{S}^N(\mathcal{R})_0$.

Another formula which will prove useful later is

$$(F(z)+I)^{-1} = \frac{1}{2}(I - S(z)) \tag{4.5}$$

whenever F and S are related as in (4.1).

We leave to the reader the verification of the fact that the formulas (4.1) and (4.2) are inverse to each other. ∎

More generally, if S is not in the normalized Schur class but is in the strict (unnormalized) Schur class, we can apply a matrix linear-fractional map mapping the unit ball of $N \times N$ matrices to itself to obtain a new $\widetilde{S}$ which is in the normalized Schur class. Indeed, given any strictly contractive $N \times N$ matrix W, the matrix linear fractional map given by

$$L_W : Z \mapsto [AZ + B][CZ + D]^{-1}. \tag{4.6}$$

Here

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} (D_{W^*})^{-1} & -(D_{W^*})^{-1}W \\ -W^*(D_{W^*})^{-1} & (D_W)^{-1} \end{bmatrix} \tag{4.7}$$

where $D_W = (I - W^*W)^{1/2}$ and $D_{W^*} = (I - WW^*)^{1/2}$ denote the invertible defect operators of W and W^*, maps the open unit ball $\mathcal{B}\mathbb{C}^{N\times N} = \{Z \in \mathbb{C}^{N\times N} : \|Z\| < 1\}$ biholomorphically to itself and maps the given strict contraction matrix W to 0 (these constructions go back at least to the paper of Phillips [28]):

$$L_W[W] = 0.$$

One can check that the linear-fractional map $(L_W)^{-1}$ mapping 0 back to W is given by

$$(L_W)^{-1} : Z' \mapsto (A - Z'C)^{-1}(B - Z'D)$$

with A, B, C, D as in (4.7), or explicitly

$$\begin{aligned} L_W^{-1} : Z' &\mapsto ((D_{W^*})^{-1} + Z'W^*(D_{W^*})^{-1})^{-1}(Z'(D_W)^{-1} + (D_{W^*})^{-1}W) \\ &= D_{W^*}(I + Z'W^*)^{-1}(Z' + W)(D_W)^{-1} \end{aligned} \tag{4.8}$$

(where we made use of the intertwining relation $(D_{W^*})^{-1}W = W(D_W)^{-1}$). This formula will prove useful below. Notice that in the scalar case with w a point in the unit disk, the matrix linear-fractional map L_W simplifies to the familiar Möbius transformation

$$L_w : z \mapsto (z - w)(1 - z\overline{w})^{-1}$$

mapping the unit disk onto itself with the point $w \in \mathbb{D}$ mapping to 0. With these observations in hand, the following is immediate.

PROPOSITION 4.2. *If the matrix function S is in the strict Schur class $\mathcal{S}^N(\mathcal{R})$, then $\widetilde{S}$ given by*

$$\widetilde{S}(z) = L_{S(0)}[S(z)] \quad \textit{with } L_{S(0)} \textit{ given as in } (4.6) \textit{ and } (4.7) \tag{4.9}$$

is in the normalized Schur class $\mathcal{S}^N(\mathcal{R})_0$.

The following formula for the defect of S in terms of the defect of $\widetilde{S}$ will also be useful below.

PROPOSITION 4.3. *Suppose* $S \in \mathcal{S}^N(\mathcal{R})$ *and* $\widetilde{S} \in \mathcal{S}^N(\mathcal{R})_0$ *are related as in Proposition* 4.2. *Then we have*

$$\begin{aligned} &I - S(z)S(w)^* \\ &= D_{S(0)^*}(I+\widetilde{S}(z)S(0)^*)^{-1}(I-\widetilde{S}(z)\widetilde{S}(w)^*)(I+S(0)\widetilde{S}(w)^*)^{-1}D_{S(0)^*}. \end{aligned} \tag{4.10}$$

Proof. From the representation (4.9) for $\widetilde{S}$ in terms of S, we solve for S to get

$$S(z) = (L_{S(0)})^{-1}[\widetilde{S}(z)].$$

We now use the explicit formula for $(L_{S(0)})^{-1}$ determined from equation (4.8) to get

$$S(z) = D_{S(0)^*}(I+\widetilde{S}(z)S(0)^*)^{-1}(\widetilde{S}(z)+S(0))(D_{S(0)})^{-1}.$$

Hence we get

$$\begin{aligned} I - S(z)S(w)^* &= I - (L_{S(0)^*})^{-1}[\widetilde{S}(z)]((L_{S(0)})^{-1}[\widetilde{S}(w)])^* \\ &= I - D_{S(0)^*}(I+\widetilde{S}(z)S(0)^*)^{-1}(\widetilde{S}(z)+S(0))(D_{S(0)})^{-1} \\ &\qquad \cdot (D_{S(0)})^{-1}(\widetilde{S}(w)^*+S(0)^*)(I+S(0)\widetilde{S}(w)^*)^{-1}D_{S(0)^*} \\ &= D_{S(0)^*}(I+\widetilde{S}(z)S(0)^*)^{-1}X(I+S(0)\widetilde{S}(w)^*)^{-1}D_{S(0)^*} \end{aligned} \tag{4.11}$$

where we set

$$\begin{aligned} X =& [I+\widetilde{S}(z)S(0)^*](D_{S(0)^*})^{-2}[I+S(0)\widetilde{S}(w)^*] \\ &- [\widetilde{S}(z)+S(0)](D_{S(0)})^{-2}[\widetilde{S}(w)^*+S(0)^*]. \end{aligned} \tag{4.12}$$

