CLARK THEORY IN THE DRURY-ARVESON SPACE

MICHAEL T. JURY

Abstract. We extend the basic elements of Clark's theory of rank-one perturbations of backward shifts, to row-contractive operators associated to de Branges-Rovnyak type spaces $\mathcal{H}(b)$ contractively contained in the Drury-Arveson space on the unit ball in $\mathbb{C}^d$. The Aleksandrov-Clark measures on the circle are replaced by a family of states on a certain noncommutative operator system, and the backward shift is replaced by a canonical solution to the Gleason problem in $\mathcal{H}(b)$. In addition we introduce the notion of a "quasi-extreme" multiplier of the Drury-Arveson space and use it to characterize those $\mathcal{H}(b)$ spaces that are invariant under multiplication by the coordinate functions.

1. Introduction

The purpose of this paper is to provide one method of extending the elementary portions of Clark's theory of rank-one perturbations of backward shifts, to the Drury-Arveson space H^2_d of the unit ball $\mathbb{B}^d \subset \mathbb{C}^d$. This is the space of functions holomorphic in $\mathbb{B}^d$ with reproducing kernel

$$\frac{1}{1 - zw^*}. \tag{1.1}$$

(Here $zw^* = \sum_{j=1}^d z_j w_j^*$ is the standard Hermitian inner product in $\mathbb{C}^d$.) When $d = 1$, this is of course the usual Hardy space H^2 in the unit disk. When $d > 1$, the space H^2_d is an analytic Besov space, but is in many ways a more appropriate higher-dimensional analog of H^2 than the classical Hardy space in the ball (which has the kernel $s(z,w) = (1 - zw^*)^{-d}$). The recent survey [24] provides an overview.

To begin with we explain what is meant by the "elementary portions of Clark's theory;" our treatment is heavily influenced by the exposition of Sarason [23] and the treatment of Aleksandrov-Clark measures in [6, Chapter 9]. In particular we take a point of view in which the de Branges-Rovnyak spaces are central. Let b be a non-constant function analytic in the unit disk $\mathbb{D} \subset \mathbb{C}$ and bounded by 1 there. (In Clark's original treatment [7] b was assumed to be an inner function; that is, $|b| = 1$ almost everywhere on the unit circle.) For this discussion we impose the simplifying normalization $b(0) = 0$. Associated to b is a reproducing kernel Hilbert space $\mathcal{H}(b)$, with kernel

$$k^b(z,w) = \frac{1 - b(z)b(w)^*}{1 - zw^*} \tag{1.2}$$

where $*$ denotes the complex conjugate. (In the inner case, $\mathcal{H}(b)$ is isometrically the orthogonal complement of the Beurling subspace bH^2.) We call $\mathcal{H}(b)$ the *deBranges-Rovnyak space*

Date: August 27, 2013.
Partially supported by NSF grant DMS-1101461.
Keywords: Aleksandrov-Clark measure, row contraction, de Branges-Rovnyak space, Gleason problem.
2000 *Mathematics subject classification:* 46E22, 47A20, 47B32.

associated to b; as a set it is contained in H^2 and for each $f \in \mathcal{H}(b)$ we have $\|f\|_{H^2} \leq \|f\|_{\mathcal{H}(b)}$, so that $\mathcal{H}(b)$ is contractively contained in H^2. The basic theory of $\mathcal{H}(b)$ spaces may be found in the original book of de Branges and Rovnyak [8], and the monograph of Sarason [23].

An important feature of the $\mathcal{H}(b)$ spaces is that they are invariant under the backward shift operator

$$S^*f(z) := \frac{f(z) - f(0)}{z} \tag{1.3}$$

Indeed, this is a chief reason for interest in the $\mathcal{H}(b)$ spaces; the operators $X := S^*|_{\mathcal{H}(b)}$ (and their analogs on the vector-valued $\mathcal{H}(b)$ spaces) serve as functional models for contractive operators on Hilbert space, see for example [4, 16]. We also note that, while b itself may or may not lie in $\mathcal{H}(b)$, it is always the case that S^*b belongs to $\mathcal{H}(b)$ [23, II-8].

By the Herglotz formula, for each unimodular scalar α there is a unique finite, positive measure μ_α on the unit circle $\mathbb{T}$ such that

$$\frac{1 + \alpha^* b(z)}{1 - \alpha^* b(z)} = \int_{\mathbb{T}} \frac{1 + z\zeta^*}{1 - z\zeta^*}\, d\mu_\alpha(\zeta) + it \tag{1.4}$$

where t is the imaginary part of the left-hand function at $z = 0$. The measures $\{\mu_\alpha\}_{\alpha\in\mathbb{T}}$ are called the *Aleksandrov-Clark measure* or *AC measures* for b. We refer to [6, Chapter 9] for a discussion of their properties. Let $P^2(\mu)$ denote the closure of the analytic polynomials in $L^2(\mu)$.

By the "elementary portions of Clark's theory" we mean the following three theorems adapted from [7].

Theorem 1.1. *For each $\alpha \in \mathbb{T}$, the formula*

$$(\mathcal{V}_\alpha f)(z) = (1 - \alpha^* b(z)) \int_{\mathbb{T}} \frac{f(\zeta)}{1 - z\zeta^*}\, d\mu_\alpha(\zeta) \tag{1.5}$$

defines a unitary operator from $P^2(\mu)$ onto $\mathcal{H}(b)$. [23, III-7]

To go further we assume that b is an extreme point of the unit ball of $H^\infty(\mathbb{D})$; this is the case if and only if $\int_{\mathbb{T}} \log(1 - |b(\zeta)|)\, dm(\zeta) = -\infty$. Since the Radon-Nikodym derivative of the measure μ_α is

$$\frac{d\mu_\alpha}{dm}(\zeta) = \frac{1 - |b(\zeta)|^2}{|1 - \alpha^* b(\zeta)|^2}, \tag{1.6}$$

it follows that $\int_{\mathbb{T}} \log\left(\frac{d\mu_\alpha}{dm}\right) dm = -\infty$. Thus by Szegő's theorem $P^2(\mu_\alpha) = L^2(\mu_\alpha)$ for each α, when b is extreme. In particular, in this case (and only this case) the isometry M_ζ acting on $P^2(\mu_\alpha)$ is unitary.

Theorem 1.2. *Let b be extreme and let X denote the backward shift operator restricted to $\mathcal{H}(b)$. Then for each α the rank-one perturbation*

$$U_\alpha^* := X + \alpha^* S^* b \otimes 1 \tag{1.7}$$

defines a unitary operator U_α. Each of these unitaries is cyclic (with cyclic vector 1), and the spectral measure of U_α with respect to 1 is the AC measure μ_α. Moreover the operator $\mathcal{V}_\alpha$ implements the spectral resolution of U_α:

$$U_\alpha \mathcal{V}_\alpha = \mathcal{V}_\alpha M_\zeta \tag{1.8}$$

where M_ζ denotes multiplication by the independent variable in $L^2(\mu_\alpha)$.[23, III-8]

Finally, one can say something about eigenvalues of U_α:

Theorem 1.3. *The number ζ is an eigenvalue of U_α if and only if b has finite angular derivative at ζ with $b(\zeta) = \alpha$; in this case the eigenspace is one-dimensional and spanned by the function*

$$k^b_\zeta(z) := \frac{1 - \alpha^* b(z)}{1 - z\zeta^*} \tag{1.9}$$

[6, Theorem 8.9.9]

Our goal, then, is to obtain analogs of Theorems 1.1–1.3 for deBranges-Rovnyak type subspaces of H^2_d. For this, we let b be a *contractive multiplier* of H^2_d. That is, b is an analytic function in the ball such that $bf \in H^2_d$ whenever $f \in H^2_d$, and multiplication by b contracts norms:

$$\|bf\|_{H^2_d} \leq \|f\|_{H^2_d}. \tag{1.10}$$

This operator is denoted M_b. In one dimension, b is a contractive multiplier if and only if b is bounded by 1 in the disk. In higher dimensions this condition is necessary but not sufficient. Nonetheless the algebra of bounded multipliers of H^2_d is in many ways a suitable analog of $H^\infty(\mathbb{D})$ in higher dimensions. It is true that b is a contractive multiplier if and only if the Hermitian kernel

$$k^b(z, w) := \frac{1 - b(z)b(w)^*}{1 - zw^*} \tag{1.11}$$

is positive semidefinite. When this is the case it is the reproducing kernel for a space $\mathcal{H}(b)$ which is contractively contained in H^2_d. Explicitly, $\mathcal{H}(b)$ is the range of the operator $I - M_b M_b^*$ on H^2_d, equipped with the unique norm making this operator a partial isometry. This norm is given by the expression

$$\|f\|^2_{\mathcal{H}(b)} = \sup_{g \in H^2_d} (\|f + bg\|^2_{H^2_d} - \|g\|^2_{H^2_d}) \tag{1.12}$$

(see [23, Chapter I]), though in this paper the description of the space in terms of its kernel will be more useful.

The extension of the Clark theory to the $\mathcal{H}(b)$ spaces just defined is not straightforward, for several reasons. First, the obvious analog of the backward shift S^* would be the d-tuple adjoints of the coordinate multipliers $M_{z_1}, \dots M_{z_d}$ on H^2_d (the *d-shift* of Arveson [1], Drury [9] and Müller-Vasilescu [15]). However the $\mathcal{H}(b)$ spaces are in general not invariant for the adjoints of the d-shift [3]. Following [2, 3], the correct operators to look for are those that solve the *Gleason problem* in $\mathcal{H}(b)$. That is, we seek operators $X_1, \dots X_d$ on $\mathcal{H}(b)$ such that for all $f \in \mathcal{H}(b)$ we have

$$f(z) - f(0) = \sum_{j=1}^d z_j (X_j f)(z) \tag{1.13}$$

and such that the tuple $(X_1, \dots X_d)$ is contractive in the sense that

$$\sum_{j=1}^d \|X_j f\|^2 \leq 1 - |f(0)|^2. \tag{1.14}$$

for all $f \in \mathcal{H}(b)$. (When $d = 1$, the restricted backward shift $X = S^*|_{\mathcal{H}(b)}$ always obeys this estimate, called the "inequality for difference quotients" in [8].) From [2, 3] we know contractive solutions always exist, but a principal difficulty is that, in general, such operators may not be unique.

The next obstacle is understanding what (if anything) can play the role of the AC measures μ_α. First consider a finite, positive measure μ on the unit sphere and define a function b in the ball by the formula

$$\frac{1+b(z)}{1-b(z)} = \int_{\partial\mathbb{B}^d} \frac{1+z\zeta^*}{1-z\zeta^*}\, d\mu(\zeta) \tag{1.15}$$

then b will be a contractive multiplier of H^2_d, but importantly, not every contractive multiplier admits such a representation [14]. The correct approach is to replace the Herglotz-like kernel $\frac{1+z\zeta^*}{1-z\zeta^*}$ with the "noncommutative" Herglotz kernel

$$\left(I + \sum_{j=1}^d z_j L_j^*\right)\left(I - \sum_{j=1}^d z_j L_j^*\right)^{-1} \tag{1.16}$$

where the L_j are Hilbert space operators obeying the relations

$$L_i^* L_j = \delta_{ij} I. \tag{1.17}$$

The measure μ must then be replaced with a positive linear functional on the operator system spanned by the NC Herglotz kernels (1.16) and their adjoints. (Such NC Herglotz kernels have been studied before, see e.g. [14, 13, 20].)

The remainder of the paper is organized as follows: in Section 2 we reprove the NC Herglotz formula from [14, 20] in the form in which we will need it, define the noncommutative AC states $\{\mu_\alpha\}_{\alpha\in\mathbb{T}}$ associated to b, and use them to define (via a GNS type construction) Hilbert spaces $P^2(\mu_\alpha)$. Using the NC Herglotz kernel we are then able to construct a "noncommutative normalized Fantappiè transform" $\mathcal{V}_\alpha$ which implements a unitary equivalence between $P^2(\mu_\alpha)$ and $\mathcal{H}(b)$. The section concludes with Theorem 2.8, which is our analog of Theorem 1.1.

In Section 3 we investigate the GNS construction in the noncommutative $P^2(\mu)$ spaces more closely and introduce the notion of a "quasi-extreme" multiplier b. It is these that will substitute for the extreme points of the unit ball of $H^\infty(\mathbb{D})$. We also introduce the coisometric d-tuples of operators $\mathbf{S}^\alpha$ which are a partial analog of the unitaries U_α in Clark's theory.

In Section 4 we consider the Gleason problem in $\mathcal{H}(b)$ and prove the crucial result that, when b is quasi-extreme as defined in Section 3, there is in fact a *unique* contractive solution $\mathbf{X} = (X_1, \dots X_d)$ to the Gleason problem in $\mathcal{H}(b)$. This result is non-trivial and uses in a fundamental way the noncommutative constructions of Section 3. (In one variable there would be nothing to do here, since the backward shift is trivially the unique solution to the Gleason problem, regardless of b.) In Section 5 we put the results of the previous two sections together to show that there is a unique rank-one perturbation of the (now unique) "backward shift" $\mathbf{X}$ that is unitarily equivalent, via the NC Fantappiè transfrom, to the adjoint of the GNS tuple, $\mathbf{S}^{\alpha*}$. This is Theorem 5.1, which is our extenstion of Theorem 1.2.

Finally, in Section 6 we prove the analog of Theorem 1.3, in which we show that for the GNS tuple $\mathbf{S}^\alpha$, the eigenvalue problem

$$\sum_{j=1}^{d} \zeta_j^* S_j^\alpha h = h \tag{1.18}$$

has a solution in $\mathcal{H}(b)$ if and only if b has a finite angular derivative at $\zeta \in \partial\mathbb{B}^d$ with $b(\zeta) = \alpha$. This is Theorem 6.2. Along the way we prove a number of other results, including a version of the Aleksandrov disintegration theorem in this setting (Theorem 2.9), and a characterization of those $\mathcal{H}(b)$ spaces which are invariant under multiplication by the coordinate functions z_j (Corollary 4.5); it turns out this is the case exactly when b is not quasi-extreme.

2. The NC Herglotz formula and NC Fantappiè transform

2.1. **Row contractions, row isometries, and dilations.** We begin by recalling some basic constructions in multivariable operator theory, in particular row isometries and the noncommutative disk algebra of Popescu [18] Let H be a Hilbert space. A *row contraction* is a d-tuple of operators $\mathbf{T} = (T_1, \dots T_d)$ in $B(H)$ satisfying

$$\sum_{j=1}^{n} T_j T_j^* \leq I. \tag{2.1}$$

In other words, the map

$$(T_1, \dots T_d) \begin{pmatrix} h_1 \\ \vdots \\ h_d \end{pmatrix} = \sum_{j=1}^{d} T_j h_j \tag{2.2}$$

is contractive from H^d to H (the direct sum of d copies of H), so when convenient we think of $\mathbf{T}$ as belonging to $B(H^d, H)$. Note that if $\mathbf{T}$ is isometric, then $\mathbf{T}^*\mathbf{T} = I_{H^d}$, and so

$$T_i^* T_j = \delta_{ij} I_H, \tag{2.3}$$

which says that the T_j are isometries with orthogonal ranges. If $\mathbf{T}$ is unitary, then also $\mathbf{T}\mathbf{T}^* = I_H$, which means equality holds in (2.1). In this case the T_j are called *Cuntz isometries*. We will consider both commuting and non-commuting row contractions; note however that if $\mathbf{T}$ is isometric then the relations (2.3) show that the T_j cannot commute.