In the computation to follow we use the intertwining relations

$$S(0)^*(D_{S(0)^*})^{-2} = (D_{S(0)})^{-2}S(0)^*, \quad (D_{S(0)^*})^{-2}S(0) = S(0)(D_{S(0)})^{-2}. \tag{4.13}$$

We now pick up the computation of X in (4.12):

$$\begin{aligned} X &= (D_{S(0)^*})^{-2} + \widetilde{S}(z)S(0)^*(D_{S(0)^*})^{-2} \\ &\quad + (D_{S(0)^*})^{-2}S(0)\widetilde{S}(w)^* + \widetilde{S}(z)S(0)^*(D_{S(0)^*})^{-2}S(0)\widetilde{S}(w)^* \\ &\quad - \widetilde{S}(z)(D_{S(0)})^{-2}\widetilde{S}(w)^* - S(0)(D_{S(0)})^{-2}\widetilde{S}(w)^* \\ &\quad - \widetilde{S}(z)(D_{S(0)})^{-2}S(0)^* - S(0)(D_{S(0)})^{-2}S(0)^* \\ &= [(D_{S(0)^*})^{-2} + \widetilde{S}(z)S(0)^*(D_{S(0)^*})^{-2}] \cdot [I+S(0)\widetilde{S}(w)^*] \\ &\quad - [S(0)(D_{S(0)})^{-2} + \widetilde{S}(z)(D_{S(0)})^{-2}] \cdot [S(0)^* + \widetilde{S}(w)^*] \\ &= (D_{S(0)^*})^{-2} - S(0)(D_{S(0)})^{-2}S(0)^* \\ &\quad + \widetilde{S}(z)S(0)^*(D_{S(0)^*})^{-2}S(0)\widetilde{S}(w)^* - \widetilde{S}(z)(D_{S(0)})^{-2}\widetilde{S}(w)^* + [\text{ cross terms }] \end{aligned}$$

where we make use of (4.13) to see that the cross terms vanish. Continuation of the computation of X and again making use of (4.13) then gives:

$$X = (D_{S(0)^*})^{-2}[I-S(0)S(0)^*] + \widetilde{S}(z)[S(0)^*S(0)-I](D_{S(0)})^{-2}\widetilde{S}(w)^* = I - \widetilde{S}(z)\widetilde{S}(w)^*.$$

Plugging X back into (4.11) gives us (4.10) as wanted. ∎

With these preliminaries out of the way, we may use the integral representation formula (3.15) for a normalized Herglotz function to arrive at the following representation for the defect kernel $I - S(z)S(w)^*$ for a normalized Schur-class function S. To this end, we associate with any point $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}^N_{\mathcal{R}}$ the normalized Schur-class function

$$S_{\mathbf{x},\mathbf{w}}(z) = (F_{\mathbf{x},\mathbf{w}}(z) + I)^{-1}(F_{\mathbf{x},\mathbf{w}}(z) - I) \tag{4.14}$$

where $F_{\mathbf{x},\mathbf{w}} \in \mathcal{H}^N(\mathcal{R})_I$ is given by (3.14).

THEOREM 4.4. *Given S in the strict Schur class $\mathcal{S}^N(\mathcal{R})$, there is a $\mathbb{C}^{N\times N}$-valued function $((\mathbf{x}, \mathbf{w}), z) \mapsto H_{\mathbf{x},\mathbf{w}}(z)$ on $\mathbb{T}^N_{\mathcal{R}} \times \mathcal{R}$, with values bounded and measurable in $(\mathbf{x}, \mathbf{w})$ for each fixed z and holomorphic in z for each fixed $(\mathbf{x}, \mathbf{w})$, along with a probability measure ν on $\mathbb{T}^N_{\mathcal{R}}$ so that*

$$I - S(z)S(w)^* = \int_{\mathbb{T}^N_{\mathcal{R}}} H_{\mathbf{x},\mathbf{w}}(z)(I - S_{\mathbf{x},\mathbf{w}}(z)S_{\mathbf{x},\mathbf{w}}(w)^*)H_{\mathbf{x},\mathbf{w}}(w)^*\, \mathrm{d}\nu(\mathbf{x}, \mathbf{w}). \tag{4.15}$$

Proof. We first consider the case where S is in the normalized Schur class $\mathcal{S}^N(\mathcal{R})_0$. Then $F := (I - S)^{-1}(I + S)$ is in the normalized Herglotz class $\mathcal{H}^N(\mathcal{R})_I$ as explained in Proposition 4.1. By Theorem 3.5 there is a probability measure ν on $\mathbb{T}^N_{\mathcal{R}}$ so that

$$F(z) = \int_{\mathbb{T}^N_{\mathcal{R}}} F_{\mathbf{x},\mathbf{w}}(z)\, \mathrm{d}\nu(\mathbf{x}, \mathbf{w}).$$