Definition 2.1. Let $\mathbf{T}$ be a row contraction on H. An *isometric dilation* of $\mathbf{T}$ is a row isometry $\mathbf{V}$ acting on a larger Hilbert space $K \supset H$ such that for each $j = 1, \dots d$, the space H is invariant for V_j^* and $V_j^*|_H = T_j^*$. The dilation is called *unitary* (or a *Cuntz dilation*) if the V_j are Cuntz isometries. The dilation is called *minimal* if

$$K = \bigvee_{w \in \mathbb{F}_d^+} V_w H. \tag{2.4}$$

(That is, the smallest $\mathbf{V}$-invariant subspace containing H is K itself.) By results of Frazho [10], Bunce [5], and Popescu [17] every row contraction admits a minimal isometric dilation, which is unique up to unitary equivalence. More precisely, if T is a row contraction and V

and V' are minimal isometric dilations of T on Hilbert spaces $K \supset H$, $K' \supset H$ respectively, then the map

$$UV_w h := V'_w h \tag{2.5}$$

extends to a unitary transformation from K onto K' satisfying $UV_j = V'_j U$ for all j.

2.2. **The noncommutative disk algebra.** Let $\mathbb{F}_d^+$ denote the *free semigroup on d letters*, that is, the set of all finite words

$$w = i_1 i_2 \cdots i_m \tag{2.6}$$

where $m \geq 1$ is an integer, and the i_j are drawn from the set $\{1, 2, \dots d\}$. We also include in $\mathbb{F}_d^+$ the *empty word* $\varnothing$. The integer m is called the *length* of the word w, by convention the empty word has length zero. If $w, v \in \mathbb{F}_d^+$ are words of lengths m and n respectively, they can be concatenated to produce the word wv, of length $m+n$. By convention $\varnothing w = w \varnothing = w$ for all $w \in \mathbb{F}_d^+$. The *full Fock space* $\mathcal{F}_d$ is the Hilbert space $\ell^2(\mathbb{F}_d^+)$, with orthonormal basis $\{\xi_w\}_{w \in \mathbb{F}_d^+}$. The Hilbert space $\mathcal{F}_d$ supports bounded operators $L_1, \dots L_d$, defined by their action on the basis vectors ξ_w:

$$L_i \xi_w = \xi_{iw}. \tag{2.7}$$

It is straightforward to check that the L_i satisfy

$$L_j^* L_i = \delta_{ij} I. \tag{2.8}$$

Thus by the discussion above $\mathbf{L}$ is a row isometry. In particular, this implies that $\mathbf{L}\mathbf{L}^* = \sum_{i=1}^d L_i L_i^*$ is an orthogonal projection in $\mathcal{F}_d$; the range of this projection is the orthogonal complement of the one-dimensional space spanned by the vacuum vector $\xi_\varnothing$.

The *noncommutative disk algebra* $\mathcal{A}_d$ (we will fix d and abbreviate this to $\mathcal{A}$ is the norm-closed algebra of operators on $\mathcal{F}_d$ generated by $L_1, \dots L_d$ and the identity operator I. We will write $\mathcal{A}^*$ for the algebra of operators which are the adjoints of the operators in $\mathcal{A}$. The C* -algebra generated by the L_i is called the *Cuntz-Toeplitz algebra* $\mathcal{E}_d$. The norm closure of $\mathcal{A}_d + \mathcal{A}_d^*$ in $\mathcal{E}_d$ is called the *Cuntz-Toeplitz operator system.* (Recall that an operator system is a unital, self-adjoint linear subspace of a unital C*-algebra.) A theorem of Popescu [19] shows that the row isometry $\mathbf{L}$ is universal, in the following sense: if $(V_1, \dots V_d)$ is any row isometry, acting on a Hilbert space H, then there is a representation of the Cuntz-Toeplitz algebra $\pi : \mathcal{E}_d \to B(H)$ such that $\pi(L_i) = V_i$ for each $i = 1, \dots d$. More generally, if $\mathbf{T} = (T_1, \dots T_d)$ is any row contraction acting in a Hilbert space H, then there is a unital, completely positive map $\rho : \mathcal{E}_d \to B(H)$ such that $\rho(L_j) = T_j$.

A particular subsystem of the Cuntz-Toeplitz operator system will be of interest. For a d-tuple of nonnegative integers $\mathbf{n} = (n_1, \dots n_d)$, and an arbitrary d-tuple of operators $\mathbf{T} = (T_1, \dots T_d)$ we define the *symmetrized monomials*

$$T^{(\mathbf{n})} := \sum T_{i_1} \cdots T_{i_{|\mathbf{n}|}}$$

where the sum is taken over all products of exactly n_1 T_1's, n_2 T_2's, etc. So, for example if $d = 2$ and $\mathbf{n} = (2, 1)$, then

$$T^{(2,1)} = T_1^2 T_2 + T_2 T_1^2 + T_1 T_2 T_1$$

By convention we put $T^{(\mathbf{0})} = I$. In particular, if $z = (z_1, \dots z_d)$ is a d-tuple of scalars, and the monomial $z^{\mathbf{n}}$ is defined in the usual multi-index notation as $z^{\mathbf{n}} = z_1^{n_1} \cdots z_d^{n_d}$, then we have for each integer $k \geq 1$

$$(z_1 T_1 + \cdots + z_d T_d)^k = \sum_{|\mathbf{n}|=k} z^{\mathbf{n}} T^{(\mathbf{n})}. \tag{2.9}$$

The *symmetric part* $\mathcal{S}$ of $\mathcal{A}$ is defined to be the closed linear span of the symmetrized monomials $\{L^{(\mathbf{n})} : \mathbf{n} \in \mathbb{N}^d\}$. Much of our interest will be in positive linear functionals μ defined on the operator system $\mathcal{S} + \mathcal{S}^* \subset \mathcal{A} + \mathcal{A}^*$.

In what follows we will use the notation

$$z\mathbf{L}^* := \sum_{j=1}^{d} z_j L_j^*. \tag{2.10}$$

It follows that for all $z, w \in \mathbb{C}^d$,

$$(z\mathbf{L}^*)(\mathbf{L}w^*) = \sum_{j,k=1}^{d} z_j w_l^* L_j^* L_k = zw^* \tag{2.11}$$

by the orthogonality relations for the L_j. In particular by putting $z = w$ we have

$$\|z\mathbf{L}^*\| = |z| \tag{2.12}$$

and hence for all $z \in \mathbb{B}^d$ the operator $I - z\mathbf{L}^*$ is invertible, with inverse given by the (norm-convergent) series

$$(I - z\mathbf{L}^*)^{-1} = \sum_{k=0}^{\infty} (z\mathbf{L}^*)^k = \sum_{\mathbf{n} \in \mathbb{N}^d} z^{\mathbf{n}} \frac{\mathbf{n}!}{|\mathbf{n}|!} L^{(\mathbf{n})*}. \tag{2.13}$$

It follows that $(I - z\mathbf{L}^*)^{-1}$ belongs to $\mathcal{S}^*$ for all $z \in \mathbb{B}^d$.

The identity (2.11) explains the appearance of noncommutative methods in our treatment of the $\mathcal{H}(b)$ spaces in H_d^2. Note that in one variable, if z, w are complex numbers and $|\zeta| = 1$ then trivially

$$(z\zeta^*)(\zeta w^*) = zw^* \tag{2.14}$$

However if $z, w \in \mathbb{C}^d$, $d > 1$, and $\zeta \in \mathbb{C}^d$ has unit norm, then

$$(z\zeta^*)(\zeta w^*) = \sum_{j,k=1}^{d} z_j w_k^* \zeta_j^* \zeta_k \neq zw^*. \tag{2.15}$$

By replacing ζ with the row isometry $\mathbf{L}$, equation (2.11) "repairs" equation (2.15). (Indeed, note that the identity $z\mathbf{T}^*\mathbf{T}w^* = zw^*$ cannot hold for any commuting tuple $\mathbf{T}$ when $d > 1$.) The identity (2.11) is thus central to our development, especially in the proof of the key algebraic results in Proposition 2.7.

The following lemma will be used several times.

Lemma 2.2. *The linear span of the set*

$$\{(I - \mathbf{L}w^*)^{-1} : w \in \mathbb{B}^d\} \tag{2.16}$$

is norm dense in $\mathcal{S}$.

Proof. Let $\mathcal{M}$ denote the closed linear span of the $(I - \mathbf{L}w^*)^{-1}$ in $\mathcal{S}$ as w ranges over $\mathbb{B}^d$. First, note that if T is any operator on Hilbert space with $\|T\| \leq 1$, then by expanding $(I - rT)^{-1}$ in a geometric series, we have for each positive integer m

$$T^m = \lim_{r\to 0} \frac{1}{r^m}\left((I - rT)^{-1} - \sum_{n=0}^{m-1} r^n T^n\right) \tag{2.17}$$

where the limit exists in the operator norm. Since $I \in \mathcal{M}$, induction on this fact with $T = \mathbf{L}w^*$ shows that $(\mathbf{L}w^*)^m$ lies in $\mathcal{M}$ for all $w \in \mathbb{B}^d$ and all $m \geq 0$.

From this, it suffices to prove that for each fixed m, the span of $\{(\mathbf{L}w^*)^m : |w| < 1\}$ is equal to the span of the set $\{L^{(\mathbf{p})} : |\mathbf{p}| = m\}$. From (2.9), the former span is contained in the latter. If they are not equal, then by linear algebra there is a set of scalars $\{c_{\mathbf{p}} : |\mathbf{p}| = m\}$, not all 0, so that

$$\sum_{|\mathbf{p}|=m} c_{\mathbf{p}} w^{\mathbf{p}} = 0 \tag{2.18}$$

for all $|w| < 1$. But this is a polynomial which vanishes on the open ball $\mathbb{B}^d$, and hence must vanish identically, so all $c_{\mathbf{p}}$ are 0, a contradiction. □

The next lemma encodes a key observation used in what will follow, namely that if p, q are polynomials in $\mathcal{S}$, then $p(\mathbf{L})^* q(\mathbf{L})$ belongs to $\mathcal{S} + \mathcal{S}^*$. (This is essentially an elaboration of (2.11) which will allow us to carry out a GNS-type construction in $\mathcal{S} + \mathcal{S}^*$ in Section 3.1.) We will do the required calculation quite explicitly. First, some notation: for d-tuples of nonnegative integers $\mathbf{m} = (m_1, \dots m_d)$, $\mathbf{n} = (n_1, \dots n_d)$, say $\mathbf{m} \leq \mathbf{n}$ if and only if $m_i \leq n_i$ for each $i = 1, \dots d$. If $\mathbf{m} \leq \mathbf{n}$, define $\mathbf{n} - \mathbf{m} = (n_1 - m_1, \dots n_d - m_d)$.

Next, we introduce the *letter counting map* $\lambda : \mathbb{F}_d^+ \to \mathbb{N}^d$, which when applied to a word w returns the d-tuple $(n_1, \dots n_d)$ where n_1 is the number of 1's appearing in w, n_2 the number of 2's, etc. It is immediate from definitions that

$$L^{(\mathbf{n})} = \sum_{\lambda(w)=\mathbf{n}} L_w. \tag{2.19}$$

Lemma 2.3. *For all* $\mathbf{m}, \mathbf{n} \in \mathbb{N}^n$*,*

$$L^{(\mathbf{n})*} L^{(\mathbf{m})} = \begin{cases} \frac{|\mathbf{n}|!}{\mathbf{n}!} L^{(\mathbf{m}-\mathbf{n})} & \text{if } \mathbf{m} \geq \mathbf{n} \\ \frac{|\mathbf{m}|!}{\mathbf{m}!} L^{(\mathbf{n}-\mathbf{m})*} & \text{if } \mathbf{n} \geq \mathbf{m} \\ \frac{|\mathbf{n}|!}{\mathbf{n}!} I & \text{if } \mathbf{m} = \mathbf{n} \\ 0 & \text{otherwise,} \end{cases} \tag{2.20}$$

and hence if p, q *are polynomials in* $\mathcal{S}$ *then* $p(\mathbf{L})^* q(\mathbf{L})$ *lies in* $\mathcal{S} + \mathcal{S}^*$*.*

Proof. First suppose $\mathbf{m} \geq \mathbf{n}$. Fix w with $\lambda(w) = \mathbf{n}$. Let $E(w)$ denote the set of words in $\lambda^{-1}(\mathbf{m})$ whose initial string is w:

$$E(w) = \{u \in \mathbb{F}_d^+ | u = wv \text{ and } \lambda(u) = \mathbf{m}\}. \tag{2.21}$$

Note that this set is alternatively defined as

$$E(w) = \{wv \in \mathbb{F}_d^+ | \lambda(v) = \mathbf{m} - \mathbf{n}\} \tag{2.22}$$

Now, if $u \in \lambda^{-1}(\mathbf{m})$, then $L_w^* L_u = L_v$ if $u \in E(w)$ and $u = wv$, while $L_w^* L_u = 0$ if $u \notin E(w)$. Thus

$$L^{(\mathbf{n})*} L^{(\mathbf{m})} = \sum_{\lambda(w)=\mathbf{n}} \sum_{\lambda(u)=\mathbf{m}} L_w^* L_u \tag{2.23}$$

$$= \sum_{\lambda(w)=\mathbf{n}} \sum_{u=wv\in E(w)} L_w^* L_u \tag{2.24}$$

$$= \sum_{\lambda(w)=\mathbf{n}} \sum_{v\in\lambda^{-1}(\mathbf{n}-\mathbf{m})} L_v \tag{2.25}$$

$$= \frac{|\mathbf{n}|!}{\mathbf{n}!} L^{(\mathbf{m}-\mathbf{n})} \tag{2.26}$$

since the cardinality of $\lambda^{-1}(\mathbf{n})$ is $\frac{|\mathbf{n}|!}{\mathbf{n}!}$. The cases $\mathbf{n} \geq \mathbf{m}$ and $\mathbf{n} = \mathbf{m}$ follow by symmetry.