If $S_{\mathbf{x},\mathbf{w}}(z)$ is given by (4.14), then we know from Proposition 4.1 that we recover $S_{\mathbf{x},\mathbf{w}}$ from $F_{\mathbf{x},\mathbf{w}}$ according to

$$S_{\mathbf{x},\mathbf{w}}(z) = (F_{\mathbf{x},\mathbf{w}}(z) + I)^{-1}(F_{\mathbf{x},\mathbf{w}}(z) - I).$$

Then we compute

$$\begin{aligned}
&I - S(z)S(w)^* \\
&\quad = 2(F(z) + I)^{-1}(F(z) + F(w)^*)(F(w)^* + I)^{-1} \quad (\text{ by } (4.4)) \\
&\quad = \frac{1}{2}(I - S(z)) \int_{\mathbb{T}^N_{\mathcal{R}}} 2(I - S_{\mathbf{x},\mathbf{w}}(z))^{-1}(I - S_{\mathbf{x},\mathbf{w}}(z)S_{\mathbf{x},\mathbf{w}}(w)^*)(I - S_{\mathbf{x},\mathbf{w}}S_{\mathbf{x},\mathbf{w}}(w)^*) \\
&\qquad\quad \mathrm{d}\nu(\mathbf{x}, \mathbf{w})(I - S(w)^*) \quad (\text{where we make use of (4.5) and (4.3)}) \\
&\quad = \int_{\mathbb{T}^N_{\mathcal{R}}} H_{\mathbf{x},\mathbf{w}}(z)(I - S_{\mathbf{x},\mathbf{w}}(z)S_{\mathbf{x},\mathbf{w}}(w)^*)H_{\mathbf{x},\mathbf{w}}(w)^*\, \mathrm{d}\nu(\mathbf{x}, \mathbf{w})
\end{aligned}$$

where we have set

$$H_{\mathbf{x},\mathbf{w}}(z) = (I - S(z))(I - S_{\mathbf{x},\mathbf{w}}(z))^{-1}.$$

To handle the case where $S \in \mathcal{S}^N(\mathcal{R})$ is not necessarily normalized, we proceed as follows. Write $S(z) = (L_{S(0)})^{-1}[\widetilde{S}(z)]$ where $\widetilde{S}$ is in the normalized Schur class $\mathcal{S}^N(\mathcal{R})_0$. Then, by the special case of Theorem 4.4 already proved, we know that there is a probability measure ν and a function $\widetilde{H}$ so that

$$I - \widetilde{S}(z)\widetilde{S}(w)^* = \int_{\mathbb{T}^N_{\mathcal{R}}} \widetilde{H}_{\mathbf{x},\mathbf{w}}(z)(I - S_{\mathbf{x},\mathbf{w}}(z)S_{\mathbf{x},\mathbf{w}}(w)^*)\widetilde{H}_{\mathbf{x},\mathbf{w}}(w)^* \, \mathrm{d}\nu(\mathbf{x},\mathbf{w}).$$

If we now use relation (4.10), we see that (4.15) holds for S with

$$H_{\mathbf{x},\mathbf{w}}(z) = D_{S(0)^*}(I + \widetilde{S}(z)S(0)^*)^{-1}\widetilde{H}_{\mathbf{x},\mathbf{w}}(z)$$

as needed. ∎

Specializing this result to the scalar-valued case ($N = 1$) recovers the following result of Dritschel–McCullough. To state the result we introduce the scalar counterpart of the functions $S_{\mathbf{x},\mathbf{w}}$ given by (4.14): for each point $\mathbf{x}$ in the $\mathcal{R}$-torus (3.7) let $s_{\mathbf{x}}$ be the scalar Schur-class function given by

$$s_{\mathbf{x}}(z) = \frac{f_{\mathbf{x}}(z) + 1}{f_{\mathbf{x}}(z) - 1} \tag{4.16}$$

where $f_{\mathbf{x}} \in \mathcal{H}(\mathcal{R})_1$ is given by (3.8).

THEOREM 4.5 (See Proposition 2.14 of [13]). *Given a function s on $\mathcal{R}$ in the scalar-valued Schur class $\mathcal{S}(\mathcal{R})$, there are complex-valued functions $(\mathbf{x}, z) \mapsto h_{\mathbf{x}}(z)$ on $\mathbb{T}_{\mathcal{R}} \times \mathcal{R}$, bounded and measurable in $\mathbf{x}$ for each fixed z and holomorphic in z for each fixed $\mathbf{x}$, and a positive probability measure on $\mathbb{T}_{\mathcal{R}}$, such that*

$$1 - s(z)\overline{s(w)} = \int_{\mathbb{T}_{\mathcal{R}}} h_{\mathbf{x}}(z)(1 - s_{\mathbf{x}}(z)\overline{s_{\mathbf{x}}(w)})\overline{h_{\mathbf{x}}(w)} \, \mathrm{d}\nu(\mathbf{x}) \tag{4.17}$$