Finally, if $\mathbf{m}$ and $\mathbf{n}$ are incomparable, then no word in $\lambda^{-1}(\mathbf{m})$ is a subword of a word in $\lambda^{-1}(\mathbf{n})$ and vice versa, so each summand $L_w^* L_u$ is 0. □

2.3. **The space $P^2(\mu)$.** Now, if μ is a positive linear functional on $\mathcal{S}^* + \mathcal{S}$, Lemma 2.3 allows us to define a pre-inner product on $\mathcal{S} \times \mathcal{S}$: for polynomials $p, q \in \mathcal{S}$, define

$$\langle p, q \rangle := \mu(q(\mathbf{L})^* p(\mathbf{L})). \tag{2.27}$$

Since μ is a positive linear functional, this map obeys the Cauchy-Schwarz inequality

$$|\mu(q(\mathbf{L})^* p(\mathbf{L}))|^2 \leq \mu(q(\mathbf{L})^* q(\mathbf{L}))\mu(p(\mathbf{L})^* p(\mathbf{L})) \tag{2.28}$$

and thus extends to a pre-inner product on all of $\mathcal{S} \times \mathcal{S}$. We will write $P^2(\mu)$ for the Hilbert space obtained by modding out null vectors and completing, and for $p \in \mathcal{S}$ we write $[p]$ for the image of p in $P^2(\mu)$.

2.4. **de Branges-Rovnyak spaces and the "noncommutative" AC measures.** Let b be a contractive multiplier of H_d^2 (hereafter we will just call b a *multiplier*. From the introduction, we have the reproducing kernel Hilbert space $\mathcal{H}(b)$ with kernel

$$k^b(z, w) = \frac{1 - b(z)b(w)^*}{1 - zw^*} \tag{2.29}$$

Central to our development will be the following noncommutative Herglotz-style formula for b. Such a formula is established in [14] and [20], we include a proof here since it is short (and to establish the role of the operator system $\mathcal{S} + \mathcal{S}^*$). The formula is based on the NC Herglotz kernel

$$H(z, L) := (I + z\mathbf{L}^*)(I - z\mathbf{L}^*)^{-1} \tag{2.30}$$

For each z in the ball, the operator $H(z, L)$ has positive real part, indeed using the relation (2.11) one finds

$$H(z, L) + H(z, L)^* = (1 - |z|^2)(I - z\mathbf{L}^*)^{-1}(I - z\mathbf{L}^*)^{*-1}. \tag{2.31}$$

Proposition 2.4. *Let b be a contractive multiplier of the Drury-Arveson space H_d^2. Then there exists a unique positive linear functional μ on $\mathcal{S} + \mathcal{S}^*$ such that*

$$\frac{1 + b(z)}{1 - b(z)} = \mu\left((I + z\mathbf{L}^*)(I - z\mathbf{L}^*)^{-1}\right) + iIm\left(\frac{1 + b(0)}{1 - b(0)}\right). \tag{2.32}$$

Proof. Consider the analytic function

$$f(z) = (1 + b(z))(1 - b(z))^{-1}$$

and observe that f belongs to the *positive Schur class*, i.e. the kernel

$$\frac{f(z) + f(w)^*}{1 - zw^*}$$

is positive. Indeed it factors as

$$(1 - b(z))^{-1}\frac{1 - b(z)b(w)^*}{1 - zw^*}(1 - b(w)^*)^{-1}$$

We may thus factor

$$f(z) + f(w)^* = F(z)[1 - zw^*]F(w)^*$$

for a holomorphic function F taking values in some auxiliary Hilbert space H. Substituting in turn $w = 0$, $z = 0$, and $z = w = 0$, we get

$$\begin{aligned} f(z) + f(0)^* &= F(z)F(0)^* \\ f(0) + f(w)^* &= F(0)F(w)^* \\ f(0) + f(0)^* &= F(0)F(0)^* \end{aligned}$$

Adding the first and last equations and subtracting the middle two leaves

$$F(z)zw^*F(w)^* = [F(z) - F(0)][F(w) - F(0)]^*$$

By the lurking isometry argument, there exists an isometric d-tuple $\mathbf{V} = (V_1, \dots V_d)$ on H such that

$$\sum_{j=1}^{d} w_j^* V_j F(w)^* = F(w)^* - F(0)^*$$

Solving for $F(w)^*$ gives

$$F(w)^* = (I - w\mathbf{V}^*)^{-1}F(0)^*$$

or

$$f(z) + f(0)^* = F(z)F(0)^* = F(0)(I - z\mathbf{V})^{-1}F(0)^* \tag{2.33}$$

By the Frazho-Bunce-Popescu dilation theorem, the tuple $\mathbf{V}$ is the image of $\mathbf{L}$ under a unital, completely positive map ψ. We now define

$$\mu(L^{(\mathbf{n})}) = F(0)\psi(L^{(\mathbf{n})})F(0)^* \tag{2.34}$$

which shows that μ is positive, since ψ is positive. With this definition and some algebra, (2.33) becomes

$$f(z) = \mu\left((I + z\mathbf{L}^*)(I - z\mathbf{L}^*)^{-1}\right) + i\mathrm{Im}f(0) \tag{2.35}$$

as desired. The uniqueness of μ is clear, since by (2.35) the value of $\mu(L^{(\mathbf{n})})$ is just the coefficient of $z^{(\mathbf{n})}$ in the Taylor expansion of f (with $\mu(I) = \mathrm{Re}f(0)$ when $\mathbf{n} = 0$). □

This proposition has a converse; namely if μ is a positive functional on $\mathcal{S}+\mathcal{S}^*$ and b is defined by (2.32) then b is a contractive multiplier of H^2_d; it follows as in [14] by reversing the steps of the above argument. The principal reason for introducting $\mathcal{S}+\mathcal{S}^*$ is that it forces the functional μ representing b to be unique; this need not be the case if we worked with $\mathcal{A}+\mathcal{A}^*$ (or, say, the whole Cuntz-Toeplitz algebra).

With b fixed and α a unimodular scalar, we can carry out the above construction with $\alpha^* b$ in place of b. We then have

Definition 2.5. Let b be a contractive multiplier of H^2_d. The *Aleksandrov-Clark state* (or *AC states*) for b are the family of states $\{\mu_\alpha\}_{\alpha\in\mathbb{T}}$ on $\mathcal{S}+\mathcal{S}^*$ such that

$$\frac{1+\alpha^* b(z)}{1-\alpha^* b(z)} = \mu_\alpha\left((I+z\mathbf{L}^*)(I-z\mathbf{L}^*)^{-1}\right) + i\mathrm{Im}\left(\frac{1+\alpha^* b(0)}{1-\alpha^* b(0)}\right) \tag{2.36}$$

as in Proposition 2.4.

If we compare the Herglotz-type formula (2.32) with the classical one-variable formula

$$\frac{1+b(z)}{1-b(z)} = \int_{\mathbb{T}} \frac{1+z\zeta^*}{1-z\zeta^*}\, d\mu(\zeta) + i\mathrm{Im}\frac{1+b(0)}{1-b(0)} \tag{2.37}$$

this suggests viewing the expression

$$(I+z\mathbf{L}^*)(I-z\mathbf{L}^*)^{-1} \tag{2.38}$$

as a noncommutative Herglotz kernel, and

$$(I-z\mathbf{L}^*)^{-1} \tag{2.39}$$

as a noncommutative Szegő kernel. This is explored further in the next section.

2.5. **The NC Fantappie transform.** Consider again the one variable case. As noted in the introduction, if μ is an AC measure for b, then Theorem 1.1 says that the normalized Cauchy transform

$$\mathcal{V}_\mu(f)(z) := (1-b(z))\int_{\mathbb{T}} \frac{f(\zeta)}{1-z\zeta^*}\, d\mu \tag{2.40}$$

implements a unitary operator from $P^2(\mu)$ onto $H(b)$. We are now ready to prove the analog of this theorem in the ball.

Definition 2.6. Let μ be a state on $\mathcal{S}+\mathcal{S}^*$, representing a multiplier b. For a polynomial $p\in\mathcal{S}$, the *normalized NC Fantappiè transform* of p is

$$\mathcal{V}_\mu(p)(z) := (1-b(z))\mu((1-z\mathbf{L}^*)^{-1}p(\mathbf{L})). \tag{2.41}$$

Using Lemma 2.3 and the fact that the series expansion of $(1-z\mathbf{L}^*)^{-1}$ is norm convergent in $\mathcal{S}^*$, one sees that $(1-z\mathbf{L}^*)^{-1}p(\mathbf{L})$ belongs to the closure of $\mathcal{S}+\mathcal{S}^*$, so $\mathcal{V}_\mu$ is defined. Our next goal is to show that $\mathcal{V}_\mu$ extends to a unitary operator from $P^2(\mu)$ onto $\mathcal{H}(b)$.

We will also use the notation

$$\mathcal{G}_\mu p(z) := \mu((1-z\mathbf{L}^*)^{-1}p(\mathbf{L})) \tag{2.42}$$

so that $\mathcal{V}_\mu p = (1-b(z))\mathcal{G}_\mu p$. Once we show that $\mathcal{V}_\mu$ extends to a unitary operator on $\mathcal{P}^2(\mu)$, it follows that $\mathcal{G}_\mu$ also extends to a well-defined linear operator taking $P^2(\mu)$ into the space of holomorphic functions on the ball.

To streamline the notation, write

$$H(z,L) = (I + z\mathbf{L}^*)(I - z\mathbf{L}^*)^{-1}. \tag{2.43}$$

Proposition 2.7. *For all $z, w \in \mathbb{B}^d$,*

$$(I - z\mathbf{L}^*)^{-1}(I - \mathbf{L}w^*)^{-1} = \frac{1}{2}\left(\frac{H(z,L) + H(w,L)^*}{1 - zw^*}\right) \tag{2.44}$$

In particular, if μ is a positive linear functional on $\mathcal{S} + \mathcal{S}^*$ and μ represents b as in (2.32), then

$$\mu((I - z\mathbf{L}^*)^{-1}(I - \mathbf{L}w^*)^{-1}) = \frac{1}{2}\frac{1}{1 - zw^*}\left(\frac{1 + b(z)}{1 - b(z)} + \frac{1 + b(w)^*}{1 - b(w)^*}\right) \tag{2.45}$$

$$= \frac{1}{(1 - b(z))(1 - b(w)^*)}\left(\frac{1 - b(z)b(w)^*}{1 - zw^*}\right) \tag{2.46}$$

Proof. Working with the right-hand side of (2.44), factor out $(I - z\mathbf{L}^*)^{-1}$ from the left and $(I - \mathbf{L}w^*)^{-1}$ from the right, leaving

$$\begin{aligned} &\frac{1}{2}\frac{H(z,L) + H(w,L)^*}{1 - zw^*} \\ &= (I - z\mathbf{L}^*)^{-1}\left(\frac{1}{2}\frac{(I + z\mathbf{L}^*)(I - \mathbf{L}w^*) + (I - z\mathbf{L}^*)(I + \mathbf{L}w^*)}{1 - zw^*}\right)(I - \mathbf{L}w^*)^{-1} \end{aligned} \tag{2.47}$$

$$= (I - z\mathbf{L}^*)^{-1}\frac{1}{2}\frac{2(I - (z\mathbf{L}^*)(\mathbf{L}w^*)}{1 - zw^*}(I - \mathbf{L}w^*)^{-1} \tag{2.48}$$

$$= (I - z\mathbf{L}^*)^{-1}(I - \mathbf{L}w^*)^{-1} \tag{2.49}$$

where the last equality follows from (2.11). Equations (2.45) and (2.46) follow immediately. ☐

Theorem 2.8. *Let b be a multiplier with AC state μ. Then the normalized NC Fantappiè transform $\mathcal{V}_\mu$ extends to a unitary operator from P^2_μ onto $\mathcal{H}(b)$.*

Proof. For each $w \in \mathbb{B}^n$, define

$$G_w(L) = (1 - b(w)^*)(I - \mathbf{L}w^*)^{-1}. \tag{2.50}$$

Let us write $[G_w]$ for the vector in $P^2(\mu)$ associated to G_w in the construction of $P^2(\mu)$. By Lemma 2.2, the span of the $[G_w]$ is dense in $P^2(\mu)$. Then (2.46) shows that $\langle [G_w], [G_z]\rangle_\mu = \langle k^b_w, k^b_z\rangle_{H(b)}$ for all $z, w \in \mathbb{B}^n$, so the map sending G_w to k^b_w is an isometry from the span of the G_w onto the span of the k^b_w, and thus extends uniquely to a unitary from P^2_μ onto $\mathcal{H}(b)$. But by (2.46) again, the map sending G_w to k^b_w just is the normalized NC Fantappiè transform. ☐

On $\mathcal{A} + \mathcal{A}^*$ there is a distinguished state called the *vacuum state*, which is the vector state induced by the vacuum vector $\xi_\varnothing$. That is, for polynomials $p, q \in \mathcal{A}$ we define

$$m_\varnothing(p + q^*) := \langle (p(\mathbf{L}) + q(\mathbf{L})^*)\xi_\varnothing, \xi_\varnothing\rangle. \tag{2.51}$$

Inspecting the moments we find that, since $\xi_\varnothing$ is a wandering vector for $\mathbf{L}$, we have $m_\varnothing(I) = 1$ and $m_\varnothing(L_w) = 0$ for $w \neq \varnothing$. Thus $m_\varnothing$ can be thought of as an analogue of Lebesgue measure m, which is the measure on $\mathbb{T}$ (or, state on $C(\mathbb{T})$) with moments $\widehat{m}(1) = 1$ and $\widehat{m}(z^n) = 0$

for $n \neq 0$. The analogy is strengthened by noting that if we restrict $m_\varnothing$ to $\mathcal{S}+\mathcal{S}^*$, then $m_\varnothing$ is an AC state for $b \equiv 0$, and hence $\mathcal{H}(b)$ is exactly the Drury-Arveson space H^2_d. Explicitly, Theorem 2.8 applied to the function $b \equiv 0$ with associated state $m_\varnothing$ says

$$\frac{1}{1-zw^*} = \langle k^b_w, k^b_z\rangle_{H^2_d} = m_\varnothing((I-z\mathbf{L}^*)^{-1}(I-\mathbf{L}w^*)^{-1}) \tag{2.52}$$

which can be compared to the classical one variable identity

$$\frac{1}{1-zw^*} = \int_{\mathbb{T}} \frac{1}{1-z\zeta^*}\frac{1}{1-\zeta w^*}\, dm(\zeta). \tag{2.53}$$

More generally, the equation (2.46) is in one variable the identity

$$\frac{1}{(1-b(z))(1-b(w)^*)}\frac{1-b(z)b(w)^*}{1-zw^*} = \int_{\mathbb{T}} \frac{1}{1-z\zeta^*}\frac{1}{1-\zeta w^*}\, d\mu(\zeta) \tag{2.54}$$

(see [23, III-6]). Indeed the identity (2.11) means that the proofs given in this section reduce to those of [23, Chapter III] when $d=1$.

Even more, the vacuum state $m_\varnothing$ supports a version of the Aleksandrov disintegration theorem for the AC states μ_α associated to a fixed b (Definition 2.5). Indeed the proof in our setting is essentially the same as that given in [6, Theorem 9.3.2] in the one-variable case.