REMARK 4.6. We note that it is not possible to use a smaller closed subset of $\mathbb{T}^N_{\mathcal{R}}$ in the integral representation (3.15) and still have the representation hold for all $F \in \mathcal{H}^N(\mathcal{R})_I$, almost by the definition of extreme point. However some reductions are always possible in the decomposition (4.15). Note that we have already imposed the normalization that $S_{\mathbf{x},\mathbf{w}}(t_0) = 0$ for all $(\mathbf{x}, \mathbf{w})$. In addition we note that the expression $I - S_{\mathbf{x},\mathbf{w}}(z)S_{\mathbf{x},\mathbf{w}}(w)^*$ is unchanged if we replace $S_{\mathbf{x},\mathbf{w}}(z)$ by $S_{\mathbf{x},\mathbf{w}}(z)U$ with U a unitary $N \times N$ matrix. This means that we may restrict the integral in (4.15) to points $(\mathbf{x}, \mathbf{w})$ such that $S_{\mathbf{x},\mathbf{w}}(\zeta_0) = I_N$ for some point ζ_0 in $\partial\mathcal{R}$ (e.g., $\zeta_0 \in \partial_0$) and consider the integral over this smaller set $\widetilde{\mathbb{T}}^N_{\mathcal{R}}$. In special situations for the $N = 1$ case (see [14] and [15]), there are results proven that, after these reductions, there is no proper closed subset $\widetilde{\widetilde{\mathbb{T}}}_{\mathcal{R}}$ of $\widetilde{\mathbb{T}}_{\mathcal{R}}$ for which a representation of the form (4.17) can hold with $\mathbb{T}_{\mathcal{R}}$ replaced by $\widetilde{\widetilde{\mathbb{T}}}_{\mathcal{R}}$. For the case $N > 1$, our description of the set $\mathbb{T}^N_{\mathcal{R}}$ (or of $\widetilde{\mathbb{T}}_{\mathcal{R}}$) is not as explicit as in the $N = 1$ case, so as of this writing it is not at all clear how to arrive at such minimality results for the matrix-valued case.

REMARK 4.7. In [13] the authors go on to use the general theory of the generalized Schur class associated with a collection Ψ of test functions (see [12], [14]) to identify the Schur class $\mathcal{S}(\mathcal{R})$ over $\mathcal{R}$ with the Schur class $\mathcal{S}_\Psi$ associated with the collection of test functions $\Psi = \{s_{\mathbf{x}} : \mathbf{x} \in \mathbb{T}_{\mathcal{R}}\}$ and thereby also to obtain transfer-function realizations for the class $\mathcal{S}(\mathcal{R})$. These results combined with Theorem 4.4 suggest that the matrix-valued Schur class $\mathcal{S}^N(\mathcal{R})$ is connected in a similar way with the collection of matrix-valued test functions $\boldsymbol{\Psi} = \{S_{\mathbf{x},\mathbf{w}} : (\mathbf{x},\mathbf{w}) \in \mathbb{T}^N_{\mathcal{R}}\}$. This is indeed the case (see [8], [21]). Similar results can be worked out for the matrix-valued Schur class associated with the algebra H^∞_B studied by Raghupathi [30] (see Remark 2.6 above).

5. THE SPECTRAL SET PROBLEM

Let $\mathcal{R}$ denote a domain in the complex plane $\mathbb{C}$ with boundary $\partial\mathcal{R}$ with closure $\mathcal{R}^-$. An operator T on a complex Hilbert space $\mathcal{H}$ is said to have $\mathcal{R}^-$ as a *spectral set* if the spectrum $\sigma(T)$ of T is contained in $\mathcal{R}^-$ and

$$\|f(T)\| \leqslant \|f\|_{\mathcal{R}} = \sup\{|f(z)| : z \in \mathcal{R}\}$$

for every rational function f with poles off of $\mathcal{R}^-$, where $f(T)$ can be defined by the Riesz functional calculus or simply as $f(T) = p(T)q(T)^{-1}$ when f is written as the ratio of polynomials $f(z) = \frac{p(z)}{q(z)}$. The operator T is said to have a $\partial\mathcal{R}$-normal dilation if there exists a Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ as a subspace so that

$$f(T) = P_{\mathcal{H}} f(N)|_{\mathcal{H}}$$

for every rational function f with poles off of $\mathcal{R}^-$ (where $P_{\mathcal{H}}$ is the orthogonal projection of $\mathcal{K}$ onto $\mathcal{H}$). It is easily seen that if T has a $\partial\mathcal{R}$-normal dilation, then $\mathcal{R}^-$ is a spectral set for T. The converse question can be formulated as:

Given that $\mathcal{R}^-$ is a spectral set for T, does it follow that T has a $\partial\mathcal{R}$-normal dilation?

This has become known as the *spectral set question* for $\mathcal{R}$ (see [6]).

For the case of the unit disk $\mathcal{R} = \mathbb{D}$, the von Neumann inequality combined with the Sz.-Nagy dilation theorem implies a positive answer to the spectral set question. For the case where $\mathcal{R} = \mathbb{A}$ is an annulus, it is a result of Agler [2] (see also [24]) that the spectral set question again has a positive answer. However, for the case of a multiply-connected domain with at least two holes, more recent work of Agler–Harland–Raphael [4] and Dritschel–McCullough [13] give two complementary approaches to showing that the spectral set question has a negative solution. In this section we discuss briefly how the ideas of this paper relate to the spectral set question.