Theorem 2.9 (Aleksandrov disintegration for AC states)**.** *Let m denote normalized Lebesgue measure on $\mathbb{T}$, $m_\varnothing$ the vacuum state on $\overline{\mathcal{S}+\mathcal{S}^*}$, and $\{\mu_\alpha\}_{\alpha\in\mathbb{T}}$ the AC states for a contractive multiplier b. Then for all $f \in \overline{\mathcal{S}+\mathcal{S}^*}$, the function $\alpha \to \mu_\alpha(f)$ is continuous in α, and*

$$\int_{\mathbb{T}} \mu_\alpha(f)\, dm(\alpha) = m_\varnothing(f). \tag{2.55}$$

Proof. Using the positivity of the μ_α and Lemma 2.2, it suffices to prove the theorem when $f = (I-z\mathbf{L}^*)^{-1}$ for fixed $|z|<1$. In this case by (2.46) we have

$$\mu_\alpha(f) = \frac{1-b(z)b(0)^*}{(1-\alpha^* b(z))(1-\alpha b(0)^*)} \tag{2.56}$$

which is continuous in α (note z is fixed here and $|b(z)|<1$). On the one hand, by definition of $m_\varnothing$ we have $m_\varnothing(f)=1$. On the other hand, integrating (2.56) we have (using the classical formula (2.53) for the inner product of Szegő kernels)

$$\int_{\mathbb{T}} \frac{1-b(z)b(0)^*}{(1-\alpha^* b(z))(1-\alpha b(0)^*)}\, dm(\alpha) = (1-b(z)b(0)^*)\int_{\mathbb{T}} \frac{1}{1-\alpha^* b(z)}\frac{1}{1-\alpha b(0)^*}\, dm(\alpha) \tag{2.57}$$

$$= \frac{1-b(z)b(0)^*}{1-b(z)b(0)^*} = 1. \tag{2.58}$$

We conclude that

$$\int_{\mathbb{T}} \frac{1}{1-b(z)\alpha^*}\, dm(\alpha) = 1 = m_\varnothing(f). \tag{2.59}$$

□

3. The GNS construction in $P^2(\mu)$

In this section we carry out a version of the GNS construction in the noncommutative $P^2(\mu)$ spaces of Section 2.3. This construction and the notions arising out of it (particularly that of a *quasi-extreme* multiplier) will be central to the rest of the paper. In one variable, if μ is a measure on the circle then multiplication by the independent variable ζ is an isometric operator on $P^2(\mu)$, which is unitary in the case that $P^2(\mu) = L^2(\mu)$ (equivalently, $P^2_0(\mu) = P^2(\mu)$). In the present setting the fact that $\mathcal{S}$ (the symmetric part of the NC disk algebra) is not an algebra will complicate matters. In the end we will obtain a contractive tuple $\mathbf{S}$ acting on a closed subspace $P^2_0(\mu)$ of $P^2(\mu)$, which will be coisometric in the case that $P^2(\mu) = P^2_0(\mu)$.

The GNS construction for states on the full Cuntz-Toeplitz operator system $\mathcal{A} + \mathcal{A}^*$ is well known; we recount it briefly. Suppose H is a Hilbert space and $A \subset B(H)$ is a linear subspace containing I. Then

$$A^* + A = \{b^* + a | a, b \in A\} \tag{3.1}$$

and

$$A^*A = \text{span}\{b^*a | a, b \in A\} \tag{3.2}$$

are operator systems containing A, and since A is unital we have $A^* + A \subseteq A^*A$.

For $A = \mathcal{A}$, the noncommutative disk algebra, consider the operator system $\mathcal{M} = \mathcal{A}^* + \mathcal{A}$. One sees easily from the relations (2.3) that for all words w, v, the operator $L_w^* L_v$ belongs either to $\mathcal{A}$ or $\mathcal{A}^*$. It follows that

$$\mathcal{A}^*\mathcal{A} \subset \overline{\mathcal{A}^* + \mathcal{A}} \tag{3.3}$$

This fact allows us to construct a "left regular representation" of $\mathcal{A}$ starting from any state ν on $\mathcal{A} + \mathcal{A}^*$. (Here we abuse the terminology slightly and allow "state" to mean any positive linear functional; it need not be normalized to have $\nu(I) = 1$.) In what follows we will often elide the distinction between positive functionals on $\mathcal{A} + \mathcal{A}^*$ and their unique extensions to positive functionals on $\overline{\mathcal{A} + \mathcal{A}^*}$; in practice this should cause no difficulties. A similar remark will of course be in force for $\mathcal{S} + \mathcal{S}^*$.

Given ν, the pairing on $\mathcal{A} \times \mathcal{A}$ given by

$$\langle b, c \rangle := \nu(c^*b) \tag{3.4}$$

is a pre-inner product on $\mathcal{A}$; quotienting by null vectors and completing gives a Hilbert space H_ν. For $a \in \mathcal{A}$, let $[a]$ denote the corresponding vector in H_ν. Now it is routine to check that for each $a \in \mathcal{A}$, the equation

$$\pi(a)[b] := [ab] \tag{3.5}$$

(that is, "left multiplication by a") defines a bounded linear operator on H_ν, and the map $\pi : \mathcal{A} \to B(H_\nu)$ is a completely contractive unital homomorphism. Moreover, it is not hard to show that the d-tuple $\pi(\mathbf{L}) = (\pi(L_1), \dots \pi(L_d))$ is a row isometry. Indeed, we have for all

$b, c \in \mathcal{A}$ and all $i, j = 1, \dots d$

$$\langle \pi(L_i)^* \pi(L_j)[b], [c] \rangle = \langle \pi(L_j)[b], \pi(L_i)[c] \rangle \tag{3.6}$$
$$= \nu(c^* L_i^* L_j b) \tag{3.7}$$
$$= \delta_{ij} \mu(c^* b) \tag{3.8}$$
$$= \delta_{ij} \langle [b], [c] \rangle \tag{3.9}$$

so $\pi(L_j)^* \pi(L_i) = \delta_{ij} I$.

By definition, for $a, b, c \in \mathcal{A}$ we have $\langle \pi(a) b, c \rangle := \nu(c^* ab)$. In particular, fixing a word w and taking $a = L_w$, $b = c = I$, each state ν is a vector state in the GNS representation:

$$\nu(L_w) = \langle \pi(L_w)[I], [I] \rangle. \tag{3.10}$$

3.1. **The GNS construction in $\mathcal{S} + \mathcal{S}^*$.** The next goal is to imitate the above construction with the NC disk algebra $\mathcal{A}$ replaced by its symmetric part $\mathcal{S}$. The fact that $\mathcal{S}$ is not an algebra means the construction must be modified; it is Lemma 2.3 that makes it possible at all.

Let $\mathcal{S}_0$ be the subspace of $\mathcal{S}$ given by

$$\mathcal{S}_0 := \text{span}\{L^{(\mathbf{n})} : |\mathbf{n}| \geq 1\}; \tag{3.11}$$

so that $\mathcal{S} = \text{span}\{I, \mathcal{S}_0\}$. Let $P_0^2(\mu)$ denote the closed subspace of $P^2(\mu)$ spanned by the set $\{[p] : p \in \mathcal{S}_0\}$ and $P_0 : P^2(\mu) \to P_0^2(\mu)$ the orthogonal projection. (It is possible that $P_0^2(\mu) = P^2(\mu)$.)

Proposition 3.1. *Let $p \in \mathcal{S}_0$ be a polynomial. For each $j = 1, \dots d$ the map*

$$[p] \to [L_j^* p(\mathbf{L})] \tag{3.12}$$

is well defined, and extends to a bounded linear operator from $P_0^2(\mu)$ to $P^2(\mu)$. Morevoer the operator $\mathbf{S} = (S_1, \dots S_d)$ defined by

$$S_j^*[p] := P_0[L_j^* p(\mathbf{L})] \tag{3.13}$$

is a row contraction on $P_0^2(\mu)$.

Proof. By Lemma 2.3, if $p \in \mathcal{S}_0$ then $L_i^* p(\mathbf{L}) \in \mathcal{S}$ for each $i = 1, \dots n$, and again by the lemma $q(\mathbf{L})^* L_i^* p(\mathbf{L}) \in \mathcal{S} + \mathcal{S}^*$, so belongs to the domain of μ. For each $i = 1, \dots d$, the pairing

$$([p], [q])_i = \mu(q(\mathbf{L})^* L_i^* p(\mathbf{L})), \qquad p, q \in \mathcal{S}_0 \tag{3.14}$$

gives a well-defined, bounded bilinear form on the span of $\{[p] : p \in \mathcal{S}_o\}$ in $P_0^2(\mu)$. Indeed, since $L_i^* L_i = I$ for each i, the Cauchy-Schwarz inequality for μ gives

$$|([p], [q])_i| = |\mu(q(\mathbf{L})^* L_i^* p(\mathbf{L}))| \leq \mu(q(\mathbf{L})^* L_i^* L_i q(\mathbf{L}))^{1/2} \mu(p(\mathbf{L})^* p(\mathbf{L}))^{1/2} = \|[p]\| \|[q]\| \tag{3.15}$$

so $(\cdot, \cdot)_i$ is well defined and bounded (with norm at most 1). Thus each of the maps (3.12) is bounded, and the operators S_j^* of (3.13) are bounded. To see that $\mathbf{S} = (S_1, \dots S_n)$ is a row contraction, we have for all $p \in \mathcal{S}_0$,

$$\begin{aligned}\langle\sum_{i=1}^{d} S_i S_i^*[p],[p]\rangle_{P_0^2(\mu)} &= \sum_{i=1}^{d}\langle P_0[L_i^*p], P_0[L_i^*p]\rangle_{P^2(\mu)} \\ &\leq \sum_{i=1}^{d}\langle [L_i^*p],[L_i^*p]\rangle_{P^2(\mu)} \\ &= \sum_{i=1}^{d}\mu(p(\mathbf{L})^* L_i L_i^* P(L)) \\ &= \mu(p(\mathbf{L})^* p(\mathbf{L})) \\ &= \langle [p],[p]\rangle_{P_0^2(\mu)}\end{aligned}$$

(Equality holds in the second-to-last line since $p \in \mathcal{S}_0$, which entails $p(\mathbf{L}) = \sum_{j=1}^{d} L_j L_J^* p(\mathbf{L})$). □

Remark: It is very important to observe that at this point, we cannot assert a GNS-style representation of μ in terms of $\mathbf{S}$; that is, the above construction does *not* imply that

$$\mu(L^{(\mathbf{n})}) = \langle S^{(\mathbf{n})}[I],[I]\rangle_\mu \tag{3.16}$$

Indeed, as things stand the equation (3.16) does not even make sense, since $\mathbf{S}$ is only defined on $P_0^2(\mu)$, which need not contain $[I]$. However such a representation of μ is available when $[I]$ belongs to $P_0^2(\mu)$ (that is, when $P_0^2(\mu) = P^2(\mu)$). To prove this it will be helpful to consider extensions ν of μ to the full Cuntz-Toeplitz operator system $\mathcal{A}+\mathcal{A}^*$, and compare the GNS tuple $\mathbf{U} := \pi(\mathbf{L})$ to $\mathbf{S}^\mu$. More precisely, let ν be a state on $\mathcal{A}+\mathcal{A}^*$ and let us write $Q^2(\nu)$ for the GNS space associated to ν. Inside $Q^2(\nu)$ there is a subspace $Q_0^2(\nu)$ formed by taking the closed span of the elements

$$\{[L_w] : |w| \geq 1\} \tag{3.17}$$

in $Q^2(\nu)$. We let Q_0 denote the orthogonal projection onto $Q_0^2(\nu)$. Now, if μ is a state on $\mathcal{S}+\mathcal{S}^*$ and ν extends μ, the inclusion $\mathcal{S} \subset \mathcal{A}$ induces isometric inclusions of the Hilbert spaces

$$P^2(\mu) \subset Q^2(\nu), \quad P_0^2(\mu) \subset Q_0^2(\nu). \tag{3.18}$$

Let us write $\mathbf{U} = (U_1,\dots,U_d) := (\pi(L_1),\dots,\pi(L_d)$ for the GNS tuple for ν acting in $Q^2(\mu)$. By construction the subspace $Q_0^2(\nu)$ is invariant for the U_j, so we can define $\mathbf{V}$ to be the restriction of $\mathbf{U}$ to $Q_0^2(\nu)$.

We now consider the following definition:

Definition 3.2. Let μ be a state on $\mathcal{S}^*+\mathcal{S}$ and ν be a state on $\mathcal{A}^*+\mathcal{A}$ extending μ, and $\mathbf{S},\mathbf{U}$ the GNS operators associated to μ and ν respectively. The extension μ will be called *tight* if $\mathbf{V} = \mathbf{U}|_{Q_0^2(\nu)}$ is a dilation of $\mathbf{S}$. A state ν on $\mathcal{A}+\mathcal{A}^*$ is called *tight* if it is a tight extension of its restriction $\mu = \nu|_{\mathcal{S}^*+\mathcal{S}}$.

In other words, starting from a state μ on the symmetric operator system $\mathcal{S}+\mathcal{S}^*$, we have two ways of constructing row contractions on $P_0^2(\mu)$. One is to construct the GNS tuple $\mathbf{S}$ of Proposition 3.1. The other is to extend the state μ to a state ν on $\mathcal{A}+\mathcal{A}^*$, form the GNS tuple $\mathbf{U}$ on $Q^2(\mu)$, then compress this tuple to $P_0^2(\mu) \subset Q^2(\mu)$. To call the extension ν tight is to say these constructions coincide. We will also see shortly that if $\mathbf{V}$ is a dilation of $\mathbf{S}$, then it is necessarily a minimal dilation of $\mathbf{S}$.

At present we do not know whether or not tight extensions always exist. The next theorem gives a somewhat more transparent spatial condition which characterizes tight extensions.

Theorem 3.3. *Let μ be a state on $\mathcal{S} + \mathcal{S}^*$ and ν an extension of μ to $\mathcal{A} + \mathcal{A}^*$. Then ν is a tight extension if and only if $P_0[I] = Q_0[I]$.*

Proof. Let $\mathbf{U} = (U_1, \dots U_d)$ be the GNS tuple for ν. By definition the extension is tight if and only if the restriction of the U's to $Q_0^2(\nu)$ form a dilation of the S's. This happens if and only if for all polynomials $p \in \mathcal{S}_0$,

$$Q_0 U_i^*[p] = S_i^*[p], \tag{3.19}$$

or more explicitly

$$Q_0[L_i^* p(\mathbf{L})] = P_0[L_i^* p(\mathbf{L})] \tag{3.20}$$

Of course, (3.20) will always hold when $L_i^* p(\mathbf{L}) \in \mathcal{S}_0$; the p's with this property are the span of the set $\{L^{(\mathbf{n})} : |\mathbf{n}| \geq 2\}$. So what is at issue are the cases $p(\mathbf{L}) = L_j$. In this case, if $i \neq j$, then both sides of (3.20) are 0, while if $i = j$ we obtain the condition $Q_0[I] = P_0[I]$. □

Proposition 3.4. *If ν is a tight extension of μ, then $\mathbf{V}$ is a minimal dilation of $\mathbf{S}$.*

Proof. We maintain the notation used above. By construction $P_0^2(\mu)$ contains the vectors $[L_1], \dots [L_n]$ (since the L_i belong to $\mathcal{S}_0$), but then

$$Q_0^2(\nu) \supset \bigvee_{w \in \mathbb{F}_+^n} U_w P_0^2(\mu) \supset \bigvee_{w \in \mathbb{F}_+^n} \{U_w[L_1], \dots U_w[L_n]\} = \bigvee_{p \in \mathcal{A}_0} [p] = Q_0^2(\mu). \tag{3.21}$$

In other words, the vectors $[L_i]$ are cyclic for the row isometry $\mathbf{U}$, but these cyclic vectors are contained in $P_0^2(\mu)$.) This says that each containment is an equality, which gives minimality. □

The point of this proposition is that it will show, for the quasi-extreme states to be defined shortly, the GNS tuple $\mathbf{U}$ will be completely determined by $\mathbf{S}$ (as the minimal dilation of $\mathbf{S}$), and hence uniquely determined by μ (equivalently, b). We will revisit this remark following the proof of Theorem 5.1.