In [6] Arveson obtained a reformulation of the spectral set question which had profound influence on subsequent work. For our purposes it is convenient to assume that $\sigma(T)$ is contained in the open domain $\mathcal{R}$ rather than in $\mathcal{R}^-$; in this case we can use the standard Riesz holomorphic functional calculus to define $s(T) \in \mathcal{L}(\mathcal{H})$ for any holomorphic function on $\mathcal{R}$, in particular, for s in the Schur class $\mathcal{S}(\mathcal{R})$. Then the condition that T has $\mathcal{R}^-$ as a spectral set can be reformulated as: *for any* $s \in \mathcal{S}(\mathcal{R})$, $\|s(T)\| \leqslant 1$. By the Arveson–Stinespring dilation theory (see [5]), the condition that T have a $\partial\mathcal{R}$-normal dilation can be reformulated as: *for any* $S \in \mathcal{S}^N(\mathcal{R})$ $(N = 1, 2, \dots)$, $\|S(T)\| \leqslant 1$. Here, for $S = [s_{ij}]_{i,j=1}^N$ in the matrix-valued Schur class $\mathcal{S}^N(\mathcal{R})$, we define $S(T)$ by

$$S(T) = [s_{ij}(T)] \in \mathcal{L}(\mathcal{H}^N).$$

Then the Arveson reformulation of the spectral set question for $\mathcal{R}$ becomes:

Given $T \in \mathcal{L}(\mathcal{H})$ *such that* $\|s(T)\| \leqslant 1$ *for all* $s \in \mathcal{S}(\mathcal{R})$, *does it follow that* $\|S(T)\| \leqslant 1$ *for all* $S \in \mathcal{S}^N(\mathcal{R})$*?*

From the result (4.10) of Proposition 4.3, we see that in fact one may restrict to the normalized Schur classes $\mathcal{S}(\mathcal{R})_0$ and $\mathcal{S}^N(\mathcal{R})_0$ in the above condition. We may then use relations (4.3) and (4.4) to get the following reformulation:

Given $T \in \mathcal{L}(\mathcal{H})$ *such that* $\operatorname{Re} f(T) \geqslant 0$ *for all* $f \in \mathcal{H}(\mathcal{R})_1$, *does it follow that* $\operatorname{Re} F(T) \geqslant 0$ *for all* $F \in \mathcal{H}^N(\mathcal{R})_I$ *for any* $N = 1, 2, \dots$ *?*

By plugging T into the decomposition (4.17) with the Riesz holomorphic functional calculus, we see that to check whether $\mathcal{R}^-$ is a spectral set for T, it suffices to check the condition $\|s(T)\| \leqslant 1$ only for $s = s_{\mathbf{x}}$ for each $\mathbf{x} \in \mathbb{T}_{\mathcal{R}}$. By using the relations (4.3) and (4.4), an equivalent condition for $\mathcal{R}^-$ to be a spectral set for T is that $\operatorname{Re} f_{\mathbf{x}}(T) \geqslant 0$ for each $\mathbf{x} \in \mathbb{T}_{\mathcal{R}}$ (see Theorem 1.6.16 of [4]). Similarly, to check that T has a $\partial\mathcal{R}$-normal dilation, it suffices to check the condition $\|S(T)\| \leqslant 1$ only for $S \in \mathcal{S}^N(\mathcal{R})_0$ of the form $S = S_{\mathbf{x},\mathbf{w}}$ for $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}_{\mathcal{R}}^N$. By using the relations (4.3) and (4.10), it is equivalent to check that $\operatorname{Re} F_{\mathbf{x},\mathbf{w}}(T) \geqslant 0$ for each $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}_{\mathcal{R}}^N$. We thus come to the following equivalent reformulations of the spectral set question:

Given $T \in \mathcal{L}(\mathcal{H})$ *for which* $\|s_{\mathbf{x}}(T)\| \leqslant 1$ *(respectively,* $\operatorname{Re} f_{\mathbf{x}}(T) \geqslant 0$*) for all* $\mathbf{x} \in \mathbb{T}_{\mathcal{R}}$, *does it follow that* $\|S_{\mathbf{x},\mathbf{w}}(T)\| \leqslant 1$ *(respectively,* $\operatorname{Re} F_{\mathbf{x},\mathbf{w}}(T) \geqslant 0$*) for all* $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}_{\mathcal{R}}^N$*?*

Let us suppose that all extremal measures μ for the set $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ are special (see Remark 2.16). Thus any $\mu_{\mathbf{x},\mathbf{w}} \in \partial_{\mathrm{e}}\mathcal{C}(\mathcal{R}, N, \phi)$ has the form $\mu_{\mathbf{x},\mathbf{w}} = \sum\limits_{k=1}^{n} \mu_{\mathbf{x}_k} W_k$ for some points $\mathbf{x}_k \in \mathbb{T}_{\mathcal{R}}$ and for matrix weights W_k such that $\{\operatorname{Ran} W_k : 1 \leqslant k \leqslant n\}$ is a weakly independent family of subspaces. We then see that

$$F_{\mathbf{x},\mathbf{w}}(z) + F_{\mathbf{x},\mathbf{w}}(w)^* = \sum_{k=1}^{n} (f_{\mathbf{x}_k}(z) + \overline{f_{\mathbf{x}_k}(w)}) W_k = \sum_{k=1}^{n} W_k^{1/2} [(f_{\mathbf{x}_k}(z) + \overline{f_{\mathbf{x}_k}(w)}) I_N] W_k^{1/2}.$$