Theorem 3.5. *If μ has a tight extenstion, then it is unique (that is, if ν_1 and ν_2 are tight extensions of μ, then $\nu_1 = \nu_2$).*

Proof. Suppose ν is a tight extension of μ and let $w = i_1 \cdots i_m$ be a word. Let $\bar{w} = i_m i_1 \dots i_{m-1}$ (remove the last letter of w and append it at the beginning). As shorthand write $i = i_m$. Then

$$\nu(L_w) = \nu(L_i^* L_{\bar{w}} L_i) \tag{3.22}$$

$$= \langle U_{\bar{w}}[L_i], [L_i] \rangle_{Q_0^2(\nu)} \tag{3.23}$$

$$= \langle [L_i], U_{\bar{w}}^*[L_i] \rangle_{Q_0^2(\nu)} \tag{3.24}$$

$$= \langle [L_i], S_{\bar{w}}^*[L_i] \rangle_{P_0^2(\mu)} \tag{3.25}$$

which shows that $\nu(L_w)$ is completely determined by μ, and hence the extension is unique. □

Question 3.6. Does every state μ on $\mathcal{S}+\mathcal{S}^*$ have a tight extension to $\mathcal{A}+\mathcal{A}^*$?

It is rather frustrating that this question is still unanswered. Indeed, the proof of the foregoing theorem tells us what the extension must be, namely

$$\nu(L_w) := \langle [L_i], S^*_{\bar{w}}[L_i]\rangle_{P^2_0(\mu)} \tag{3.26}$$

The difficulty is in showing that this defines a *positive* linear functional.

We can now give a sufficient condition for the existence of a tight extension, in terms of the GNS space.

Definition 3.7. A state μ on $\mathcal{S}+\mathcal{S}^*$ will be called *quasi-extreme* if $P^2_0(\mu) = P^2(\mu)$.

Remark. The name "quasi-extreme" is chosen by analogy with the one-variable case. Indeed it is an easy consequence of the Szegő theorem that a function b is an extreme point of the unit ball of $H^\infty(\mathbb{D})$ if and only if for some (equivalently, all) $\alpha \in \mathbb{T}$, one has $P^2(\mu_\alpha) = L^2(\mu_\alpha)$. By a standard backward-shift argument, this latter condition is in turn equivalent to the equality $P^2_0(\mu) = P^2(\mu)$. So a state on $C(\mathbb{T})$ (that is, a probability measure on $\mathbb{T}$) is quasi-extreme by the above definition if and only if it is an AC measure for an extreme point of the ball of H^∞. We do not know if there is any relation between extreme points of the unit ball and quasi-extreme states in higher dimensions.

Theorem 3.8. *Every quasi-extreme state on $\mathcal{S}+\mathcal{S}^*$ has a unique extension to a state on $\mathcal{A}+\mathcal{A}^*$, and this extension is tight.*

Proof. The quasi-extremality assumption implies that the projection P_0 is the identity operator, hence $[I] \in P^2_0(\mu)$, but then if ν is *any* extension, we have $[I] \in Q^2_0(\nu)$, so $P_0[I] = [I] = Q_0[I]$. Thus by Theorem 3.3 ν is a tight extension, but then Theorem 3.5 gives that ν is unique. ☐

There is an operator-theoretic characterization of quasi-extremity, using the GNS tuple $\mathbf{S}$:

Lemma 3.9. *The state μ is quasi-extreme if and only if its GNS tuple $\mathbf{S} = (S_1, \dots S_d)$ is co-isometric.*

Proof. First assume μ is quasi-extreme. It suffices to show that

$$\sum_{j=1}^{d} \|S^*_j[p]\|^2 = \|[p]\|^2 \tag{3.27}$$

for all polynomials $p \in \mathcal{S}_0$, since by hypothesis these vectors are dense in $P^2(\mu)$. For this, first note that we can write

$$p(\mathbf{L}) = \sum_{j=1}^{d} L_j p_j(\mathbf{L}) \tag{3.28}$$

with $p_j \in \mathcal{A}_0$. Then by the orthogonality relations for the L_i,

$$\sum_{i=1}^{d} L_i L^*_i p(\mathbf{L}) = \sum_{i=1}^{d}\sum_{j=1}^{d} L_i L^*_i L_j p_j(\mathbf{L}) = \sum_{j=1}^{d} L_j p_j(\mathbf{L}) = p(\mathbf{L}). \tag{3.29}$$

Thus,

$$\sum_{j=1}^{d} \|S_j^*[p]\|^2 = \sum_{j=1}^{d} \mu(p(\mathbf{L})^* L_j L_j^* p(\mathbf{L})) = \mu(p(\mathbf{L})^* p(\mathbf{L})) = \|[p]\|^2. \tag{3.30}$$

For the converse, recall the proof of Proposition 3.1, which established (for any state μ and any $p \in \mathcal{S}_0$) the inequalities

$$\langle \sum_{i=1}^{n} S_i S_i^*[p], [p]\rangle_{P_0^2(\mu)} = \sum_{i=1}^{n} \langle P_0[L_i^* p], P_0[L_i^* p]\rangle_{P^2(\mu)} \tag{3.31}$$

$$\leq \sum_{i=1}^{n} \langle [L_i^* p], [L_i^* p]\rangle_{P^2(\mu)} \tag{3.32}$$

$$= \langle [p], [p]\rangle_{P_0^2(\mu)} \tag{3.33}$$

If $\mathbf{S}$ is coisometric, then equality holds in (3.32). Specializing to $p(\mathbf{L}) = L_1$, we have $\|P_0[I]\|_\mu^2 = \|[I]\|_\mu^2 = 1$, so $P_0[I] = [I]$ and hence μ is quasi-extreme. □

The fact that $\mathbf{S}$ is coisometric forces $\mathbf{U}$ to be a row unitary (that is, a system of Cuntz isometries):

Proposition 3.10. *If μ is a quasi-extreme state on $\mathcal{S} + \mathcal{S}^*$ then the GNS tuple $\mathbf{U}$ belonging to ν is a row unitary.*

Proof. Since ν is a tight extension of μ and μ is quasi-extreme, it follows $Q_0[I] = P_0[I] = [I]$, and thus $Q_0^2(\nu) = Q^2(\nu)$. Imitating the proof of Lemma 3.9 we see that the tuple $\mathbf{U}$ is coisometric; since it is already isometric by the GNS construction, it is unitary. □

We can now prove that in the quasi-extreme case the state μ has an honest GNS representation in terms of $\mathbf{S}$.

Proposition 3.11. *If μ is a quasi-extreme state on $\mathcal{S} + \mathcal{S}^*$, then μ is a vector state in the GNS representation, that is*

$$\mu(p(\mathbf{L})) = \langle p(\mathbf{S})[I], [I]\rangle \tag{3.34}$$

for all polynomials $p \in \mathcal{S}$. Moreover the GNS tuple $\mathbf{S}$ is cyclic, with cyclic vector $[I]$.

Proof. Let ν be the unique extension of μ to a state on $\mathcal{A} + \mathcal{A}^*$ coming from Theorem 3.8. Since the extension is tight, the restricted GNS tuple $\mathbf{V} = \mathbf{U}|_{Q_0^2(\nu)}$ for ν dilates $\mathbf{S}$. Fix a polynomial $p \in \mathcal{S}$. Then for any polynomial $q \in \mathcal{S}$, we have

$$\langle [q(\mathbf{L})], p(\mathbf{S})[I]\rangle_\mu = \langle p(\mathbf{S})^*[q(\mathbf{L})], [I]\rangle_\mu \tag{3.35}$$

$$= \langle p(V)^*[q(\mathbf{L})], [I]\rangle_\nu \tag{3.36}$$

$$= \nu(p^* q) \tag{3.37}$$

$$= \mu(p^* q) \tag{3.38}$$

$$= \langle [q(\mathbf{L})], [p(\mathbf{L})]\rangle. \tag{3.39}$$

Since this holds for all q, we conclude that $p(\mathbf{S})[I] = [p(\mathbf{L})]$. The identity (3.34) now follows by taking $q(\mathbf{L}) = [I]$. Since the $[p(\mathbf{L})]$ are dense in $P^2(\mu) = P_0^2(\mu)$ by definition, we also have that $[I]$ is a cyclic vector for $\mathbf{S}$. □

In the one-dimensional case, the theory of the de Branges-Rovnyak spaces $\mathcal{H}(b)$ often splits into the extreme and non-extreme cases. For example, b itself belongs to $\mathcal{H}(b)$ if and only if b is not extreme [23, IV-4, V-3]. It turns out that the notion of quasi-extreme introduced above is the correct one in this context.

Definition 3.12. A contractive multiplier b of H^2_d will be called *quasi-extreme* if and only if the state μ representing b as in (2.32) is quasi-extreme.

Theorem 3.13. *Let b be a contractive multiplier of H^2_d. Then $b \in \mathcal{H}(b)$ if and only if b is not quasi-extreme.*

Proof. Recall the normalized NC Fantappiè transform $\mathcal{V}_\mu$. Assume $b \in \mathcal{H}(b)$ with AC state μ. Then $\mathcal{V}_\mu(x) = b$ for some $x \in P^2(\mu)$, so $\mathcal{G}_\mu(x) = \frac{b}{1-b}$. But also $\mathcal{G}_\mu((1-\overline{b(0)})[I]) = \frac{1-b(0)^*b}{1-b}$, so it follows that

$$1 = \frac{1-b(0)^*b}{1-b} - (1-b(0)^*)\frac{b}{1-b} = (1-b(0)^*)\mathcal{G}_\mu([I]-x); \tag{3.40}$$

that is, the constant function 1 lies in the image of $P^2(\mu)$ under $\mathcal{G}_\mu$. By expanding C_z in a power series and putting $y = (1-\overline{b(0)})([I]-x) \in P^2(\mu)$, it follows from (3.40) and the definition of $\mathcal{G}_\mu$ that

$$1 = \mathcal{G}_\mu(y)(z) = \sum_{\mathbf{n}\in\mathbb{N}^d} z^{\mathbf{n}} \langle y, [L^{(\mathbf{n})}]\rangle_{P^2(\mu)}. \tag{3.41}$$

In other words, y is orthogonal in $P^2(\mu)$ to each symmetric monomial $L^{(\mathbf{n})}$ with $|\mathbf{n}| \geq 1$, so y is a nonzero vector orthogonal to $P^2_0(\mu)$, which means μ is not quasi-extreme. Conversely, the steps of this argument reverse to show that if b is not quasi-extreme (so that there is some nonzero $y \in P^2_0(\mu)^\perp \subset P^2(\mu)$), then 1 lies in the range of $\mathcal{G}_\mu$ and hence $b \in \mathcal{H}(b)$. ☐

It is worth noting that while the proof given here works in one variable, it is quite different from the proof in [23].

Corollary 3.14. *If b is quasi-extreme then so is αb for every unimodular $\alpha \in \mathbb{C}$.*

Proof. This is immediate from Theorem 3.13, since $\mathcal{H}(b) = \mathcal{H}(\alpha b)$. ☐

It also follows that the family of AC states $\{\mu_\alpha\}$ associated to a given b are either all quasi-extreme, or all not, a fact which was not obvious from the definition. Unfortunately, at present we do not know if there is any connection between being quasi-extreme, and being an extreme point of the set of contractive multipliers of H^2_d when $d > 1$ (as noted above these notions coincide when $d = 1$).

Question 3.15. If b is quasi-extreme, then is it an extreme point of the set of contractive multipliers of H^2_d? or vice-versa?

It would be very desirable to have some other characterization of the quasi-extreme multipliers when $d > 1$. A different characterization of the extreme b in one variable is the following: b is non-extreme if and only if $1 - |b|^2$ is log-integrable, which happens if only if there is an outer function $a \in H^\infty$ such that $|a|^2 + |b|^2 = 1$ on $\mathbb{T}$. This is in turn equivalent to saying that there is an a satisfying the operator identity

$$M_a^* M_a + M_b^* M_b = I. \tag{3.42}$$

However, this identity can never hold between multipliers of H^2_d when $d > 1$, unless a and b are both constant [11, Theorem 2.3].

3.2. **Examples.** At present we do not know any function-theoretic characterization of the quasi-extreme b when $d > 1$, but it is possible to give a few examples (and non-examples). As noted in the introduction, if μ is a positive measure on $\partial\mathbb{B}^d$, and b is given by the formula

$$\frac{1+b(z)}{1-b(z)} = \int_{\partial\mathbb{B}^d} \frac{1+z\zeta^*}{1-z\zeta^*}\, d\mu(\zeta) \tag{3.43}$$

then b is a contractive multiplier of H^2_d, though not every such b is representable in this form. Every such measure of course gives rise to a unique state $\widetilde{\mu}$ on $\mathcal{S}+\mathcal{S}^*$ representing b as in (2.32), and by comparing Taylor coefficients one finds that

$$\widetilde{\mu}(L^{(\mathbf{n})}) = \int_{\partial\mathbb{B}^d} \zeta^{\mathbf{n}}\, d\mu(\zeta). \tag{3.44}$$

In particular if we take μ to be the point mass at a fixed $\zeta \in \partial\mathbb{B}^d$, the resulting state on $\mathcal{S}+\mathcal{S}^*$ is called the *Cuntz state* ω_ζ. The corresponding b is $b(z) = \langle z, \zeta\rangle$ and it is easy to see this b is quasi-extreme, since $[I] = \left[\sum_{j=1}^d \zeta_j^* L_j\right]$ in $P^2_0(\omega_\zeta)$. (Indeed $\|[I] - [\sum \zeta_j^* L_j]\|^2_{\omega_\zeta} = 2 - 2\mathrm{Re}\,\omega_\zeta(\sum \zeta_j^* L_j) = 0$.) We will see later that all of the $\mathcal{H}(b)$ spaces are infinite dimensional, which gives another indication that the classical measure μ is inadequate for our purposes—in this example, $L^2(\mu)$ is of course one-dimensional so there can be no identification of $L^2(\mu)$ with $\mathcal{H}(b)$.

If in the above construction we take μ to be a measure supported on the circle $z_2 = \cdots z_d = 0$, then the resulting b is a function of z_1 alone, and any $b(z) = b(z_1)$ can equal any function in the unit ball of $H^\infty(\mathbb{D})$. In this case b will be quasi-extreme if and only if $b(z_1)$ is an extreme point of the unit ball of H^∞.