The assumption that $\operatorname{Re} f_{\mathbf{x}}(T) \geqslant 0$ for each $\mathbf{x} \in \mathbb{T}_{\mathcal{R}}$ then leads to the conclusion that $\operatorname{Re} F_{\mathbf{x},\mathbf{w}}(T) \geqslant 0$ for each $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}_{\mathcal{R}}^N$, and hence that the spectral set question has an affirmative answer for $\mathcal{R}$, in contradiction with the results of [4], [13]. These observations lead to the following corollary concerning the structure of the set of extreme points for a compact convex set of the form $\mathcal{C}(\partial\mathcal{R}, N, \boldsymbol{\phi})$ with $\boldsymbol{\phi}$ as in (3.4).

COROLLARY 5.1. *There are multiply-connected domains $\mathcal{R}$ (with at least two holes) such that not all extremal measures μ for $\mathcal{C}(\mathcal{R}, N, \boldsymbol{\phi})$ are special (as defined in Remark* 2.16*).*

REMARK 5.2. We note that for the case where $\mathcal{R}$ is the unit disk $\mathbb{D}$, it is the case that all extremal measures are special; hence the argument leading to Corollary 5.1 yields yet another proof that the spectral set question for $\mathcal{R} = \mathbb{D}$ has an affirmative answer. Alternatively, one could simply use the Herglotz integral representation formula for the matrix-valued Herglotz class $\mathcal{H}^N(\mathbb{D})_I$ and work with quantum probability measures, i.e., positive operator measures on $\mathbb{T}$ with $\nu(\mathbb{T}) = I_N$, rather than probability measures:

$$F(z) = \int_{\mathbb{T}} \frac{\zeta + z}{\zeta - z}\, \mathrm{d}\nu(\zeta). \tag{5.1}$$

The representation (3.10) for the scalar-valued Herglotz class over $\mathcal{R}$ suggests a representation for $F \in \mathcal{H}^N(\mathcal{R})_I$ analogous to (5.1) for the disk case $\mathcal{R} = \mathbb{D}$:

$$F(z) = \int_{\mathbb{T}_{\mathcal{R}}} f_{\mathbf{x}}(z) \mathrm{d}\nu(\mathbf{x}) \tag{5.2}$$

with ν a quantum probability measure. However an argument similar to that leading to Corollary 5.1 shows that such a representation cannot possibly be true in general when $\mathcal{R}$ has at least two holes. Indeed, if F has a representation as in (5.2), it is convenient first to rewrite it in the form

$$\int_{\mathbb{T}_{\mathcal{R}}} G(\mathbf{x})(f_{\mathbf{x}}(z) I_N) G(\mathbf{x})^*\, \mathrm{d}\mu(\mathbf{x}) \tag{5.3}$$

where we have set μ equal to the scalar measure $\mu(E) = \mathrm{trace}(\nu(E))$ and where we then use a matrix-valued Radon–Nikodym theorem to represent ν as $\mathrm{d}\nu(\mathbf{x}) = G(\mathbf{x})G(\mathbf{x})^* \mathrm{d}\mu(\mathbf{x})$. If we then let S be the matrix Schur-class function $S = (F + I_N)^{-1}(F - I_N)$, we obtain a so-called Agler decomposition for S as in (4.15), but using only the scalar functions $s_{\mathbf{x}}$ (4.16) rather than the full complement of extremal Schur-class functions as described in (4.14):

$$I - S(z)S(w)^* = \int_{\mathbb{T}_{\mathcal{R}}} H(z, \mathbf{x})((1 - s_{\mathbf{x}}(z)\overline{s_{\mathbf{x}}(w)}) I_N) H(w, \mathbf{x})^*\, \mathrm{d}\mu(\mathbf{x}) \tag{5.4}$$

where, explicitly for the record,

$$H(z, \mathbf{x}) = \sqrt{2}(F(z) + I_N)^{-1} G(\mathbf{x}) \frac{1}{1 - s_{\mathbf{x}}(z)}.$$

From the representation (5.4) we see that $\|S(T)\| \leqslant 1$ whenever $\|s_{\mathbf{x}}(T)\| \leqslant 1$ for each $\mathbf{x} \in \mathbb{T}_{\mathcal{R}}$ and we conclude that every $F \in \mathcal{H}^N(\mathcal{R})_I$ having a representation as in (5.2) leads to a positive solution of the spectral set question in contradiction with the results of [2], [13]. One would also arrive at a representation of the form

(5.4) if one assumed that the associated Herglotz function F had an integral representation (3.15) with the measure ν supported only on points of $\mathbb{T}^N_{\mathcal{R}}$ associated with special measures.

The particular type of Agler decomposition appearing in (5.4) comes up in the work of Dritschel–McCullough [13] and is associated with matrix inner functions S on $\mathcal{R}$ which are *diagonalizable* (i.e., $S(z) = U^* D(z) V$ with U and V constant unitary matrices and $D(z)$ pointwise diagonal). It is exactly the existence of matrix inner functions on $\mathcal{R}$ which are not diagonalizable which leads to a counterexample for the spectral set question on $\mathcal{R}$ in [13].

REMARK 5.3. We have not determined if all extremal measures are special for the case of an annulus $\mathcal{R} = \mathbb{A}$. However there is a somewhat different way to reduce the matrix-valued Herglotz (or Schur) class to the scalar Herglotz class for the case $\mathcal{R} = \mathbb{A}$ due to McCullough [24] which we now describe.

For $\mathbb{A}$ equal to the annulus $\mathbb{A}_q = \{z \in \mathbb{C} : q < |z| < 1\}$ (where $0 < q < 1$), it is shown in [24] that there is a curve $\{\varphi_\zeta : \zeta \in \mathbb{T}\}$ of inner functions over $\mathbb{A}_q$ (constructed from the Ahlfors function φ based at the point $\sqrt{q} \in \mathbb{A}_q$) with the following special property. First as a matter of notation, for $t = (t_1, \dots, t_N) \in \mathbb{T}^N$, let us let $\Phi_t(z)$ be the diagonal matrix inner function over $\mathbb{A}_q$ given by

$$\Phi_t(z) = \begin{bmatrix} \varphi_{t_1}(z) & & \\ & \ddots & \\ & & \varphi_{t_N}(z) \end{bmatrix} \quad \text{if } t = (t_1, \dots, t_N), \tag{5.5}$$

and, for U a unitary $N \times N$ matrix and $t \in \mathbb{T}^N$, let us set

$$R_{U,t}(z) = (I_N + U\Phi_t(z))(I - U\Phi_t(z))^{-1}.$$

Then one of the main results from [24] is: *given a point $(\mathbf{x}, \mathbf{w}) \in \mathbb{T}^N_{\mathbb{A}_q}$ with associated extremal Herglotz function $F_{\mathbf{x},\mathbf{w}}$, there is a $t \in \mathbb{T}^N$, an $N \times N$ unitary matrix U, and an invertible $N \times N$ matrix X so that*

$$\operatorname{Re} F_{\mathbf{x},\mathbf{w}}(z) = X(\operatorname{Re} R_{U,t}(z))X^* \tag{5.6}$$

for all $z \in \mathbb{A}_q$. Note that an easy computation gives

$$R_{U,t}(z) + R_{U,t}(w)^* = 2(I - U\Phi_t(z))^{-1} U[I - \Phi_t(z)\Phi_t(w)^*]U^*(I - \Phi_t(w)^* U^*)^{-1}.$$

Now suppose that $T \in \mathcal{L}(\mathcal{K})$ with $\sigma(T) \subset \mathbb{A}_q$ has $\mathbb{A}_q$ as a spectral set, so $\|s(T)\| \leqslant 1$ for all s in the scalar Schur class $\mathcal{S}(\mathbb{A}_q)$. Then an immediate consequence of the formula (5.6) combined with the diagonal form (5.5) of Φ_t is that it then follows that $\operatorname{Re} F_{\mathbf{x},\mathbf{w}}(T) \geqslant 0$ as well, and thus the spectral set question has an affirmative answer for the annulus $\mathbb{A}_q$. This line of reasoning arguably provides some simplifications to the solutions of the spectral set question for the annulus given in [2], [24]. We were informed by the referee that Jim Agler has unpublished work which works out an alternative simplification of the McCullough argument.

Acknowledgements. We would like to thank David Sherman of the University of Virginia for several helpful discussions on various topics developed in this paper, as well as the anonymous referee whose perceptive comments lead to a number of improvements in this final version of the paper.

REFERENCES

[1] M.B. ABRAHAMSE, The Pick interpolation theorem for finitely connected domains, *Michigan Math. J.* **26**(1979), 195–203.

[2] J. AGLER, Rational dilation on an annulus, *Ann. of Math.* **121**(1985), 106–148.

[3] J. AGLER, On the representation of certain holomorphic functions defined on a polydisc, in *Topics in Operator Theory: Ernst D. Hellinger Memorial Volume*, Oper. Theory Adv. Appl., vol. 48, Birkhäuser-Verlag, Basel 1990, pp. 47–66.

[4] J. AGLER, J. HARLAND, B.J. RAPHAEL, Classical function theory, operator dilation theory, and machine computation on multiply-connected domains, *Mem. Amer. Math. Soc.* **191**(2008), no. 892.

[5] W.B. ARVESON, Subalgebras of C^*-algebras, *Acta Math.* **123**(1969), 141–224.

[6] W.B. ARVESON, Subalgebras of C^*-algebras. II, *Acta Math.* **128**(1972), 271–308.

[7] J.A. BALL, V. BOLOTNIKOV, S. TER HORST, A constrained Nevanlinna–Pick interpolation problem for matrix-valued functions, *Indiana Univ. Math. J.* **59**(2010), 15–51.