A more sophisticated example, in this case for $d = 2$, comes by considering the state $\mu = \frac{1}{2}(\omega_{e_1} + \omega_{e2})$ on $\mathcal{S}+\mathcal{S}^*$. The resulting b is

$$b(z_1, z_2) = \frac{z_1 + z_2 - z_1 z_2}{2 - z_1 - z_2} \tag{3.45}$$

It is now less obvious, but this b is quasi-extreme; this follows from the fact that for the polynomial

$$p(\mathbf{L}) = \frac{1}{\sqrt{6}}\left(\sum_j L_j - \sum_{j,k} L_j L_k + \sum_{j,k,l} L_j L_k L_l\right) \tag{3.46}$$

one may verify that $\|[I] - [p(\mathbf{L})]\|^2_\mu = 0$.

In the other direction, if $b_1, \ldots b_d$ are functions in $H^\infty(\mathbb{D})$ and each is not an extreme point, then the product

$$b(z) = b_1(z_1)\cdots b_d(z_d) \tag{3.47}$$

is not quasi-extreme.

4. Canonical functional models and the Gleason problem in $\mathcal{H}(b)$

The goal of this section is to establish the uniqueness of the contractive solution to the Gleason problem in $\mathcal{H}(b)$ when b is quasi-extreme, and study some of its properties. In the next section we will show that this solution admits rank-one coisometric perturbations. If f is a holomorphic function in $\mathbb{B}^d$, we say that a d-tuple of holomorphic functions $f_1, \dots f_d$ *solves the Gleason problem for f* if

$$f(z) - f(0) = \sum_{j=1}^{d} z_j f_j(z). \tag{4.1}$$

Similarly, a d-tuple of linear operators $A_1, \dots A_d$ is said to solve the Gleason problem in a holomorphic space H if

$$f(z) - f(0) = \sum_{j=1}^{d} z_j (A_j f)(z) \tag{4.2}$$

for all $f \in H$.

Notice that it one variable, it is trivial that the Gleason problem for f has a unique solution, given by the backward shift $f \to (f(z) - f(0))/z$. Likewise the backward shift is the only operator solving the Gleason problem in a holomorphic space H, so questions about it focus on boundedness, etc. In contrast, in the multivariable setting solutions to the Gleason problem for a given f are never unique, so the goal is to establish existence (and perhaps uniqueness) of solutions satisfying some additional conditions, typically membership in some space of functions. It was proved by Ball and Bolotnikov [2] that contractive solutions to the Gleason problem in $\mathcal{H}(b)$ always exist. In this section we study some of these solutions in more detail. We prove that every such solution can be split into a sum of two operators; these being a rank-one operator and the adjoint of a multiplication operator (each is possibly unbounded). This structure result will be applied to obtain a Clark-type theorem on rank-one perturbations, and to characterize the z-invariant $\mathcal{H}(b)$ spaces.

4.1. Functional models.

In this subsection we recall a result of Ball and Bolotnikov [2] on solutions to the Gleason problem in the $\mathcal{H}(b)$ spaces. We begin with their definition of a *canonical functional model realization*.

Definition 4.1. Given a multiplier b, say that the block operator matrix

$$\mathbf{U} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} : \mathcal{H}(b) \oplus \mathbb{C} \to \mathcal{H}(b)^d \oplus \mathbb{C} \tag{4.3}$$

is a *canonical functional model realization* for b if the following conditions are satisfied:

(1) $\mathbf{U}$ is contractive,
(2) the d-tuple $A : \mathcal{H}(b) \to \mathcal{H}(b)^n$ solves the Gleason problem for $\mathcal{H}(b)$,
(3) $B : \mathbb{C} \to \mathcal{H}(b)^n$ solves the Gleason problem for b,
(4) the operators $C : \mathcal{H}(b) \to \mathbb{C}$ and $D : \mathbb{C} \to \mathbb{C}$ are given by

$$C : f \to f(0), \qquad D : \lambda \to b(0)\lambda \tag{4.4}$$

respectively, for all $f \in \mathcal{H}(b)$ and all $\lambda \in \mathbb{C}$.

To say that $\mathbf{U}$ is a *realization* of b means that for all $z \in \mathbb{B}^n$

$$b(z) = D + C\left(I - \sum_{j=1}^{d} z_j A_j\right)^{-1}\left(\sum_{j=1}^{d} z_j b_j\right). \tag{4.5}$$

where we have written B as a column vector $(b_1, \dots b_d)^T$ with $b_j \in \mathcal{H}(b)$. The fact that $\mathbf{U}$ is contractive then entails

$$B^*B \leq 1 - D^*D, \tag{4.6}$$

or

$$\sum_{j=1}^{n} \|b_j\|^2_{\mathcal{H}(b)} \leq 1 - |b(0)|^2. \tag{4.7}$$

Moreover, since C can be expressed as $C : f \to \langle f, k_0^b\rangle_{\mathcal{H}(b)}$, we can write $C^*C = k_0^b \otimes k_0^b$, and contractivity also entails

$$A^*A \leq I_{\mathcal{H}(b)} - C^*C, \tag{4.8}$$

or

$$\sum_{j=1}^{d} A_j^* A_j \leq I - k_0^b \otimes k_0^b. \tag{4.9}$$

Following [2], when (4.9 holds we say A is a *contractive solution to the Gleason problem in* $\mathcal{H}(b)$. One of the main results of [2] is

Theorem 4.2. *For each contractive multiplier b, there exists a canonical functional model realization $\mathbf{U}$. In particular, for each b there exists a contractive solution to the Gleason problem in $\mathcal{H}(b)$, and there exist functions $b_1, \dots b_d \in \mathcal{H}(b)$ satisfying*

(i) $b(z) - b(0) = \sum_{j=1}^{d} z_j b_j(z)$, *and*

(ii) $\sum_{j=1}^{d} \|b_j\|^2_{\mathcal{H}(b)} \leq 1 - |b(0)|^2$

The functions b_j play the same role in the present development as S^*b does in the one-variable case; in particular (ii) is the "inequality for difference quotients" in this setting.

We record one more result from [2], namely that the reproducing kernel for $\mathcal{H}(b)$ can be expressed in terms of a functional model. In particular we have for all $f \in \mathcal{H}(b)$

$$f(z) = C(I - z\mathbf{A})^{-1} f \tag{4.10}$$

or more explicitly

$$f(z) = \langle (I - z\mathbf{A})^{-1} f, k_0^b \tag{4.11}$$

and thus

$$k_z^b = (I - \mathbf{A}^* z^*)^{-1} k_0^b. \tag{4.12}$$

4.2. **Unique contractive solutions in the quasi-extreme case.**

Definition 4.3. A contractive solution $\mathbf{X} = (X_1, \dots X_d)$ to the Gleason problem in $\mathcal{H}(b)$ will be called *extremal* if

$$\sum_{j=1}^{d} X_j^* X_j = I - k_0^b \otimes k_0^b. \tag{4.13}$$

(That is, equality holds in (4.9).) The next theorem characterizes contractive and extremal solutions by their action on reproducing kernels. To avoid trivialities we assume b is non-constant.

Theorem 4.4. *A d-tuple $\mathbf{X} = (X_1, \dots X_d)$ is a contractive solution to the Gleason problem in $\mathcal{H}(b)$ if and only if it acts on the reproducing kernels k_w^b by the formula*

$$(X_j k_w^b)(z) = w_j^* k_w^b(z) - b_j(z) b(w)^*. \tag{4.14}$$

where $b_1, \dots b_d$ are functions in $\mathcal{H}(b)$ satisfying

(i) $b(z) - b(0) = \sum_{j=1}^{d} z_j b_j(z)$, *and*

(ii) $\sum_{j=1}^{d} \|b_j\|^2_{\mathcal{H}(b)} \leq 1 - |b(0)|^2$

The solution is extremal if and only if equality holds in (ii).

Proof. First, suppose b_j exist satisfying (i) and (ii), and the X_j are defined by (4.14). Then for all $z, w \in \mathbb{B}^d$,

$$\sum_{j=1}^{d} z_j (X_j k_w^b)(z) = \sum_{j=1}^{d} z_j \left(w_j^* k_w^b(z) - z_j b_j(z) b(w)^* \right) \tag{4.15}$$

$$= zw^* \frac{1 - b(z)b(w)^*}{1 - zw^*} - (b(z) - b(0)) b(w)^* \tag{4.16}$$

$$= zw^* \frac{1 - b(z)b(w)^*}{1 - zw^*} + (1 - b(z)b(w)^*) - (1 - b(0)b(w)^*) \tag{4.17}$$

$$= k_w^b(z) - k_w^b(0) \tag{4.18}$$

where we have used property (i) of the b_j. Thus the X_j solve the Gleason problem on the linear span of the k_w^b. Once we show that $\sum X_j^* X_j \leq I - k_0^b \otimes k_0^b$ on this span, it follows that the X_j are bounded and have unique bounded extensions to $\mathcal{H}(b)$; a routine approximation argument shows that these extensions solve the Gleason problem for all of $\mathcal{H}(b)$. So, compute:

$$\langle X_j k_w^b, X_j k_z^b \rangle = \langle w_j^* k_w^b - b_j b(w)^*, z_j^* k_z^b - b_j b(z)^* \rangle \tag{4.19}$$

$$= z_j w_j^* k(z, w) - z_j b_j(z) b(w)^* - w_j^* b_j(w)^* b(z) + \|b_j\|^2 b(z) b(w)^*. \tag{4.20}$$

Summing over j, and using properties (i) and (ii) of the b_j, we have

(4.21)
$$\langle\sum_{j=1}^d X_j^*X_jk_w^b, k_z^b\rangle \leq zw^*k(z,w) - (b(z)-b(0))b(w)^* - b(z)(\overline{b(w)-b(0)}) + (1-|b(0)|^2)b(z)b(w)^*$$

(4.22)
$$= zw^*k(z,w) + (1-b(z)b(w)^*) - (1-b(z)b(0)^*)(1-b(0)b(w)^*)$$

(4.23)
$$= \langle k_w^b, k_z^b\rangle - \langle k_w^b, k_0^b\rangle\langle k_0^b, k_z^b\rangle$$

(4.24)
$$= \langle (I - k_0^b\otimes k_0^b)k_w^b, k_z^b\rangle$$

This shows $\sum X_j^*X_j \leq I - k_0^b\otimes k_0^b$, with equality if and only if equality holds in (4.21), if and only if equality holds in (ii).

Conversely, suppose X_j are operators on $\mathcal{H}(b)$ which solve the Gleason problem and satisfy $\sum X_j^*X_j \leq I - k_0^b\otimes k_0^b$. For each $j=1,\dots d$ and each $w\in\mathbb{B}^d$, define a function $f_{j,w}\in\mathcal{H}(b)$ by the formula

(4.25)
$$f_{j,w}(z) = w_j^*k_w^b(z) - (X_jk_w^b)(z).$$

We must show $f_{j,w} = b_jb(w)^*$ for some b_j satisfying (i) and (ii). A computation similar to the verification in the first part of the proof shows that, since the X_j are assumed to solve the Gleason problem, we must have

(4.26)
$$\sum_{j=1}^d z_jf_{j,w}(z) = (b(z)-b(0))b(w)^*.$$

Using this identity, and again imitating the algebra in the first part of the proof, the hypothesis $\sum X_j^*X_j \leq I - k_0^b\otimes k_0^b$ entails the kernel inequalities

(4.27)
$$k(z,w) - k_0^b(z)k_0^b(w)^* \geq \sum_{j=1}^d\langle X_jk_w^b, X_jk_z^b\rangle$$

(4.28)
$$= zw^*k(z,w) - 2b(z)b(w)^* + \sum_{j=1}^d\langle f_{j,w}, f_{j,z}\rangle.$$

This simplifies to

(4.29)
$$\sum_{j=1}^d\langle f_{j,w}, f_{j,z}\rangle \leq (1-|b(0)|^2)b(z)b(w)^*.$$

This inequality implies, via Douglas's factorization lemma, that there is a contractive linear map from $\mathbb{C}$ to the direct sum of d copies of $\mathcal{H}(b)$ taking $b(w)^*$ to the column vector whose j^{th} entry is $(1-|b(0)|^2)^{-1/2}f_{j,w}$. Such a map must send the scalar 1 to a vector in the unit ball of $\mathcal{H}(b)^d$. If we write a_j for the j^{th} entry of this vector, then $\sum\|a_j\|^2\leq 1$, and we have

(4.30)
$$(1-|b(0|^2)^{-1/2}f_{j,w} = a_jb(w)^*.$$

Rescale: put $b_j = (1-|b(0)|^2)^{1/2}a_j$; then $\sum\|b_j\|^2 \leq 1-|b(0)|^2$ and

(4.31)
$$f_{j,w} = b_jb(w)^*.$$

But now from (4.26) we have for all $z, w \in \mathbb{B}^d$

$$\sum_{j=1}^{d} z_j b_j(z) b(w)^* = (b(z) - b(0)) b(w)^*. \tag{4.32}$$

Since b is not identically 0, we conclude $\sum z_j b_j(z) = b(z) - b(0)$, so the b_j satisfy (i) and (ii), and from (4.25), the X_j have the claimed form. In the case of equality $\sum_{j=1}^d X_j^* X_j = I - k_0^b \otimes k_0^b$,, we have also equality in (4.29), and it is a straightforward matter to verify that this propagates through the calculation to give equality in (ii) (in this case the contractive linear map which produces the b_j is isometric). □

The above theorem also lets us obtain a formula for the action of the X_j^* on arbitrary elements of $\mathcal{H}(b)$. Indeed using Theorem 4.4 and the relation

$$X_j^* f(z) = \langle X^* f, k_z^b \rangle = \langle f, X_j k_z^b \rangle$$

we have for all $f \in \mathcal{H}(b)$

$$X_j^* f(z) = z_j f(z) - \langle f, b_j \rangle_{\mathcal{H}(b)} b(z). \tag{4.33}$$

This makes the next corollary almost immediate.

Corollary 4.5. *Let b be a contractive multiplier of $\mathcal{H}_d^2$. Then the following are equivalent:*

i) $z_j \mathcal{H}(b) \subset \mathcal{H}(b)$ *for all* $j = 1, \dots d$
ii) $b \in \mathcal{H}(b)$
iii) b *is not quasi-extreme.*

Proof. The equivalence of (ii) and (iii) is Theorem 3.13. For the equivalence of (i) and (ii), first recall that by Theorem 4.2, for any b there exists a contractive solution to the Gleason problem $(X_1, \dots X_d)$. From the formula (4.33 we see that if $b \in \mathcal{H}(b)$, then $z_j f \in \mathcal{H}(b)$ for all $f \in \mathcal{H}(b)$ and all j. Conversely, if $\mathcal{H}(b)$ is z_j-invariant for each j, since b is non-constant we can choose j such that $b_j \neq 0$. Then specializing to $f = b_j$ we have

$$\|b_j\|^2 b(z) = z_j b_j(z) - (X_j^* b_j)(z). \tag{4.34}$$

and the right side lies in $\mathcal{H}(b)$ by hypothesis. □

Remark: Note that it is possible that $\mathcal{H}(b)$ is z_j-invariant for some j but not others. A simple example in two variables is $b(z_1, z_2) = z_1$: then we can take $b_1(z) = 1, b_2(z) = 0$. This b is associated to the Cuntz state ω_{e_1} and is quasi-extreme (Example 3.2). However from (4.33) we see that $\mathcal{H}(b)$ is invariant for z_2 but not z_1.