[8] J.A. BALL, M.D. GUERRA HUAMAN, Test functions, Schur–Agler classes and transfer-function realizations: the matrix-valued case, *Complex Anal. Oper. Theory* **7**(2013), 529–575.

[9] E. BISHOP, K. DE LEEUW, The representations of linear functionals by measures on sets of extreme points, *Ann. Inst. Fourier (Grenoble)* **9**(1959), 305–331.

[10] G. CHOQUET, Existence et unicité des représentations intégrales au moyen des points extrémaux dans les cônes convexes, *Seminaire Bourbaki*, vol. 139 (Dec. 1956), 15 pp.

[11] K.R. DAVIDSON, V. PAULSEN, M. RAGHUPATHI, D. SINGH, A constrained Nevanlinna–Pick interpolation problem, *Indiana Univ. Math. J.* **58**(2009), 709–732.

[12] M.A. DRITSCHEL, S. MARCANTOGNINI, S. MCCULLOUGH, Interpolation in semigroupoid algebras, *J. Reine Angew. Math.* **606**(2007), 1–40.

[13] M.A. DRITSCHEL, S. MCCULLOUGH, The failure of rational dilation on a triply connected domain, *J. Amer. Math. Soc.* **18**(2005), 873–918.

[14] M.A. DRITSCHEL, S. MCCULLOUGH, Test functions, kernels, realizations and interpolation, in *Operator Theory, Structured Matrices, and Dilations: Tiberiu Constantinescu Memorial Volume*, Theta, Bucharest 2007, pp. 153–179.

[15] M.A. DRITSCHEL, J. PICKERING, Test functions in constrained interpolation, *Trans. Amer. Math. Soc.* **364**(2012), 5589–5604.

[16] D.R. FARENICK, P.B. MORENZ, C^*-extreme points of some compact C^*-convex sets, *Proc. Amer. Math. Soc.* **118**(1993), 765–775.

[17] D.R. FARENICK, P.B. MORENZ, C^*-extreme points in the generalized state space of a C^*-algebra, *Trans. Amer. Math. Soc.* **349**(1997), 1725–1748.

[18] S.D. FISHER, *Function Theory on Planar Domains: A Second Course in Complex Analysis*, Wiley and Sons, New York 1983; Second Edition: Dover Publ., New York 2007.

[19] M.C. GREGG, On C^*-extreme maps and $*$-homomorphisms of a commutative C^*-algebra, *Integral Equations Operator Theory* **63**(2009), 337–349.

[20] H. GRUNSKY, *Lectures on Theory of Functions in Multiply Connected Domains*, , Studia Math., Skript 4, Vandenhoeck and Ruprecht, Göttingen 1978.

[21] M.D. GUERRA HUAMAN, Schur-class of finitely connected planar domains: the test-function approach, Ph.D. Dissertation, Virginia Tech., Blacksburg, VA 2011.

[22] A. HOPENWASSER, R.L. MOORE, V.I. PAULSEN, C^*-extreme points, *Trans. Amer. Math. Soc.* **266**(1981), 291–307.

[23] R.I. LOEBL, V.I. PAULSEN, Some remarks on C^*-convexity, *Linear Algebra Appl.* **35**(1981), 63–78.

[24] S. MCCULLOUGH, Matrix functions of positive real part on an annulus, *Houston J. Math.* **21**(1995), 489–506.

[25] S. MCCULLOUGH, Isometric representations of some quotients of H^∞ of an annulus, *Integral Equations Operator Theory* **39**(2001), 335–362.

[26] S. MCCULLOUGH, V. PAULSEN, C^*-envelopes and interpolation theory, *Indiana Univ. Math. J.* **51**(2002), 479–505.

[27] R.R. PHELPS, *Lectures on Choquet's Theorem*, Second edition, Lecture Notes in Math., vol. 1757, Springer, New York 2001.

[28] R.S. PHILLIPS, On symplectic mappings of contraction operators, *Studia Mathematica* **31**(1968), 15–27.

[29] J. PICKERING, Counterexamples to rational dilation on symmetric multiply connected domains, *Complex Anal. Oper. Theory* **4**(2010), 55–95.

[30] M. RAGHUPATHI, Nevanlinna–Pick interpolation for $\mathbb{C} + BH^\infty$, *Integral Equations Operator Theory* **63**(2009), 103–125.

[31] W. RUDIN, *Functional Analysis*, Second edition, McGraw-Hill, New York 1991.

[32] M. TAKESAKI, *Theory of Operator Algebra*. I, Encylopaedia Math. Sci., vol. 124, Springer, New York 1979.

JOSEPH A. BALL, DEPARTMENT OF MATHEMATICS, VIRGINIA TECH, BLACKSBURG, VA 24061-0123, U.S.A.
E-mail address: joball@math.vt.edu

MOISÉS D. GUERRA HUAMÁN, CAL. JOSE A. QUINONES 135, URB. LUCYANA, DIST. CARABAYLLO LIMA, LIMA 6, PERU
E-mail address: moisesgggg@gmail.com

Received September 21, 2011.