We also observe that once $b \in \mathcal{H}(b)$, then $\mathcal{H}(b)$ also contains the constant functions, and therefore, by z-invariance, all polynomials. In one variable the polynomials are dense in $\mathcal{H}(b)$ when b is non-extreme, but so far we have been unable to prove this when $d > 1$ (though it seems very likely to be true).

Question 4.6. If b is non-extreme, are the polynomials dense in $\mathcal{H}(b)$?

We are now in a position to prove the uniqueness of the contractive solution (which will in fact be extremal) to the Gleason problem when b is quasi-extreme. By Theorem 4.4, it suffices to produce the functions b_j. The next lemma shows how to do this, starting from the AC state for b. We make use of the normalized NC Fantappiè transform $\mathcal{V}_\mu$ and the GNS tuple $\mathbf{S}$ associated to the state μ.

Lemma 4.7. *Let b be a quasi-extremal multiplier, with AC state μ. Then the functions*

$$b_j := (1 - b(0)^*)\mathcal{V}_\mu(S_j^*[I]) \tag{4.35}$$

belong to $\mathcal{H}(b)$, satisfy $\sum_j \|b_j\|^2_{\mathcal{H}(b)} = 1 - |b(0)|^2$, and solve the Gleason problem for b; that is

$$b(z) - b(0) = \sum_{j=1}^{d} z_j b_j(z). \tag{4.36}$$

Proof. From (2.46) we have

$$\mu((I - z\mathbf{L}^*)^{-1})(z) = \frac{1}{1 - b(0)^*}\frac{1 - b(0)^* b(z)}{1 - b(z)} \tag{4.37}$$

and

$$\mu(I) = \frac{1 - |b(0)|^2}{|1 - b(0)|^2}. \tag{4.38}$$

It follows that

$$\mu(\sum_{k=1}^{\infty}(z\mathbf{L}^*)^k) = \mu((I - z\mathbf{L}^*)^{-1}) - \mu(I)) \tag{4.39}$$

$$= \frac{1}{1 - b(0)^*}\frac{b(z) - b(0)}{1 - b(z)} \tag{4.40}$$

By the assumption that b is quasi-extreme, there is a sequence of polynomials $p_m \in \mathcal{S}_0$ such that $[p_m] \to [I]$ in $P^2(\mu)$. Then for each integer $n \geq 1$,

$$\begin{aligned}
\mu((z\mathbf{L}^*)^n) &= \langle [I], [(z\mathbf{L}^*)^n] \rangle_H \\
&= \lim_{m\to\infty} \langle [p_m], [(z\mathbf{L}^*)^n] \rangle \\
&= \lim_{m\to\infty} \mu((z\mathbf{L}^*)^n p_m(L)) \\
&= \lim \sum_{j=1}^{d} z_j \mu((z\mathbf{L}^*)^{n-1} L_j^* p_m(L)) \\
&= \lim \sum_{j=1}^{d} z_j \langle S_j^*[p_m], (z\mathbf{L}^*)^{n-1} \rangle \\
&= \sum_{j=1}^{d} z_j \langle S_j^*[I], (z\mathbf{L}^*)^{n-1} \rangle
\end{aligned}$$

Summing from $n = 1$ and multiplying by $(1 - b)$, we obtain

$$\frac{b(z) - b(0)}{1 - b(0)^*} = \sum_{j=1}^{d} z_j (1 - b(z))\mathcal{G}_\mu(S_j^*[I]) \tag{4.41}$$

$$= \sum_{j=1}^{d} z_j \mathcal{V}_\mu(S_j^*[I]). \tag{4.42}$$

Multiplying this by $1 - b(0)^*$ and applying (4.40) and the definition of b_j, we have

$$b(z) - b(0) = \sum_{j=1}^{d} z_j b_j(z), \tag{4.43}$$

so the functions b_j solve the Gleason problem as claimed. They belong to $\mathcal{H}(b)$ since they lie in the range of $\mathcal{V}_\mu$. For the norm computation we have

$$\sum_{j=1}^{d} \|S_j^*[I]\|^2 = \|[I]\|^2 = \mu(I) = \frac{1 - |b(0)|^2}{|1 - b(0)|^2}. \tag{4.44}$$

by Lemma 3.9 and equation (4.38). Thus using the definition of b_j and the fact that $\mathcal{V}_\mu$ is unitary,

$$\begin{aligned}\sum_{j=1}^{d} \|b_j\|^2 &= |1 - b(0)|^2 \sum_{j=1}^{d} \|\mathcal{V}_\mu(S_j^*[I])\|^2 \\ &= |1 - b(0)|^2 \sum_{j=1}^{d} \|S_j^*[I]\|^2 \\ &= 1 - |b(0)|^2.\end{aligned}$$

□

Theorem 4.8. *If b is quasi-extreme, then there is a unique contractive solution to the Gleason problem in $\mathcal{H}(b)$, and this solution is extremal.*

Proof. Define

$$X_j k_w^b = w_j^* k_w^b - b(w)^* b_j \tag{4.45}$$

where the b_j are chosen as in Lemma 4.7. It is then immediate from Theorem 4.4 that $\mathbf{X} = (X_1, \ldots X_j)$ is an extremal solution to the Gleason problem in $\mathcal{H}(b)$.

Uniqueness will be proved by contradiction. Suppose there are two Gleason tuples $\mathbf{X}, \widetilde{\mathbf{X}}$. These must be defined as in equation (4.14), for functions b_j and $\widetilde{b}_j$ satisfying conditions (i) adn (ii) of Theorem 4.4. But then for each j the densely defined operator

$$(X_j - \widetilde{X}_j) k_w^b = (b_j - \widetilde{b}_j) b(w)^* \tag{4.46}$$

is bounded on $\mathcal{H}(b)$, and is nonzero for some j. Fix such a j; put $g = b_j - \widetilde{b}_j$. So $k_w^b \to b(w)^* g$ is a bounded rank-one operator, which means that $k_w^b \to b(w)^*$ extends to a bounded linear functional on $\mathcal{H}(b)$. Then there is an $h \in \mathcal{H}(b)$ with

$$b(w)^* = \langle k_w^b, h \rangle h(w)^* \tag{4.47}$$

and it follows that $h = b$, so $b \in \mathcal{H}(b)$. Since b was assumed quasi-extreme, this contradicts Theorem 3.13. □

5. Rank-one perturbations and intertwining

The goal of this section is to prove Theorem 5.1, which is the analog of the one-variable Theorem 1.2. Fix a quasi-extreme multiplier b with its family of AC states $\{\mu_\alpha\}$. To unclutter the notation we will write $\mathcal{V}_\alpha$ for the Fantappie transform $\mathcal{V}_{\mu_\alpha}$. As before, $\mathbf{X}$ denotes the unique solution to the Gleason problem in $\mathcal{H}(b)$, and we write $\mathbf{S}^\alpha = (S_1^\alpha, \dots S_d^\alpha)$ for the co-isometric GNS tuple acting on the GNS space $P^2(\mu_\alpha)$.

Theorem 5.1. *Let b be a quasi-extreme multiplier of H^2_d. Then the rank-one perturbation of $\mathbf{X}$ defined by*

$$X_j + \alpha^*(1-\alpha^* b(0))^{-1} b_j \otimes k_0^b \tag{5.1}$$

is cyclic, isometric, and unitarily equivalent to $\mathbf{S}^{\alpha*}$ under the normalized Fantappiè transform $\mathcal{V}_\alpha$:

$$\mathcal{V}_\alpha S_j^{\alpha*} = (X_j + \alpha^*(1-\alpha^* b(0))^{-1} b_j \otimes k_0^b)\mathcal{V}_\alpha. \tag{5.2}$$

Moreover, if ν_α is the unique extension of μ_α to $\mathcal{A}+\mathcal{A}^*$, then the GNS construction applied to ν_α produces a Cuntz tuple $\mathbf{U}^\alpha$, which is unitarily equivalent to the minimal isometric dilation of $\mathbf{S}^\alpha$.

Proof. Since we already know $\mathbf{S}$ is cyclic and coisometric (Lemma 3.9 and Proposition 3.11), everything follows once we prove the intertwining property; and in fact the intertwining holds even when b is not quasi-extreme.

To prove the intertwining relation, recall from the proof of Theorem 2.8 that the NC kernel functions

$$G_w^\alpha = (1-\alpha b(w)^*)[(1-\mathbf{L}w^*)^{-1}] \tag{5.3}$$

are dense in $P^2(\mu_\alpha)$, and $\mathcal{V}_\alpha$ takes G_w^α onto the reproducing kernel k_w^b of $\mathcal{H}(b)$. To compute the action of $S_j^{\alpha*}$ on G_w^α, we first have for integers $n \geq 1$

$$S_j^{\alpha*}[(\mathbf{L}w^*)^n] = [L_j^*(\mathbf{L}w^*)^n] \tag{5.4}$$

$$= w_j^*[(\mathbf{L}w^*)^{n-1}] \tag{5.5}$$

and, when $n=0$, from the definition of b_j in Lemma 4.7

$$S_j^{\alpha*}[I] = \alpha^*(1-\alpha^* b(0))^{-1}\mathcal{V}_\alpha^{-1} b_j. \tag{5.6}$$

Summing over n, we obtain

$$S_j^{\alpha*} G_w^\alpha = (1-\alpha b(w)^*)\sum_{n=0}^\infty S_j^{\alpha*}[(\mathbf{L}w^*)^n] \tag{5.7}$$

$$= (1-\alpha b(w)^*)\left(\alpha^*(1-\alpha^* b(0))^{-1}\mathcal{V}_\alpha^{-1} b_j + w_j^* \sum_{k=0}^\infty [(\mathbf{L}w^*)^n]\right) \tag{5.8}$$

and therefore

$$\mathcal{V}_\alpha S_j^{\alpha*} G_w^\alpha = \alpha^*(1-\alpha^* b(0))^{-1}(1-\alpha b(w)^*) b_j + w_j^* k_w^b \tag{5.9}$$

On the other hand, by the definition of X_j,

(5.10)

$$(X_j + \alpha^*(1-\alpha^* b(0))^{-1} b_j \otimes k_0^b)\mathcal{V}_\alpha G_w^\alpha = (X_j + \alpha^*(1-\alpha^* b(0))^{-1} b_j \otimes k_0^b) k_w^b$$

$$= w_j^* k_w^b - b_j b(w)^* + \alpha^*(1-\alpha^* b(0))^{-1}(1 - b(0)b(w)^*) b_j \quad (5.11)$$

$$= w_j^* k_w^b + \alpha^*(1-\alpha^* b(0))^{-1}(1 - \alpha b(w)^*) b_j \quad (5.12)$$

which agrees with (5.9).

Finally, the claims about the Cuntz tuple $\mathbf{U}^\alpha$ follow from the fact that μ is quasi-extreme and Proposition 3.10. □

Let us recapitulate the relationship between the function b, the state μ on $\mathcal{S}+\mathcal{S}^*$ representing b, and the operator tuples $\mathbf{S}^\alpha$, $\mathbf{U}^\alpha$. Starting with b one obtains the AC states μ_α via the NC Herglotz representation. Since b is quasi-extreme, each μ_α is quasi-extreme and determines a coisometric tuple $\mathbf{S}^\alpha$. This $\mathbf{S}^\alpha$ has a minimal row-unitary dilation $\mathbf{U}^\alpha$. On the other hand, μ_α has a unique extension to a positive functional ν_α on the full Cuntz-Toeplitz operators system $\mathcal{A}+\mathcal{A}^*$. Applying the GNS construction to ν_α gives $\mathbf{U}^\alpha$ again. In this sense we think of ν_α as the "spectral measure" of the row unitary $\mathbf{U}^\alpha$. Moreover, a suitable rank-one perturbation of $\mathbf{S}^{\alpha *}$ is unitarily equivalent, via the NC Fantappie transform $\mathcal{V}_\alpha$, to the unique contractive solution to the Gleason problem in $\mathcal{H}(b)$.

The only difference between this picture and the one-variable situation is, of course, that there is no distinction between $\mathcal{S}+\mathcal{S}^*$ and $\mathcal{A}+\mathcal{A}^*$; they are both just (dense subspaces of) $C(\mathbb{T})$, and $\mathbf{S}$ and $\mathbf{U}$ are both just the unitary operator M_ζ acting on $P^2(\mu)=L^2(\mu)$.

A natural question which arises at this point is: which unitaries $\mathbf{U}$ can arise by this construction? In one variable the answer is simple: every cyclic unitary operator. In the present setting, the answer is somewhat more delicate, in that the row unitary $\mathbf{U}$ must not only be cyclic (thus determining a "spectral measure" ν), but $\mathbf{U}$ must also be the minimal dilation of its compression to the subspace $P^2(\mu)\subset Q^2(\nu)$. This will be explored further in a separate paper examining the characteristic functions associated to rank-one perturbations of $\mathbf{S}$ and $\mathbf{U}$.

6. Spectral results

Finally, we examine the spectra of the solutions $\mathbf{X}$ to the Gleason problem and the GNS tuples $\mathbf{S}$. We begin with some preliminaries on angular derivaties in the ball, in particular for multipliers of H^2_d.

Say that a point $\zeta\in\partial\mathbb{B}^d$ is a *C-point* for b if

$$\liminf_{z\to\zeta} \frac{1-|b(z)|^2}{1-|z|^2} = L < \infty. \quad (6.1)$$

By [, Theorems...], ζ is a C-point for b if and only if b and its directional derivative $D_\zeta b$ both have finite limits as $z\to\zeta$ non-tangentially, with $\lim_{z\to\zeta}|b(z)|=1$ and $\lim_{z\to\zeta} D_\zeta b(z)>0$ (briefly, b has a *finite angular derivative* at ζ). The general theory of angular derivatives for functions in the ball may be found in [21, Chapter 8]. However when the function b is a contractive multiplier of H^2_d, a somewhat stronger theorem is available (see [12]). In particular there is a connection between angular derivatives of b and the $\mathcal{H}(b)$ spaces which closely parallels the one-dimensional results of Sarason [23, Chapter VI].

We summarize the results needed from [12] in the following theorem:

Theorem 6.1. *Let b be a contractive multiplier of the Drury-Arveson space H^2_d and let $\zeta \in \partial\mathbb{B}^d$. The following are equivalent:*

i) *ζ is a C-point for b*

ii) *there exists $\alpha \in \mathbb{T}$ such that the function*

$$k^b_\zeta(z) := \frac{1 - b(z)\alpha^*}{1 - z\zeta^*} \tag{6.2}$$

belongs to $\mathcal{H}(b)$.

When these occur, b has nontangential limit α at ζ, and additionally every $f \in \mathcal{H}(b)$ has a finite nontangential limit at ζ, equal to $\langle f, k^b_\zeta\rangle_{\mathcal{H}(b)}$. Moreover we have $\|k^b_\zeta\|^2 = L$.

It what follows we will abuse the notation slightly and write S^*_j for the rank-one perturbation of X_j in (5.1).

Theorem 6.2. *Let b be quasi-extreme with AC states $\{\mu_\alpha\}_{\alpha\in\mathbb{T}}$ and $\mathbf{S}^\alpha$ the Clark tuple for μ_α. For fixed $\zeta \in \partial\mathbb{B}^d$, the eigenvalue problem*

$$\sum_{j=1}^d \zeta^*_j S^\alpha_j h = h \tag{6.3}$$

has a solution in $\mathcal{H}(b)$ if and only if b has finite angular derivative at ζ and $b(\zeta) = \alpha$, in which case the eigenspace is one-dimensional and spanned by k^b_ζ.

Proof. First assume (6.3) has a nonzero solution $h \in \mathcal{H}(b)$. Write $\widetilde{b}(z) = \sum_{j=1}^d \zeta_j b_j(z)$. Then using (5.1) and (4.33) to compute S^α_j, we have

$$h(z) = \left(\sum_{j=1}^d \zeta^*_j S^\alpha_j h\right)(z) \tag{6.4}$$

$$= z\zeta^* h(z) - \langle h, \widetilde{b}\rangle_{\mathcal{H}(b)} b(z) + \alpha\langle h, \widetilde{b}\rangle_{\mathcal{H}(b)} k^b_0 \tag{6.5}$$

and solving for h we find

$$h(z) = \frac{\alpha\langle h, \widetilde{b}\rangle}{1 - b(0)^*\alpha}\frac{1 - b(z)\alpha^*}{1 - \langle z, \zeta\rangle} = ck^b_\zeta(z) \tag{6.6}$$

for some nonzero c. Thus by Theorem 6.1, b has an angular derivative at ζ with $b(\zeta) = \alpha$. Conversely, suppose the angular derivative condition holds at ζ, with $b(\zeta) = \alpha$. Then by Theorem 6.1 the function k^b_ζ lies in $\mathcal{H}(b)$. Note also that by the reproducing property of k^b_ζ at ζ, we have

$$\langle k^b_\zeta, \widetilde{b}\rangle = \sum_{j=1}^d \zeta^*_j b_j(\zeta)^* = b(\zeta)^* - b(0)^* = \alpha^* - b(0)^*. \tag{6.7}$$

With this in hand, repeating the calculation in the first part of the proof shows that

$$\sum_{j=1}^d \zeta^*_j S^\alpha_j k^b_\zeta = k^b_\zeta. \tag{6.8}$$

□

Finally, we include a result on the essential Taylor spectrum of $\mathbf{X}$. For this result we do not need to assume b is quasi-extreme, and $\mathbf{X}$ can be any contractive solution to the Gleason problem in $\mathcal{H}(b)$. First let us note that while the operators X_j do not commute, we see from Theorem 4.4 that the commutators $[X_i, X_j]$ have finite rank. Thus if we let π denote the quotient map to the Calkin algebra, then the $\pi(X_j)$ form a commuting row contraction, and it then makes sense to talk about its Taylor spectrum. It turns out that we do not need the definition of the Taylor spectrum in the proof of the next theorem, only the fact that the spectral mapping theorem holds for it (and even this we need only for polynomial mappings; which means that Theorem 6.3 is valid for the Harte spectrum as well). That is, if $\sigma(T_1, \dots T_d)$ denotes the Taylor spectrum of a commuting d-tuple of operators $T_1, \dots T_d$, then for any analytic polynomial p in d variables we have

$$p(\sigma(T_1, \dots T_d)) = \sigma(p(T_1, \dots T_d)). \tag{6.9}$$

Theorem 6.3. *Let X be a contractive solution to the Gleason problem in $\mathcal{H}(b)$. Then the Taylor spectrum of $\pi(X)$ contains the unit sphere $\partial\mathbb{B}^d$.*

In one variable, Sarason proves in [23, Theorem V-8] that an open arc $I \subset \mathbb{T}$ lies in the resolvent set of X^* if and only if every function in $\mathcal{H}(b)$ can be analytically continued across I. In higher dimensions, our result says that this is still true, though in a vacuous way: the spectrum of $\pi(\mathbf{X})$ contains the entire sphere, and it will turn out that there is no open set of $\partial\mathbb{B}^d$ across which all $f \in \mathcal{H}(b)$ can be continued.

We begin with two lemmas; it is the second lemma that does most of the work.

Lemma 6.4. *A point $\zeta \in \mathbb{B}^d$ belongs to the Taylor spectrum of $(T_1, \dots T_d)$ if and only if $(I - \mathbf{T}\zeta^*)$ is not invertible.*

Proof. This follows immediately from the spectral mapping property (6.9) applied to T and the polynomial $p(z) = 1 - z\zeta^*$. □

Lemma 6.5. *Let $\mathbf{X}$ be any contractive solution to the Gleason problem in $\mathcal{H}(b)$ and let $\zeta \in \mathbb{B}^d$. If $I - \zeta\mathbf{X}^*$ has closed range, then ζ is a C-point for b.*

Proof. Notice that the quantity in the definition of C-point (6.1) is nothing but $\|k_z^b\|^2$. Now from the expression for the reproducing kernel in terms of $\mathbf{X}$, we have

$$k_z^b = (I - z^*\mathbf{X}^*)^{-1} k_0^b. \tag{6.10}$$

First we show that if $I - \zeta\mathbf{X}^*$ has closed range, then its range contains k_0^b. For this it suffices to show that k_0^b is always orthogonal to the kernel of $I - \zeta^*\mathbf{X}$, or what is the same, that if $f \in \ker(I - \zeta^*\mathbf{X})$, then $f(0) = 0$. To see this, for such f we have

$$f(z) = \sum_{j=1}^{d} \zeta_j^* (X_j f)(z) \tag{6.11}$$

so in particular

$$f(0) = \sum_{j=1}^{d} \zeta_j^* (X_j f)(0). \tag{6.12}$$

Now apply $\zeta_k^* X_k$ to (6.11), sum over k, and evaluate at $z = 0$. We get

$$f(0) = \sum_{k=1}^{d} \zeta_k^* X_k f)(0) = \sum_{j,k=1}^{d} \zeta_k^* \zeta_j^* (X_k X_j f)(0). \tag{6.13}$$

Continuing in this manner, we see that for each integer $m \geq 0$ we have

$$f(0) = \sum_{|\mathbf{n}|=m} \zeta^{\mathbf{n}*} (X^{(\mathbf{n})} f)(0). \tag{6.14}$$

Using the Taylor expansion for f in terms of the X's, we conclude that for this ζ and all $0 \leq r < 1$,

$$f(r\zeta) = \sum_{m=0}^{\infty} r^m \sum_{|\mathbf{n}|=m} \zeta^{\mathbf{n}*} (X^{(\mathbf{n})} f)(0) = (1-r)^{-1} f(0). \tag{6.15}$$

But f belongs to $\mathcal{H}(b)$ and hence also to H^2_d, so it must satisfy the estimate

$$|f(z)| = o((1-|z|)^{-1}) \quad \text{as } |z| \to 1. \tag{6.16}$$

This is only possible in (6.15) if $f(0) = 0$.

So, assuming $(I - \zeta^* X)$ has closed range, we conclude that there exists a function $h \in \mathcal{H}(b)$ so that $k_0^b = (I - \zeta \mathbf{X}^*)h$. Substitute this into the expression (6.10), and let $z = r\zeta$ for $r < 1$. Then

$$k_z^b = (I - r\zeta \mathbf{X}^*)^{-1} (I - \zeta \mathbf{X}^*) h. \tag{6.17}$$

Now if T is any contractive operator, one easily checks that

$$(I - rT)^{-1}(I - T) = I - (1-r)(I - rT)^{-1} T, \tag{6.18}$$

and that $\|(I - rT)^{-1}\| = O((1-r)^{-1})$. Applying this to $T = \zeta \mathbf{X}^*$, we see from (6.17) that $\|k_z^b\|$ stays bounded as $z \to \zeta$ along a radius, and hence ζ is a C-point for b. $\square$

The last ingredient we need is the following result on the boundary behavior of bounded analytic functions in the ball, due to Rudin [22, Theorem 1.2].

Theorem 6.6. *Suppose that*

- Γ *is a nonempty open set in* $\partial \mathbb{B}^d$,
- r_j *increases to* 1 *as* $j \to \infty$,
- f *is a nonconstant holomorphic function bounded by* 1 *in* $\mathbb{B}^d$, *and* $\lim_{r\to 1} |f(r\zeta)| = 1$ *for a.e.* $\zeta \in \Gamma$.

Then Γ *has a dense* G_δ *subset* H *such that the set*

$$\{f(r_j \zeta) : j = 1, 2, 3 \ldots\} \tag{6.19}$$

is dense in the unit disk for every $\zeta \in H$.

In particular, under the conditions of this theorem we see that

$$\limsup_{r \to 1} |(D_\zeta f)(r\zeta)| = +\infty \quad \text{for every } \zeta \in H. \tag{6.20}$$

Proof of Theorem 6.3. We suppose $\zeta_0 \in \partial\mathbb{B}^d$ does not lie in the joint spectrum of $\pi(X)$ and derive a contradiction. If this were the case, then by Lemma 6.4 the element $I - \zeta_0^* \mathbf{X}$ would be invertible modulo compacts, as would $I - \zeta \mathbf{X}_0^*$, and hence there would exist an open set $\Gamma \subset \partial\mathbb{B}^d$ containing ζ_0 for which $I - \zeta \mathbf{X}^*$ was invertible modulo compacts for every $\zeta \in \Gamma$. In particular, each of the operators $I - \zeta\mathbf{X}^*$ would be Fredholm and hence have closed range. Thus by Lemma 6.5, each $\zeta \in \Gamma$ would be a C-point for b, and thus b and Γ would satisfy the hypotheses of Theorem 6.6 for any sequence $r_j \to 1$, but also $\lim_{r\to 1}(D_\zeta f)(r\zeta)$ would exist and be finite for each $\zeta \in \Gamma$. This obviously contradicts (6.20). ☐

References

[1] William Arveson. Subalgebras of C^*-algebras. III. Multivariable operator theory. *Acta Math.*, 181(2):159–228, 1998.

[2] Joseph A. Ball and Vladimir Bolotnikov. Canonical de Branges-Rovnyak model transfer-function realization for multivariable Schur-class functions. In *Hilbert spaces of analytic functions*, volume 51 of *CRM Proc. Lecture Notes*, pages 1–39. Amer. Math. Soc., Providence, RI, 2010.

[3] Joseph A. Ball, Vladimir Bolotnikov, and Quanlei Fang. Schur-class multipliers on the Fock space: de Branges-Rovnyak reproducing kernel spaces and transfer-function realizations. In *Operator theory, structured matrices, and dilations*, volume 7 of *Theta Ser. Adv. Math.*, pages 85–114. Theta, Bucharest, 2007.

[4] Joseph A. Ball and Thomas L. Kriete, III. Operator-valued Nevanlinna-Pick kernels and the functional models for contraction operators. *Integral Equations Operator Theory*, 10(1):17–61, 1987.

[5] John W. Bunce. Models for n-tuples of noncommuting operators. *J. Funct. Anal.*, 57(1):21–30, 1984.

[6] Joseph A. Cima, Alec Matheson, and William T. Ross. The Cauchy transform. In *Quadrature domains and their applications*, volume 156 of *Oper. Theory Adv. Appl.*, pages 79–111. Birkhäuser, Basel, 2005.

[7] Douglas N. Clark. One dimensional perturbations of restricted shifts. *J. Analyse Math.*, 25:169–191, 1972.

[8] Louis de Branges and James Rovnyak. *Square summable power series*. Holt, Rinehart and Winston, New York, 1966.

[9] S. W. Drury. A generalization of von Neumann's inequality to the complex ball. *Proc. Amer. Math. Soc.*, 68(3):300–304, 1978.

[10] Arthur E. Frazho. Models for noncommuting operators. *J. Funct. Anal.*, 48(1):1–11, 1982.

[11] Kunyu Guo, Junyun Hu, and Xianmin Xu. Toeplitz algebras, subnormal tuples and rigidity on reproducing $\mathbf{C}[z_1, \dots, z_d]$-modules. *J. Funct. Anal.*, 210(1):214–247, 2004.

[12] Michael T. Jury. An improved Julia-Caratheodory theorem for Schur-Agler mappings of the unit ball. `http://adsabs.harvard.edu/abs/2007arXiv0707.3423J`.

[13] Michael T. Jury. Operator-valued Herglotz kernels and functions of positive real part on the ball. *Complex Anal. Oper. Theory*, 4(2):301–317, 2010.

[14] John E. McCarthy and Mihai Putinar. Positivity aspects of the Fantappiè transform. *J. Anal. Math.*, 97:57–82, 2005.

[15] V. Müller and F.-H. Vasilescu. Standard models for some commuting multioperators. *Proc. Amer. Math. Soc.*, 117(4):979–989, 1993.

[16] N. K. Nikolskiĭ and V. I. Vasyunin. Notes on two function models. In *The Bieberbach conjecture (West Lafayette, Ind., 1985)*, volume 21 of *Math. Surveys Monogr.*, pages 113–141. Amer. Math. Soc., Providence, RI, 1986.

[17] Gelu Popescu. Isometric dilations for infinite sequences of noncommuting operators. *Trans. Amer. Math. Soc.*, 316(2):523–536, 1989.

[18] Gelu Popescu. Non-commutative disc algebras and their representations. *Proc. Amer. Math. Soc.*, 124(7):2137–2148, 1996.

[19] Gelu Popescu. Universal operator algebras associated to contractive sequences of non-commuting operators. *J. London Math. Soc. (2)*, 58(2):469–479, 1998.

[20] Gelu Popescu. Free holomorphic functions and interpolation. *Math. Ann.*, 342(1):1–30, 2008.

[21] Walter Rudin. *Function theory in the unit ball of* $\mathbf{C}^n$, volume 241 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York, 1980.

[22] Walter Rudin. *New constructions of functions holomorphic in the unit ball of* $\mathbf{C}^n$, volume 63 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1986.

[23] Donald Sarason. *Sub-Hardy Hilbert spaces in the unit disk*. University of Arkansas Lecture Notes in the Mathematical Sciences, 10. John Wiley & Sons Inc., New York, 1994. A Wiley-Interscience Publication.

[24] Orr Shalit. Operator theory and function theory in Drury-Arveson space and its quotients. `http://adsabs.harvard.edu/abs/2013arXiv1308.1081S`.

Department of Mathematics, University of Florida, Gainesville, Florida 32611-8105
E-mail address: `mjury@math.ufl.edu`